Allele-specific transcriptional regulation of IRF4 in melanocytes is mediated by chromatin looping of the intronic rs12203592 enhancer to the IRF4 promoter

Mijke Visser, Robert-Jan Palstra† and Manfred Kayser*

Department of Forensic Molecular Biology, Erasmus MC University Medical Centre Rotterdam, Wytemaweg 80, 3015 CN Rotterdam, The Netherlands

*To whom correspondence should be addressed. Tel: +31 107038073; Fax: +31 107044575; Email: m.kayser@erasmusmc.nl

Abstract

The majority of significant single-nucleotide polymorphisms (SNPs) identified with genome-wide association studies are located in non-coding regions of the genome; it is therefore possible that they are involved in transcriptional regulation of a nearby gene rather than affecting an encoded protein’s function. Previously, it was demonstrated that the SNP rs12203592, located in intron 4 of the IRF4 gene, is strongly associated with human skin pigmentation and modulates an enhancer element that controls expression of IRF4. In our study, we investigated the allele-specific effect of rs12203592 on IRF4 expression in epidermal skin samples and in melanocytic cells from donors of different skin color. We focused on the characteristics and activity of the enhancer, and on long-range chromatin interactions in melanocytic cells homozygous and heterozygous for rs12203592. We found that, irrespective of the trans-activating environment, IRF4 transcription is strongly correlated with the allelic status of rs12203592, the activity of the rs12203592 enhancer and that the chromatin features depend on the rs12203592 genotype. Furthermore, we demonstrate that the rs12203592 enhancer physically interacts with the IRF4 promoter through an allele-dependent chromatin loop, and suggest that subsequent allele-specific activation of IRF4 transcription is stabilized by another allele-specific loop from the rs12203592 enhancer to an additional regulatory element in IRF4. We conclude that the non-coding SNP rs12203592 is located in a regulatory region and affects a wide range of enhancer characteristics, resulting into modulation of the enhancer’s activity, its interaction with the IRF4 promoter and subsequent allele-specific transcription of IRF4. Our findings provide another example of a non-coding SNP affecting skin color by modulating enhancer-mediated transcriptional regulation.

Introduction

Single-nucleotide polymorphisms (SNPs) linked to specific phenotypes in genome-wide association studies (GWA studies) are often located in non-coding regions of the genome. In contrast to coding variants affecting protein function, these non-coding SNPs are more complicated to functionally characterize (1). The SNP rs12203592 is located in intron 4 of the interferon regulatory factor 4 (IRF4) gene, encoding a member of a helix-loop-helix family of DNA-binding transcription factors involved in downstream regulation of interferon signaling. IRFs are primarily associated with immune system development and response (2). IRF4 is predominantly expressed in lymphocytes, macrophages, B-cells and dendritic cells, but also in melanocytic lineages (3). Involvement of this gene has been described in several types of lymphoma and leukemia (3), including childhood acute lymphoblastic leukemia (ALL) (4), but also in non-hematopoietic diseases (3). IRF4 was initially implicated in pigmentation through the association of DNA variants located outside this gene, particularly with freckling in humans (5). In succession, the intronic SNP
rs12203592 was found to be strongly associated with hair, eye, skin color, tanning response and nevus count (6–10). This non-coding SNP was also identified to be associated with (male-specific) childhood ALL and in that study, the first evidence was given that rs12203592 is located within a regulatory element, controlling expression of IRF4 through binding of transcription factor TFAP2α (4). In Burkitt Lymphoma B-cells (Raji), HEK293T human embryonic kidney cells and in NCI-H295R human adrenal cells, the rs12203592 region containing the C-allele was shown to strongly bind the transcription factor TFAP2α and repress IRF4 promoter activity, while the rs12203592 region with the T-allele binds TFAP2α with less affinity, and has a less repressive effect on IRF4 expression (4). More recently, it was shown that in skin melanocytes, rs12203592 is located within a melanocyte-specific enhancer regulating the expression of IRF4 (11). The transcription factor TFAP2α binds to the enhancer element containing the C-allele with high affinity, and together with MITF, transcription of IRF4 is activated. When the rs12203592 T-allele is present, TFAP2α-binding is reduced, which in turn leads to a decreased expression of IRF4 (11). In addition, the lighter pigmentation phenotype associated with the rs12203592 T-allele was shown to result from impaired activation of tyrosinase (TYR) transcription due to reduced IRF4 levels (11). The difference between the two studies, describing either transcriptional repression (4) or activation (11), might be due to cell-type-specific effects mediated by transcription factors other than TFAP2α and to the fact that TFAP2α can behave as an activator as well as a repressor in different contexts (12,13).

In the present study, we extend on these previously published results (4,11) and provide detailed novel insight in the molecular mechanism of IRF4 activation by the rs12203592 enhancer. Using several molecular approaches, we investigated the transcriptional regulation of IRF4 in melanocytes, with a special focus on allelic differences of the rs12203592 enhancer. We studied the expression of IRF4 and transcription factors potentially involved in its transcriptional regulation in epidermal skin samples from donors of different skin color phenotypes and rs12203592 genotypes, in cultured skin melanocytes derived from differently pigmented skin as well as in two melanoma cell lines. We surveyed the chromatin profile of the IRF4 locus focusing at specific regulatory sites within the gene using chromatin immune-precipitation (ChIP) of H3K27Ac in combination with deep sequencing (ChIP-seq) and the use of ENCODE data, and we mapped the chromatin folding of the IRF4 locus in cultured skin melanocytes and melanoma cell lines using chromosome conformation capture (3C) techniques. We demonstrate an allele-dependent activity of the rs12203592 enhancer, and subsequent differential IRF4 expression. Furthermore, we show that the rs12203592 enhancer physically interacts with the IRF4 promoter and we provide evidence for an interaction with a potential additional regulatory element in intron 7 of IRF4, both interactions depend on the rs12203592 genotype. We propose a model for the transcriptional regulation of IRF4, in which we suggest that the activation of IRF4 transcription in melanocytes is determined by the rs12203592 enhancer/IRF4 promoter chromatin-loop, while the potential rs12203592 enhancer/intron-7 regulatory-element chromatin-loop mediates stabilization of the chromatin structure resulting into an increased and stable transcription of IRF4.

**Results**

**Expression levels of IRF4 depend on the allelic status of rs12203592 in epidermal skin samples**

To study the correlation between the rs12203592 genotype and IRF4 gene expression levels, we used a panel of 29 epidermal skin samples obtained from donors with different pigmentation phenotypes (12 dark and 17 light) and three commercially available primary Human Epidermal Melanocytes of neonatal origin derived from donors with different pigmentation phenotypes; one light pigmented donor (LP22) and two dark pigmented donors (DP74 and DP80) (see 14 for category separation of the epidermal skin samples and characterization of all samples).

DNA and RNA were co-extracted from the epidermal layers of the skin biopsy samples and mRNA levels of IRF4 were measured in the individual samples using quantitative (q) reverse transcriptase (RT)–PCR. DNA samples were used for HiRisPlex analysis, and as rs12203592 is included in this assay, the genotypes of rs12203592 were obtained from the HiRisPlex results for all epidermal skin samples (14,15). Analyzing this SNP in a world-wide set of individuals from 51 populations (HGDP-CEPH) (16) reveals that the minor-allele-frequency of this SNP is relatively low; it is present in 12% of the average European populations and it is rare in African population (0.6%) and in Asian populations (1%). Similar results are obtained when analyzing data from the 1000 Genomes project (17). This is reflected in the rs12203592 genotypes of the epidermal skin samples we used in this study; only 1 out of 29 epidermal skin samples is homozygote TT, and 3 epidermal skin samples are heterozygote CT, all 4 are of European decent (Supplementary Material, Table S1).

We observed statistically significant higher expression levels of IRF4 in the samples carrying the homozygote CC-genotype (n = 23) than in the combined sample set with the heterozygote CT- and homozygote TT-genotypes (n = 4) (Fig. 1A, P = 0.02). This confirms previous notions that transcription of IRF4 depends on the allelic status of rs12203592 (11). To exclude the possibility that the detected differential IRF4 expression between CC and CT/TT samples is caused by the dark phenotype of a large fraction (48%) of the rs12203592-CC epidermal skin samples, we analyzed the correlation between IRF4 expression and the rs12203592 genotype in the light sample set only. This analysis revealed that IRF4 is still differentially expressed between rs12203592 CC and CT/TT alleles when only the light samples are considered, although the difference is less significant (P = 0.048) (Fig. 1B), which is most probably due to the smaller sample size (n = 17). Furthermore, we observed no significant expression differences for IRF4 between the light and dark epidermal skin samples (Fig. 1C).

We next investigated the expression patterns of IRF4 in the melanocyte cell lines LP22, DP74 and DP80. DNA and RNA were co-extracted from the individual cell lines, mRNA levels were measured using RT–qPCR and DNA samples were used to obtain rs12203592 genotypes using the HiRisPlex (14,15). Interestingly, this revealed that not a lightly pigmented but the darkly pigmented melanocytic cell line DP74 is heterozygous for rs12203592 (Supplementary Material, Table S1). Despite the dark phenotype of DP74, expression of IRF4 is lower in this cell line when compared with the expression of IRF4 in LP22 and DP80 (Fig. 1D), which are both homozygote CC for rs12203592 (Supplementary Material, Table S1).

From the results with both samples sets, we conclude that IRF4 is differentially expressed depending on the allelic status of rs12203592, confirming previous observations (11) and extending on this knowledge, we demonstrate that this differential expression of IRF4 does not depend on the pigmentation phenotype of the epidermal skin samples and the melanocytic cell lines used.

**Characterization of intronic regulatory elements in IRF4**

The region around rs12203592 was shown to act as an intronic enhancer for IRF4 in melanocytes (11), and we demonstrated in
The previous paragraph demonstrates reduced IRF4 expression when the T-allele is present. (C) RT-qPCR analysis of IRF4 transcripts in all epidermal skin samples and (D) in only light epidermal skin samples with the rs12203592 CC-genotype (all: n = 25; lights: n = 13) or with the combinational rs12203592 CT- and TT-genotypes (n = 4). Together, we conclude from these observations that the differences in H3K27Ac enrichment observed in the IRF4 region are not due to technical aspects, but rather to the different rs12203592 genotypes.

Accessibility of chromatin is another important characteristic of enhancer elements—regions that are active, are generally more open and accessible for protein complexes like the transcription machinery to bind—and this can be detected by several techniques. DHS mapping detects regions of chromatin that are more accessible for the DNA digesting activity of the DNase I enzyme. In a publicly available data set for DHS signals obtained from epidermal skin melanocytes, such a prominent DHS signal is present at the rs12203592 enhancer (Fig. 2B). To investigate this in our melanocytic cell system, we used a commercially available assay that directly measures the accessibility of chromatin to nucleases (EpiQ-assay; BioRad). This revealed that the chromatin at the rs12203592 enhancer is more open and accessible in the cell lines with the CC-genotype (LP22 and DP80) than in the cell line with the CT-genotype (DP74). Data are represented as mean ± SEM; *p < 0.05.
IRF4 promoter and one in intron 7 of IRF4. Notably, both YY1 peaks coincide with robust DHS peaks and H3K27Ac-signals in the melanocytic cell lines LP22 and DP74 (Fig. 2C). Although the H3K27Ac peak in intron 7 is only moderate, it differs in amplitude between the two cell lines. Furthermore, this DHS peak in intron 7 resembles the tissue-specificity of the DHS in intron 4; it is present in melanocytes and absent in the majority of the other cell types tested (Supplementary Material, Fig. S2). However, in contrast to the DHS signal in intron 4, the DHS peak in intron 7 is not present in melanoma cells (Supplementary Material, Fig. S2). Based on the observed features of this small conserved region in intron 7 that is associated with active enhancers (18), we suggest that this region

**Figure 2.** Identification of three regulatory elements in the IRF4 locus; the IRF4 promoter, the rs12203592 enhancer and a potential regulatory element in intron 7. (A) IGV genome browser (24) shows a 39 kb window of the IRF4 locus. To investigate the chromatin for features of enhancer elements, the following tracks are included: ChIP-seq analysis in LP22 and DP74 of acetylated histone H3 (H3K27Ac), an active chromatin mark (Palstra et al., manuscript in preparation); DHSs in epidermal skin melanocytes (20); ChIP-seq data for the transcription factor MITF in melanocytic cells (21); ChIP-seq data in MALME-3M melanoma cells for the transcription factor YY1 (22); and Phastcons conserved elements inferred from 46 way alignments of placental mammals (23). IGV genome browser shows (B) a zoomed-in frame of the region around the rs12203592 enhancer and (C) a zoomed-in frame of the region around a potential regulatory element in intron 7. The rs12203592 enhancer displays apparent characteristics of an active enhancer; strong H3K27Ac signals that are differential between LP22 and DP74, a robust DHS peak is detected and binding of the transcription factors MITF and YY1 is observed. Binding of YY1 is also observed in the potential regulatory region in intron 7, as well as a robust DHS peak, the H3K27Ac signals are modest, but differential between LP22 and DP74.
represents a potential additional regulatory element for IRF4 in skin melanocytes.

The detection of three regions that harbor chromatin modifications specific for regulatory elements, including binding of the transcription factor YY1, demonstrate that the IRF4 locus contains three regulatory elements; its promoter, the region around rs12203592 and an additional (potential) regulatory region at intron 7 (from here onwards referred to as the ‘intron-7 YY1-element’). Moreover, the region around rs12203592 acts as a (melanocyte-specific) enhancer, and the activity of this enhancer depends on the allelic status of the rs12203592 SNP.

The rs12203592 enhancer interacts with the IRF4 promoter

Regulatory elements, in particular enhancers, control transcription of their target gene by physically interacting with the gene promoter via chromatin loops (25), and it has been reported that multiple distal enhancers can regulate a target gene simultaneously or cooperatively (26). Furthermore, cell-type-specific transcription factors, such as MITF or YY1 not only determine the activity of the enhancer, but they are also involved in formation of the chromatin loops, either by recruiting co-factors involved in loop formation or by directly mediating loop formation (25). It has been suggested that the multifunctional transcription factor YY1 directs long-range chromatin interactions (27) and it was further demonstrated that YY1 physically interacts and co-localizes with, for example, cohesin and other proteins known to be involved in chromatin looping (27). We therefore reasoned that YY1 binding to the IRF4 promoter, the rs12203592 enhancer and the intron-7 YY1 element could mediate loop formation between these elements. In order to study this, we used 3C technology (28); this method enables detection of long-range chromatin interactions by trapping the interactions between chromatin segments that are in close proximity with formaldehyde. The cross-linked chromatin is subsequently digested using a restriction enzyme followed by intermolecular ligation under dilute conditions. With qPCR, the relative abundance of ligation products is determined, and this is proportional to the frequency with which the different restriction fragments interact. Locus-wide 3C analysis using three different Apol restriction fragments containing the IRF4 promoter (P), the rs12203592 enhancer (E) and the additional YY1 element (Y), respectively, as anchor point in the melanocytic cell line LP22 (Fig. 4A–C) revealed robust interactions between the IRF4 promoter and the rs12203592 enhancer. Interestingly, interactions

![Figure 3. The region around rs12203592 displays features of an allele-dependent enhancer. (A) EpiQ analysis in the melanocyte cell lines reveals that region around rs12203592 is more open and accessible in LP22 (rs12203592-CC) and DP80 (rs12203592-CC), than in DP74 (rs12203592-CT) (*P < 0.001). (B) Luciferase reporter assay demonstrates that the activity of the rs12203592 enhancer is specific for the melanocyte lineage and differential between the rs12203592 CC-allele and the TT-allele (**P < 0.05). Data are represented as mean ± S.E.M.](image1)

![Figure 4. A long-range chromatin loop is formed between the rs12203592 enhancer and the IRF4 promoter in skin melanocytes. Locus-wide cross-linking frequencies observed in LP22 cells (rs12203592-CC). The analyzed region of the IRF4 locus is shown on the top of each graph, the x-axis depicts the location of the fragments analyzed; C1–C4, control regions 1–4; P, IRF4 promoter; E, rs12203592 enhancer; Y, intron-7 YY1-element. Black shading, the position of the ‘fixed’ restriction fragment (the anchor point), gray-shaded dotted lines, position of the other restriction fragments analyzed. Cross-linking frequencies are shown for a restriction fragment containing (A) the IRF4 promoter, (B) the rs12203592 enhancer and (C) the intron-7 YY1-element. Robust interactions are observed between the IRF4 promoter and the rs12203592 enhancer (A) and (B), only modest interactions are observed between the rs12203592 enhancer and the intron-7 YY1-element (B) and (C), while no interaction is detected between the IRF4 promoter and the intron-7 YY1-element (A) and (C). Data are represented as mean ± S.E.M.](image2)
between the rs12203592 enhancer and downstream located restriction fragments, which include the intron-7 YY1-element, remained high, while the interaction between the IRF4 promoter and these downstream restriction fragments appeared to be low. Locus-wide 3C analysis in another light-pigmented melanocytic cell line (LP89; rs12203592-CC) revealed similar interaction patterns (data not shown), in agreement with the data obtained with LP22.

To gain further insight into the effect of the chromatin loops in IRF4, we studied two melanoma cell lines BLM and G361. Characterization of BLM and G361 using HiRisPlex analysis (19) predicted light pigmentation phenotypes for both melanoma cell lines (data not shown). Furthermore, both melanoma cell lines were homozygote CC for rs12203592 (Supplementary Material, Table S1). Recently, a whole-transcriptome RNA-seq study described reduced IRF4 expression in melanoma cells when compared with that in skin melanocytes (29). Indeed, we detected borderline statistically significantly reduced mRNA levels of IRF4 in the G361 cells when compared with that in the melanocytic cell lines also carrying the rs12203592 CC-genotype (P = 0.056), while the expression of IRF4 is fully absent in the BLM melanoma cell line (Fig. 1D).

We next inquired how this on/off difference of IRF4 expression in the melanoma cell lines is reflected by the formation of chromatin loops in the IRF4 locus. Locus-wide 3C analysis using the three different Apol restriction fragments containing the IRF4 promoter, the rs12203592 enhancer and the additional YY1 element, respectively, as anchor point in the melanoma cell lines G361 and BLM revealed that in G361 cells, a chromatin loop is formed between the rs12203592 enhancer and the IRF4 promoter, while in BLM cells, this chromatin loop is absent (Fig. 5A and B), which agrees with the difference in IRF4 expression observed between G361 and BLM cells. However, in contrast to what we observed in our melanocyte cell lines, the interaction between the rs12203592 enhancer and the intron-7 YY1 element containing downstream restriction fragments appeared to be low in both G361 and BLM cells (Fig. 5B and C), which in combination with the DNaseI Hypersensitivity data (Supplementary Material, Fig. S2) suggests that a melanocyte-specific interaction between the rs12203592 enhancer and the intron-7 YY1 element exists.

From these data, we conclude that the rs12203592 enhancer regulates IRF4 expression by directly contacting the IRF4 promoter by formation of a chromatin loop. Furthermore, a possible interaction was detected between the rs12203592 enhancer and the intron-7 YY1 element in skin melanocytes; remarkably, the presence of this interaction was not detected in the melanoma cell lines possibly reflecting the absence of DHS signals in this region.

Differential IRF4 expression does not correlate with the availability of key transacting factors

Differential availability of (tissue-specific) TFs in the different samples, rather than modulation of TF binding at the rs12203592 enhancer, could in principle also explain the differential expression of IRF4 we and others observed. Using online-available ChIP-seq data sets for the transcription factors MITF (21) and YY1 (22), we observed binding signals for both TFs in the rs12203592 enhancer region, which confirms previous data (11). Furthermore, it was previously demonstrated that the transcription factor TFAP2α binds to this region, and it has been suggested that the binding sequence of TFAP2α is altered by rs12203592 (4,11). We therefore studied the expression of TFAP2α, MITF and YY1 in relation to the pigmentation phenotypes as well as to the rs12203592 genotypes in our samples sets. In the epidermal skin samples, we observed no correlation between the expression patterns of TFAP2α or MITF and the rs12203592 genotype (Supplementary Material, Figs S3A and S4A) or the pigmentation phenotype (Supplementary Material, Figs S3B and S4B). Furthermore, we also did not detect an effect of the expression patterns of YY1 on the transcription levels of IRF4, even though...
YY1 expression correlated with the pigmentation phenotype of the samples (Supplementary Material, Figs S5A and B). In the melanocytic cell lines, the expression levels of MITF and YY1 correlated with the pigmentation phenotypes; however, this correlation was not reflected by either the rs12203592 genotypes of the cell lines or the expression levels of IRF4 (Supplementary Material, Figs S4C and S5C). Furthermore, the expression levels of TFAP2α did not correlate with either the pigmentation phenotypes or the rs12203592 genotypes of the cell lines (Supplementary Material, Fig. S3C). In the G361 cells, MITF expression was statistically significantly up-regulated (P < 0.005), while IRF4 expression was reduced when compared with the average expression of MITF and IRF4 in the melanocytic cell lines with similar rs12203592 genotypes (Supplementary Material, Fig. S4C and Fig. 1D). Consistent with the previous findings (29), MITF expression was dramatically reduced (Supplementary Material, Fig. S4C) and YY1 expression was strongly elevated (Supplementary Material, Fig. S5C) in BLM cells in which IRF4 expression was fully absent, when compared with the expression of MITF and YY1 (P < 0.001 and <0.04) in melanocytic cell lines.

Based on the obtained expression analyses in the epidermal skin samples, the melanocytic cells and the melanoma cell lines, we conclude that the expression levels and subsequent availability at the enhancer region of the transcription factors TFAP2α, MITF and YY1 do not explain the observed differences in activity of the rs12203592 enhancer and transcription levels of IRF4 in our samples.

Allele-specific characterization of the intronic rs12203592 enhancer in IRF4

Activity of an enhancer is mediated by a broad variety of factors, both ubiquitous and cell-type specific, that modulate the specific features of the enhancer, such as modifications of histones, accessibility of the chromatin and loop formation. We have demonstrated that the availability of the three transacting factors that are presumably involved in the transcriptional regulation of IRF4 does not correlate with the allele-dependent differences in IRF4 expression. However, it remains possible that additional unknown factors modulate the transcriptional regulation of IRF4 instead of the allele differences of rs12203592. In principle, if a signal, such as H3K27Ac enrichment or chromatin accessibility, is dependent on the allelic status of an SNP, this should be detectable using an allele-specific approach in a cell-line heterozygous for this SNP. In contrast, if the signal is independent of the allelic status of an SNP, the allelic balance of the DNA fragments that give rise to the signal would be equal in a heterozygote cell line for this SNP since the entire cellular environment including potentially involved trans-acting factors are the same for both alleles. In order to test this hypothesis, we used the heterozygote melanocytic cell line DP74 to detect allele-specific primary IRF4 transcripts, and H3K27Ac, FAIRE and EpiQ signals at the rs12203592 enhancer. We PCR-amplified the region around rs12203592 that was captured with the different assays and sequenced the amplicons. Subsequently, the balance between the fragments with either the rs12203592 T-allele or the rs12203592 C-allele was determined. As expected, the C/T ratio in the control samples (ChIP input; without H3K27Ac-enrichment and undigested EpiQ samples) was ∼50/50 (Fig. 6A and B). Conversely, in the H3K27Ac-ChIP sample, the C/T ratio was 79/21 (Fig. 6A and Supplementary Material, Fig. S6A), indicating that the C-allele is more enriched for H3K27Ac than the T-allele. In contrast, in the EpiQ sample, the C/T ratio was 7/93, which implies that the T-allele is protected from digestion indicating a closed chromatin conformation around the rs12203592 enhancer (Fig. 6B and Supplementary Material, Fig. S6B). Likewise, FAIRE analysis followed by sequencing analysis to determine the allelic ratio of rs12203592 confirmed that the C-allele has an open chromatin configuration, while the T-allele is inaccessible (Fig. 6C and Supplementary Material, Fig. S6C).

Since the rs12203592 enhancer is located in intron 4 of the IRF4 gene, it is transcribed as part of the primary transcript. The C/T balance of rs12203592 in the primary transcripts in DP74 cells indicates that the majority of the RNA molecules are transcribed from the C-allele (Fig. 6D and Supplementary Material, Fig. S6D), which is in agreement with the notion that the C-allele of the rs12203592 enhancer is more active than the T-allele. We next investigated whether these allelic dis-balances are also reflected in differences in loop formation between the regulatory elements. For this, we prepared 3C samples from DP74 cells, used the rs12203592 enhancer as anchor point and sequenced the 3C PCR-products of five distinct captured regions in the IRF4 locus. For all restriction fragments analyzed, including the control regions (C1, C3 and C4), the C/T ratio was higher than the allelic ratio of 1. Importantly however, for the fragments that contain the IRF4 promoter and the intron-7 YY1-element, this ratio was shifted toward 2.5 (Fig. 6E and Supplementary Material, Fig. S6E). Our allele-specific 3C analysis thus demonstrates that the C-allele of rs12203592 is more actively involved in chromatin looping, and the chromatin loops are preferentially formed with the IRF4 promoter and the intron-7 YY1-element.

Taken together, we demonstrated that the characteristics of the rs12203592 enhancer strongly depend on the allelic status of rs12203592; with the C-allele present, the region around rs12203592 is highly accessible, it is enriched for the active enhancer mark H3K27Ac, and the chromatin configuration, while the T-allele is inaccessible (Fig. 6C and Supplementary Material, Fig. S6C). Since the rs12203592 enhancer is located in intron 4 of the IRF4 gene, it is transcribed as part of the primary transcript. The C/T balance of rs12203592 in the primary transcripts in DP74 cells indicates that the majority of the RNA molecules are transcribed from the C-allele (Fig. 6D and Supplementary Material, Fig. S6D), which is in agreement with the notion that the C-allele of the rs12203592 enhancer is more active than the T-allele. We next investigated whether these allelic dis-balances are also reflected in differences in loop formation between the regulatory elements. For this, we prepared 3C samples from DP74 cells, used the rs12203592 enhancer as anchor point and sequenced the 3C PCR-products of five distinct captured regions in the IRF4 locus. For all restriction fragments analyzed, including the control regions (C1, C3 and C4), the C/T ratio was higher than the allelic ratio of 1. Importantly however, for the fragments that contain the IRF4 promoter and the intron-7 YY1-element, this ratio was shifted toward 2.5 (Fig. 6E and Supplementary Material, Fig. S6E). Our allele-specific 3C analysis thus demonstrates that the C-allele of rs12203592 is more actively involved in chromatin looping, and the chromatin loops are preferentially formed with the IRF4 promoter and the intron-7 YY1-element.

Taking together, we demonstrated that the characteristics of the rs12203592 enhancer strongly depend on the allelic status of rs12203592; with the C-allele present, the region around rs12203592 is highly accessible, it is enriched for the active enhancer mark H3K27Ac, and the chromatin configuration, while the T-allele is inaccessible (Fig. 6C and Supplementary Material, Fig. S6C). Since the rs12203592 enhancer is located in intron 4 of the IRF4 gene, it is transcribed as part of the primary transcript. The C/T balance of rs12203592 in the primary transcripts in DP74 cells indicates that the majority of the RNA molecules are transcribed from the C-allele (Fig. 6D and Supplementary Material, Fig. S6D), which is in agreement with the notion that the C-allele of the rs12203592 enhancer is more active than the T-allele. We next investigated whether these allelic dis-balances are also reflected in differences in loop formation between the regulatory elements. For this, we prepared 3C samples from DP74 cells, used the rs12203592 enhancer as anchor point and sequenced the 3C PCR-products of five distinct captured regions in the IRF4 locus. For all restriction fragments analyzed, including the control regions (C1, C3 and C4), the C/T ratio was higher than the allelic ratio of 1. Importantly however, for the fragments that contain the IRF4 promoter and the intron-7 YY1-element, this ratio was shifted toward 2.5 (Fig. 6E and Supplementary Material, Fig. S6E). Our allele-specific 3C analysis thus demonstrates that the C-allele of rs12203592 is more actively involved in chromatin looping, and the chromatin loops are preferentially formed with the IRF4 promoter and the intron-7 YY1-element.

Discussion

Enhancers are DNA elements involved in transcriptional regulation, which contain short DNA motifs that act as binding sites for sequence-specific transcription factors (25). Upon binding of specific transcription factors to the enhancer, co-activators and/or co-repressors are recruited. This combination of cues resulting from factors either directly or indirectly interacting with the enhancer determines the activity of that enhancer (18) and as such, controls transcription of the target gene through physical contacts with the core promoter of that gene (30). The enhancer can act on a gene even if the distance between the enhancer and the promoter is relatively large. It has been shown in various studies that this physical distance is overcome by the formation of long-range chromatin loops that bring regulatory elements like enhancers and promoters in close proximity with each other (31).

Several studies showed that non-coding SNPs that were identified through GWAS studies to be highly associated with a certain phenotype can be located in enhancer elements regulating transcription of a target gene, and the activities of those enhancers are modulated depending on the allelic status of the trait-associated SNPs (14,32–35). Similarly, the SNP rs12203592 in intron 4
of IRF4, which had been significantly associated with human pigmentation phenotypes (5–10), was recently described to modulate an enhancer controlling expression of IRF4 (4,11). In the present study, we further investigated the transcriptional regulation of IRF4 in detail and extend on the previously published knowledge regarding the rs12203592 enhancer (4,11). We found that transcription of IRF4 not only correlates with the rs12203592 genotype in skin melanocytic cell lines but importantly also in a large set of epidermal skin samples. We demonstrated that the region around rs12203592 displays features of an enhancer element, and that these features strongly depend on the allelic status of rs12203592. Furthermore, we show that the rs12203592 enhancer physically interacts with the IRF4 promoter via a long-range chromatin loop. Our data, especially from the allele-specific analysis, suggest that this rs12203592 enhancer also interacts in an allele-dependent manner with an additional regulatory element in intron 7 of IRF4, thereby folding the IRF4 locus into an intricate structure. Based on previously published studies and the data obtained in the present study, we propose a model for the transcriptional regulation of IRF4 in melanocytes. With the rs12203592 C-allele present, the transcription factor TFAP2α binds to the enhancer in intron 4 (11), which allows for the recruitment of MITF and YY1, formation of a chromatin loop to the IRF4 promoter and potentially to the YY1 element in intron 7, resulting into proper and stable transcription of IRF4. With the rs12203592 T-allele, TFAP2α is unable to bind (11), which leads to reduced recruitment of MITF and YY1, less stable chromatin loops and consequently, decreased IRF4 expression. Furthermore, the observed interaction differences of the rs12203592 enhancer with the IRF4 promoter and the YY1 element in intron 7 might be due to a shift in equilibrium between the unfolded and interacting state that depends on the rs12203592 alleles (Fig. 7).

It is likely that other transcription factors are additionally involved in the mediation and stabilization of the chromatin loops and subsequent transcriptional regulation of IRF4. Analysis of transcription factor binding [using online available TF-binding prediction software, ‘Promo’ (http://alggen.lsi.upc.es/cgi-bin/promo_v3/promo/promoinit.cgi?dirDB=TF_8.3) (36)] predicts LEF1 binding at both regulatory regions in IRF4, as well as in the IRF4 promoter region (Supplementary Material, Fig. S7A–C), suggesting that LEF1 might also be involved in the loop formation and transcriptional activation of IRF4 (Fig. 7). LEF1 has been
shown to be involved in chromatin looping (37) and to interact with MITF (38), resulting into transcriptional activation of two other human pigmentation genes DCT (38) and OCA2 (32).

The interaction of the rs12203592 enhancer with the potential regulatory element in intron 7 appears not to be crucial for IRF4 activation, as we found this interaction to be almost absent in G361 melanoma cells, while IRF4 was still transcribed. However, IRF4 expression is reduced in this cell line when compared with that in melanocytic cell lines with the same rs12203592 genotype (CC), suggesting that this interaction is probably necessary for...
stabilization of the chromatin structure, resulting into proper and stable IRF4 expression. The biological reason why this interaction with the regulatory region in intron 7 is lost in G361 cells and the exact mode-of-action of this potential enhancer element remain to be elucidated.

We recently demonstrated that it is valuable to not only focus on the associated SNP identified with a GWA study, but also to study the SNPs in linkage disequilibrium (LD) in order to discover, for example, the actual causal and functional variant underlying the studied associated phenotype [14]. Although it appears to be straightforward in the present study that rs12203592 and not one of its LD partners is the functional SNP, we considered the possibility of an LD SNP to contribute to the transcriptional regulation of IRF4. Not many SNPs are in LD with IRF4, and of those only one, in relatively low LD (LD = 0.242) is potentially interesting, as this SNP (rs3778607) is located at the border of the YY1 binding site in intron 7 of IRF4. It may be that this SNP modulates the activity of the additional regulatory element by altering the binding sequence of YY1, thereby contributing to the (de)stabilization of the chromatin structure and subsequent transcriptional regulation of IRF4, which should be investigated in future studies.

In the BLM melanoma cells, we detected a linear chromatin organization in the IRF4 locus and that IRF4 expression was totally absent. Notably, in this cell line, the master regulator of melanocyte development MITF is silenced, probably caused by and/or resulting into a complex cascade of (silencing) events, that eventually leads to a completely deregulated melanomatous landscape (Fig. 7) and overall gene expression profile, which are typical features of metastatic cancer cells [39]. The exact mechanism behind the loss of the IRF4 chromatin loops, and whether the loss of these loops indeed directly results into a fully inactive gene, has to be further investigated, but it seems likely that the loss of MITF expression in these cells plays a crucial role.

Genes often display expression patterns that are highly cell-type specific, ranging from fully silenced in one cell type to highly expressed in another. In order to achieve these differential expression patterns, transcription also needs to be regulated in a cell-type specific manner. As part of the IRF gene family, IRF4 is, besides its well-established association with pigmentation phenotypes [5–10], mainly involved in immune system development and response [3] and rs12203592 was shown to be associated with (male-specific) childhood ALL [4]. This study is the first to demonstrate that rs1203592 is located in an element that regulates IRF4 expression, however with a repressive activity. With the rs12203592 C-allele present, the enhancer has a repressing effect on IRF4 expression in three different, non-melanocytic cell lines, rather than the activating effect we and others [11] detected in skin melanocytic cells. This difference is probably due to the transcription factors involved in the activity of the rs12203592 enhancer, with TFAP2a as common transcription factor shown to be involved in all cell types tested and MITF being highly specific for skin melanocytes. Binding of this transcription factor and recruitment of other pigmentation-related transcription factors like YY1 can drive the activity of the enhancer in a direction more favorable for transcription when compared with that in cells in which MITF is absent and other cell-type-specific transcription factors, possibly with repressive functions are recruited.

The IRF4 protein encoded by the IRF4 gene is also a transcription factor, which is thought to play a crucial role in the development of lymphoid cells. Since this gene is also expressed in melanocytic cells, it is probable that IRF4 is also involved in pigmentation biology or melanocyte development. Indeed, it was shown recently [11] that expression of TYR, an essential enzyme in melanogenesis, is regulated by cooperation between MITF and IRF4, which eventually leads to differences in pigmentation. As a consequence, the differences in TYR expression might therefore explain the strong association of rs12203592 with human skin color variation. Although it was not the scope of our study, we investigated the expression patterns of TYR in our sample sets and found that the expression of TYR indeed seems to correlate, although not statistically significantly, with the rs12203592 genotype in the epidermal skin samples (data not shown).

We did not observe a correlation between pigmentation phenotype and rs12203592 genotype in our skin biopsy sample set; however, we believe that this is most probably due to the small sample size of only 27 individuals rather than the complete absence of the correlation. The association of rs12203592 with human pigmentation phenotypes was repeatedly shown in several independent studies [6–10], that used thousands of individuals as needed for obtaining reliable outcomes in GWA studies of complex traits such as pigmentation. Additionally, a recent multivariate analysis including nine SNPs from nine genes that are strongly associated with skin color variation, demonstrated that the T-allele of rs12203592 has a significant effect on lighter skin color (F. Liu et al., manuscript in preparation). Owing to a stronger effect of other SNPs on skin color (such as rs12913832 in HERC2 and rs4268748 in MC1R), it is however unlikely to detect a significant effect of rs12203592 in our small sample set consisting of only four rs12203592 CT/TT-allele samples.

Since the rise of GWA studies and the emerging data resulting therefrom, several studies have been conducted to functionally analyze the (strongly) associated DNA variants. Several of those located in non-coding regions either intergenic or intronic were demonstrated to sit in enhancer elements regulating transcription of a nearby gene. Recently, the intronic SNP rs12203592 which is highly associated with skin pigmentation was shown to be located in and modulate the activity of an enhancer controlling the gene IRF4 [11]. Here, we demonstrate that this SNP not only modulates transcription factor binding and subsequent enhancer activity, but that this enhancer also physically interacts with the IRF4 gene promoter in an allele-specific manner and with another potential regulatory element to stabilize the overall chromatin structure in order to achieve stable and cell-type-specific expression of the target gene IRF4.

Materials and Methods
Skin sample collection and epidermal layer separation
Collection and processing of the skin samples were done as described previously [14], in short; non-sun-exposed skin material that was left-over from patients undergoing plastic surgery or dermatological surgery were obtained under informed consent and patients were asked to fill in a questionnaire to obtain information concerning the pigmentation of their skin, eye and hair and more basic information as age, sex and bio-geographic ancestry. Skin color reflectance was measured at the inner side of the upper arm with spectrophotometer (Konica Minolta CM-600d), the obtained L, a and b values were used to calculate the individual typology angles (ITA) [40,41] in order to categorize the skin samples according to the ITA separation of light (above 41°) and dark (under 28°) [14]. This study was approved by the Medical Ethical Committee of the Erasmus MC (number MEC-2012-067).

Upon receiving the skin samples, they were cleaned and cut into small pieces, and stored in RNA later at −4°C until further processing. The epidermal layer of the skin, including the
pigment-containing melanocytic layer, was removed from the dermal layer by incubating the skin samples cut into small pieces of ±1 mm wide and 7 mm long for 30 min at room temperature in Ammonium Thiocyanate solution (3.8% ATC in PBS) (42). The epidermal layers were then peeled off carefully and stored in RNA later at 4°C until RNA/DNA co-isolation, or at –80°C for long-term storage.

**Cell culture**

HEMn-LP22 (C-0025C, lot # 200708522; Cascade Biolgies, Invitrogen), HEMn-DP74 (C-2025C, lot # 6C0447; Cascade Biologies, Invitrogen) and HEMn-DP80 (C-2025C, lot # 20070980 Cascade Biologies, Invitrogen) were grown in Medium 254 supplemented with HMGS according to the manufacturer’s instructions (Cascade Biologies, Invitrogen). G361, HEK293 and BLM cells were cultured in DMEM/10% FCS at 37°C/5% CO2.

**RNA/DNA co-isolation from skin samples and from cultured cells**

Epidermal skin tissue was lysed and RNA and DNA were co-isolated with the Qiagen Allprep mini kit according to the manufacturer’s instructions for fibrous tissue. From the different cell lines, total cellular RNA and genomic DNA were isolated with TriPure Isolation Reagent according to the manufacturer’s protocol (Roche Diagnostics). DNA and RNA samples were column purified to remove PCR inhibiting substances like melanin (OneStep™ PCR Inhibitor Removal Kit, Zymo Research Corp) and subsequent DNasel digestion of the RNA samples was performed using Ambion’s Turbo DNA-free kit (Applied Biosystems) following the instructions of the manufacturers.

**Transcription analysis**

The RT reaction was performed using RevertAid™ H Minus First Strand cDNA Synthesis Kit (Fermentas GmbH) according to the manufacturer’s instructions. Quantitative real-time PCR reactions for gene expression analysis were performed using the iTaq Universal SYBR Green Supermix (Bio-Rad Laboratories) with the following PCR parameters: initial denaturation at 95°C for 5 min, followed by 45 cycles at 95°C for 1 s and 60°C for 30 s, followed by a melting curve analysis. Bio-Rad Software v1.5 in combination with Excel was used to analyze the qPCR data. The reference gene ACTB was used to normalize the amplification signal between samples, differences in treatment and amount of input cDNA. Primers are available on request. Student’s two-tailed t-test was used to determine statistical significance.

**SNP genotyping**

The HiPlex assay and the interactive HiPlex prediction tool were used as described by (15) to predict hair and eye color for the cell line samples, to confirm the obtained eye and hair color phenotype information of all skin biopsy samples and, additionally, to obtain genotypes of the SNP rs12203592. previously (43). Antibody used: acetylated Histone H3 K27 (#4729ab) from Abcam.

**Epiq analysis**

Q-PCR-based EpiQ chromatin analysis assay (Bio-Rad Laboratories) was performed according to the manufacturer’s protocol. To remove melanin, DNA was column purified prior to PCR (OneStep™ PCR Inhibitor Removal Kit, Zymo Research). Quantitative real-time PCR was performed using Universal SYBR green master mix (Bio-Rad Laboratories) under the following cycling conditions: 95°C for 5 min, 45 cycles of 10 s at 95°C, 30 s at 60°C and followed by a melting curve analysis. PCR primers are available on request.

** Luciferase assays**

Custom500 bp gBlocks Gene Fragments (Integrated DNA Technologies) containing either the C- or T-allele of the rs12203592 enhancer region and Xhol and BglIII restriction sites were cloned into the Xhol/BglIII sites of a modified pGL3-promoter vector (in which the SV40 promoter and the SV40 3'UTR were replaced by an HSP promoter and an HSP 3'UTR) (Promega). Inserts in each construct were verified by sequencing (Basesclear). Constructs were transfected into HEK293 or G361 melanoma cells using Lipofectamine LTX (Invitrogen), and luciferase expression was normalized to Renilla luciferase expression. Primer sequences are available on request. Data represent at least five independent experiments. Student’s two-tailed t-test was used to determine statistical significance.

**Chromosome conformation capture analysis**

3C analysis was performed essentially as described (44,45) using Apol as the restriction enzyme. To remove melanin, DNA was column purified prior to PCR (OneStep™ PCR Inhibitor Removal Kit, Zymo research). Quantitative real-time PCR (CFX96TM Real Time System, BioRad) was performed using iTaq SYBR Green Supermix with ROX (BioRad), under the following cycling conditions: 50°C for 2 min, 95°C for 10 min, 45 cycles of 15 s at 95°C, 1 min at 60°C, followed by a melting curve analysis. Random template was generated from a 975 bp custom-designed gBlocks Gene Fragments (Integrated DNA Technologies). gBlock were verifived as described in the Millipore protocol (http://www.millipore.com/userguides/tech1/mcproto407), except that samples were cross-linked with 2% formaldehyde for 10 min at room temperature. To remove melanin, DNA was column purified prior to PCR (OneStep™ PCR Inhibitor Removal Kit, Zymo Research). ChIP-seq analysis was performed as described

**Allele-specific analysis**

For allele-specific analysis, respective templates (cDNA, gDNA and DNA enriched in FAIRE, EpiQ and ChIP experiments) were amplified with primers covering the rs12203592 SNP using the Expand Long Template PCR system (Roche). Amplicons were gel purified and sequenced (Basesclear). Allelic ratios were calculated from the peak plot files. PCR primers are available on request.

**Supplementary Material**

Supplementary Material is available at HMG online.
Acknowledgements

We are very thankful to all volunteers who agreed their materials to be used for this study. We acknowledge the staff, especially Jan Maerten Smit and Han van Neck, from the Department of Plastic Surgery, and Tamar Nijsten, Leonie Jacobs and Emmilia Dowlatshahi from the Department of Dermatology, both of Erasmus MC, who helped with collecting epidermal skin samples. We additionally thank the staff of the Erasmus Center of Biomics, especially Wilfred van IJcken and Rutger Brouwer, for sequencing and data-processing of the Chip-seq samples. Lakshmi Chaitanya and Arwin Ralf from the Department of Forensic Molecular Biology of Erasmus MC are acknowledged for help with genotyping.

Conflict of Interest statement. None declared.

Funding

This study was supported by the Erasmus MC, and in part by a grant from the Netherlands Genomics Initiative (NGI)/Netherlands Organization for Scientific Research (NWO) within the framework of the Forensic Genomics Consortium Netherlands (FGCN).

References
