The *CDKN2A/p16*^{INK4a} 5'UTR sequence and translational regulation: impact of novel variants predisposing to melanoma

Virginia Andreotti[1,2,°], Alessandra Bisio[3,°], Brigitte Bressac- de Paillerets[4], Mark Harland[5], Odile Cabaret[4], Julia Newton-Bishop[5], Lorenza Pastorino[1,2], William Bruno[1,2], Roberto Bertorelli[6] ,Veronica De Sanctis[6], Alessandro Provenzani[7], Chiara Menin[8], Gilberto Fronza[9], Paola Queirolo[10], Robert C. Spitale[11], Giovanna Bianchi-Scarrà[1,2], Alberto Inga[3, #], Paola Ghiorzo[1,2, #*]

[1]Department of Internal Medicine and Medical Specialties, DiMI, University of Genoa, Genoa, 16132, Italy; [2]Genetics of Rare Cancers, IRCCS AOU San Martino-IST, Genoa, 16132, Italy; [3]Laboratory of Transcriptional Networks, Centre for Integrative Biology, CIBIO, University of Trento, Trento, 38123, Italy; [4]Department of Biopathology and INSERM U1186, Gustave Roussy, Villejuif, F-94805, France; [5]Section of Epidemiology and Biostatistics, Leeds Institute of Cancer and Pathology, University of Leeds, LS9 7TF , UK; [6]NGS Core Facility, Centre for Integrative Biology, CIBIO, University of Trento, Trento, 38123, Italy; [7]Laboratory of Genomic Screening, Centre for Integrative Biology, CIBIO, University of Trento, Trento, 38123, Italy;[8]Immunology and Molecular Oncology Unit, Veneto Institute of Oncology IOV-IRCCS, Padua, 35128, Italy; [9]Mutagenesis Unit, IRCCS AOU San Martino -IST, Genoa, 16132, Italy; [10]Medical Oncology Unit, IRCCS AOU San Martino-IST, Genoa, 16132, Italy; [11]Department of Pharmaceutical Sciences, University of California, Irvine. CA 92697-1075.

° The first two authors equally contributed and should be regarded as joint First Authors

#The last two authors jointly directed this work and should be regarded as joint Last Authors

*To whom correspondence should be addressed (fax: +390103538887, tel.: +390103538949, email: paola.ghiorzo@unige.it)

Summary

Many variants of uncertain functional significance in cancer susceptibility genes lie in regulatory regions, and clarifying their association with disease risk poses significant challenges. We studied 17 germline variants (9 of which were novel) in the *CDKN2A* 5'UTR with independent approaches, which included mono and bicistronic reporter assays, western blot of endogenous protein, and allelic representation after polysomal profiling to investigate their impact on *CDKN2A* mRNA translation regulation. Two of the novel variants (c.-27del23,c.-93-91delAGG) were classified as causal mutations (score ≥3), along with the c.-21C>T,c.-34G>T and c.-56G>T, which had already been studied by a subset of assays. The novel c.-42T>A as well as the previously described c.-67G>C were classified as potential mutations (score 1 or 2). The remaining variants (c.-14C>T,c.-20A>G, c.-25C>T+c.-180G>A,c.-30G>A,c.-40C>T,c.-45G>A,c.-59C>G,c.-87T>A,c.-252A>T) were classified as neutral (score 0). In conclusion, we found evidence that nearly half of the variants found in this region had a negative impact on *CDKN2A* mRNA translation, supporting the hypothesis that 5'UTR can act as a cellular Internal Ribosome Entry Site (IRES) to modulate $p16^{INK4a}$ translation.

Significance

The aim of this study was to develop assays to clarify the pathogenicity of 17 5'UTR *CDKN2A* variants found during routine screening in melanoma patients and at the same time give insights into *CDKN2A* mRNA translation regulation. To suggest a functional classification useful to inform genetic counseling we assigned a score based on concordance among the different tests. We propose that the 5'UTR of *CDKN2A* should be included in routine screening of melanoma patients as we found evidence that nearly half of the variants found in this region had a negative impact on the post-transcriptional dynamics of the *CDKN2A* mRNA.

Keywords melanoma susceptibility; germline mutation; 5' untranslated region (5'UTR); variants with unknown functional significance; reporter assays; polysomal imbalance; *CDKN2A*

Running title: New CDKN2A 5'UTR variants predisposing to melanoma

Total word count: 7932

Introduction

An increasing number of variants of uncertain functional significance (VUS) are emerging from genome-wide association (GWAS) and next generation sequencing (NGS) studies in cancer genetics. Many of these VUS lie in regulatory regions, and clarifying their association with disease risk poses significant challenges (Edwards et al., 2013).

In the specific case of melanoma, to date it has been possible to interpret the significance of a small number of VUS that lie in the regulatory regions of high or low susceptibility genes. A recent study that performed linkage analysis and high throughput sequencing in a melanoma-prone family identified a variant in the promoter of the telomerase reverse transcriptase gene (TERT) associated with the generation of consensus motifs for transcription factors and increased transcriptional activity (Horn et al., 2013), which indicates that TERT may be a novel, although rare, high penetrance susceptibility gene for melanoma. Another recent report described a functional single nucleotide polymorphism (SNP) in the promoter region of the gene encoding the β1 immunoproteasome catalytic subunit (PSBM9), which was identified through mining of GWAS data in the search for low penetrance melanoma risk alleles. This promising SNP was predicted to be located at putative transcription factor binding sites and associated with increased gene expression through electrophoretic mobility shift assay (EMSA) and luciferase assays and mRNA quantitation in preliminary functional studies (Qian et al., 2013). At a somatic level, analysis of whole genome sequencing data from malignant melanomas revealed two telomerase reverse transcriptase (TERT) gene promoter mutations, indicative of ultraviolet (UV) light-induced damage, in 17 of 19 (89%) cases examined; these mutations generated *de novo* consensus binding motifs for Ets transcription factors, and increased transcriptional activity from the TERT promoter by 2 to 4-fold, similarly to the germline variant (Huang et al., 2013). In addition, another somatic mutation in the 5'UTR of ribosomal gene RPS27, at the major transcription start site of the gene, has also been recently described as a recurrent event in melanoma. The mutation expands the 5'TOP element, a motif known to regulate the expression of most of the ribosomal protein family (including RPS27) (Dutton-Regester et al., 2014), which is required for the translational control of the protein in a growth-dependent manner.

The results of these studies emphasize that germline as well as somatic mutations in regulatory regions of the genome may represent an important tumorigenic mechanism. Therefore, identifying the causal variation associated with risk Single Nucleotide Variants (SNVs) alleles is crucial to establish the mechanistic link between the genotype and the phenotype, and indicate that the function of SNVs in regulatory regions in high as well as low risk genes needs to be further explored.

Functional tests, *in vitro* analyses, and combined approaches for determining the causality of variants in the coding region have been developed and are continuously optimized for the major melanoma susceptibility

gene, (Jenkins et al., 2013; McKenzie et al., 2009, 2010; Scaini et al., 2009, 2014) *CDKN2A* (cyclin-dependent kinase inhibitor 2A), which is found to be germline mutated in approximately 40% of melanoma families worldwide and in Italy (Bruno et al., 2009; Ghiorzo et al., 1999, 2006; Goldstein et al., 2006). The *CDKN2A* locus regulation is particularly complex, due to the use of alternative promoters, splicing and reading frames of shared coding regions for encoding two structurally unrelated proteins involved in cell cycle regulation, namely $p16^{INK4a}$ and $p14^{ARF}$. Loss of $p16^{INK4a}$ leads to inactivation of the Rb pathway while inactivation of $p14^{ARF}$ affects the p53 pathway; thus, alteration of this single locus is a highly efficient means of compromising two important tumor suppressor pathways (Aoude et al., 2015). It's still not clear whether mutations affecting both proteins have a stronger effect or predispose to more than one cancer (*i.e.*, melanoma and pancreatic cancer) (Goldstein et al., 2007). The high intrinsic mutation rate of the locus amplifies this effect: mutations, deletions and promoter methylation are common in a variety of human tumors (Hall and Peters, 1996; Liggett and Sidransky, 1998; Rocco and Sidransky, 2001) and make this locus one of the most frequently mutated loci in human cancers (Kandoth et al., 2013).

The regulatory regions of *CDKN2A/p16*^{INK4a}, including the promoter and the 5'UTR, 3'UTR and intronic regions, have also been investigated, and a causal germline mutation in the 5'UTR at nucleotide -34 (c.-34G>T), which gives rise to an alternative translating initiation codon able to decrease translation from the wild-type ATG (Liu et al., 1999) has been identified, as well as three deep intronic mutations (Harland et al., 2001, 2005). Recently, one novel intronic variant (IVS2+82C>T) of the *CDKN2A* gene and a small deletion (c.−20677−20682delGTACGC) in its promoter region were found (Veinalde et al., 2013). These findings support the hypothesis that variants in the regulatory regions of *CDKN2A* may be involved in melanoma predisposition, but their rarity and the lack of tools to assess their causality limit our ability to interpret them.

We previously studied a panel of 7 variants in the 5'UTR from the Italian melanoma population using specific reporter assays combined with polysomal profiling. Because a subset of these variants had a negative impact on the post-transcriptional dynamics of the *CDKN2A/p16*^{INK4a} mRNA, we concluded that they might have clinical significance (Bisio et al., 2010).

Building on this background, the present study evaluated a panel of 9 additional novel *CDKN2A/p16*^{INK4a} 5'UTR variants found during routine testing of *CDKN2A* in Italian as well as in English and French melanoma patients by different groups belonging to GenoMEL, the International Melanoma Genetics Consortium. We further developed tools to assess the 5'UTR variants functional significance and enhance mechanistic insights into the regulation of *CDKN2A*. Other 8 variants previously identified but not completely characterized were also investigated more in-depth, for a total of 17 germline variants studied in *CDKN2A* 5'UTR.

We identified a subset of variants which showed an impact at the level of the translational efficiency of the associated mRNA both in gene reporter assays and at the level of the endogenous p16 transcript, and thus

gained interesting insights into p16 mRNA translation regulation, as well as into clinical use of these variants in genetic counseling.

Results

Novel germline *CDKN2A/p16*^{INK4a} 5'UTR variants in melanoma patients

Following an initial evaluation of 5'UTR variants in Italian melanoma cases and controls that was conducted within the framework of a multi-center study with another Italian group (Veneto Institute of Oncology of Padua) and with the GenoMEL groups from the Gustave Roussy in Villejuif, France and the University of Leeds, UK, we identified 9 novel variants in the 5'UTR of *CDKN2A/p16*^{INK4a} after surveying 5889 melanoma patients (probands of melanoma families and hospital-based sporadic cases) and 1709 controls overall. The localization and frequency of the novel 5'UTR variants is shown in Figure 1A and Table 1, together with the other 5'UTR variants previously identified and analyzed in this study (c.-252A>T, c.-180G>A + c.-25C>T, c.-67G>C, c.-56G>T, c.-34G>T, c.-21C>T, c.-14C>T). All of the variants, except the c.-67G>C, were identified in melanoma patients and not in controls. The c.-93-91delAGG, c.-87T>A , c.-59C>G, c.-27del23 were found for the first time in probands from melanoma families, so it was possible to verify cosegregation among affected members. The c.-93-91delAGG variant was found in a French proband and cosegregated in one affected sister (Figure 1B). The c.-87T>A variant was found in a French proband and in only one of the two affected sons, so cosegregation was not complete (Figure 1C). The c.-59C>G variant was found in two unrelated French patients, one MPM non-familial and one familial. Interestingly, the variant did not cosegregate in the affected relatives of the familial patient, while the CDKN2A p.Gly101Trp coding region mutation, one of the most frequent CDKN2A mutations observed in melanoma families, did, so cosegregation could be excluded (Figure 1D,E). The c.-27del23 variant was found in two unrelated probands, one from France, the second from Italy. Cosegregation analysis was not possible for this variant (Figure 1F,G).

The c.-45G>A, c.-42T>A c.-40C>T, c.-30G>A, c.-20A>G, c.-14C>T variants were found in French, English and Italian sporadic melanoma patients, most of which were affected by MPM (c.-14 C>T, c.-40C>T and c.-42T>A). Interestingly, the c.-14C>T variant was also found in two unrelated English sporadic SPM patients and it had previously been described in one MPM from the Swedish population. (Hashemi et al., 2000) The novel c.-42T>A variant was found in one patient who developed two melanomas and renal cancer. The patient carried *in cis* the c.-25C>T & the c.-180G>A variants on the other allele (Figure 1H). We previously identified the variants at position -25 and -180 in one Italian sporadic melanoma patients. (Bisio et al., 2010) Another one was identified in this study in an Italian sporadic SPM patient. The c.-40C>T variant was found in a French MPM patient (Figure 1I) while the c.-20A>G and c.-30G>A were identified in two sporadic English

patients. Interestingly, the c.-21C>T variant, which was previously described to cosegregate in an Italian family, was found in a French proband affected by four melanomas, with a family history of neuroblastoma, ovarian and breast cancer (Figure 1J), and also in a sporadic English SPM patient, so that, overall, three cases have been identified to date with this candidate mutation (Bisio et al., 2010).

The c.-34A>G mutation, which is common among English melanoma families (Liu et al., 1999), was also observed in one sporadic English patient and in one Italian melanoma family in the present study.

The c.-56G>T and c.-67G>C were previously described in Italian patients (Bisio et al., 2010). Finally, we also included the c.-252A>T variant, which we did not observe but has been previously identified in the Australian population, showing no evidence on cosegregation (Pollock et al., 2001) (Figure 1A).

Overall the observed novel variants are rare in melanoma cases, as few carriers were observed after surveying 5889 melanoma cases, cosegregation analysis was often not possible due to lack of family history or testing information in affected probands, and none of these variants were found in 1709 controls with the exception of c.-67G>C found in one sample. Hence, only functional analyses could help in interpreting the potential clinical significance of these noncoding variations at the *CDKN2A* 5'UTR.

Functional analysis of *CDKN2A/p16 5'UTR* variants using gene reporter assays: impact on cap-independent translation

Since we previously observed that the impact of the p16 5'UTR alleles was mainly at the post transcriptional level, we specifically developed two plasmid-based assays, using a modified version of the pGL3 promoter and pRuF bicistronic dual luciferase reporter vectors (Supplemental Methods).

The pGL3-promoter vector contains the wild type or variant 5'UTR cloned just downstream of the viral promoter, and is thus expected to act as the 5'UTR of the firefly luciferase reporter whose activity is determined. In the pRuF bicistronic vector the wild type or variant 5'UTR is cloned as an intervening sequence between the Renilla reniformis and the Firefly luciferase, which are transcribed as part of a single transcript. Expression of the second cistron documents the ability of the inserted sequence to promote internal ribosome loading and translation independent from the first cistron. Firefly luciferase activity obtained with this construct is thus suggestive of the presence of an IRES-like function in the intervening sequence.

With each of these constructs we transfected three melanoma cell lines that are null for $p16^{INK4a}$ protein expression (WM266-4, G361, SKMel-5). Dual luciferase assays were conducted after transient transfection, and, to assess the potential impact of each variant, the relative activities of the two luciferase enzymes in the presence of wild type or variant 5'UTRs were compared. Each cell line showed distinct transfection efficiency and

plasmid expression levels, but the relative impacts of the different 5'UTR alleles compared to the wild type were very consistent. Several of the variants analyzed were *in cis* with the common polymorphism at nt -191 (c.-191 A>G). We first verified that the two -191 alleles showed no differences in reporter assays with the promoter and the pRuF vectors (Supplemental Data).

The post-transcriptional impact of the wild type 5'UTR was confirmed using the three melanoma cell lines with the promoter-derived constructs (Table S1, Table 2, Figure S1). We also confirmed that the c.-21C>T and c.-34G>T variant had a severe impact on luciferase activity (mean, 51 and 81% of reduction in activity compared to the wild type, respectively). This was further confirmed using the two new modified plasmids (Supplemental Methods) while we observed that all the other new variants had a weaker impact (mean, up to 30% of reduced activity compared to the wild type). Only the c.-30G>A variant had a positive impact compared to wild type; the effect was stronger in WM266-4, compared to G361 and SKMel5.

When we tested the wild type and variant pRuF constructs in the three melanoma cell lines, we found that the wild type 5'UTR strongly stimulated the Firefly luciferase indicating that the whole 5'UTR sequence stimulated ribosome binding through a cap-independent translation (Figure 2). Furthermore, in the three cell-lines we found that the c.-21C>T and c.-34G>T variants had a negative impact (mean, 53% and 57% of the wild-type, respectively), as expected, as did c.-27del23 and c.-93-91delAGG (mean, 49% and 46% of the wild-type). The reduction resulting from these two variants is comparable to that observed for the mutation at nt -34 and the candidate mutation we previously characterized at nt -21 (Tables S1 and 2).

The remaining variants only showed a slightly reduced or increased activity compared to the wild type in the three cell lines. Specifically, c.-14C>T, c.-20A>G, c.-45G>A, c.-59C>G showed a mean reduced activity of up to 18% compared to the wild type, while c.-40C>T, c.-42T>A, c.-87T>A, c.-252A>T showed a mean increased activity of up to 19%. However, the c.-56G>T and c.-67G>C variants previously studied only in one melanoma cell line (Bisio et al., 2010), were here tested in the three melanoma cell lines and showed a mean reduced activity of 40% and 53%, respectively (Figure S2) .

The variant at c.-42T>A deserved further attention. As described in Figure 1, this variant was found *in trans* with c. -25 C>T+c.-180G>A. In a separate study, we are showing that this variant can affect a YBX1 binding site (Bisio et al., 2015). The c.-42T>A variant alone does not reduce 5'UTR activity in the pRuF construct. The -180 variant *in cis* with the -25 has been previously shown to have no impact as a single variant, as confirmed here (Figure S1).

In order to rule out that the differences observed in our reporter assays were caused by different transcription rates or by an alternative splicing events, we performed qualitative and quantitative RT-PCR

starting from RNAs from the cell lines transfected with wild type or variant pGL3 promoter constructs. As previously described (Bisio et al., 2010), comparable mRNA expression levels were observed for the wild type and the transfected variant constructs (Figure S4) and differently sized RNAs, to be ascribed at alternative splicing events, were not observed (data not shown). Overall, when we compared the luciferase protein activities to the amounts of luciferase mRNA to establish whether the impact of the 5'UTR variants occurred at a transcriptional or post-transcriptional level, or both, we again confirmed that the impact was mainly post-transcriptional, supporting the hypothesis that 5'UTR can act as a cellular IRES.

Impact of *CDKN2A/p16*^{INK4a} 5'UTR variants on p16^{INK4a} endogenous levels

To evaluate the effect of *CDKN2A/p16* 5'UTR variants on the expression of endogenous p16, mRNA and protein levels were determined taking advantage of the available melanoma patients-derived LCLs harboring the aforementioned 5'UTR variants at heterozygous level. A reduction in p16 mRNA relative amounts for few LCLs was apparent (4/14, c.-14C>T, c.-27del23, c.-40C>T and c.-45G>A). Direct sequencing confirmed bi-allelic expression (data not shown). Surprisingly, a cell line with homozygous A/A polymorphism at position -191 showed to a considerable decrease in p16 mRNA level (around 50%) compared to a homozygous G/G or an heterozygous A/G line (Figure 3A). However, p16 protein levels were more comparable among the three c.-191 cell lines. The c.-14C>T, c.-27del23, c.-40C>T 5'UTR variants showed lower p16 mRNA and even lower protein levels. Notably, the c.-20A>G, c.-21C>T, c.-34G>T, c.-42T>A, c.-93-91delAGG variants showed comparable relative amounts of mRNA to their c.-191 matched wild-type *CDKN2A* 5'UTR (G/G or A/G, none of the variants was A/A), but a decrease in p16 protein amount, indicating a negative impact of these 5'UTR variants at post-transcriptional level (Figure 3B, 3C and 3D). Instead the cell line of the c.-59G>C and in part also that of c.-45G>A showed higher p16 protein expression (Table2).

Allelic imbalance analyses quantify the impact of p16^{INK4a} 5'UTR variants on polysomal loading

To better understand the impact of the *CDKN2A/p16* 5'UTR variants on p16^{INK4a} mRNA translation efficiency, polysomal profiling was performed on RNA extracted from heterozygous LCLs followed by RNA-sequencing using Ion Torrent PGM. In this approach, an imbalance at polysomal levels would indicate an impact of 5'UTR sequence variants on the efficiency of polysomal loading. For this analysis we selected two of the most negatively impacting 5'UTR variants in the gene reporter assays (c.-27del23 and c.-93-91delAGG). The c.-34G>T variants that generated an upstream open reading frame was included as a control (Figure 4). While for the c.-191A/G polymorphism both alleles were equally represented both on the subpolysomal and polysomal fractions, with a slight preference for the A allele in the latter. The c.-27del23, c.-34G>T and c.-93-91delAGG showed imbalance also at subpolysomal level, suggesting an impact of the variants also at the level of mRNA stability. For the c.-34G>T variants this can be due to nonsense mediated decay, as the variant generates an

upstream AUG whose utilization leads to premature truncation (Liu et al., 1999). The imbalance was higher in the polysomal fractions and in all cases the variant alleles were less abundant compared to the wild type.

For eight additional *CDKN2A/p16* 5'UTR variants (c.-14C>T, c.-27del23, c. -25 C>T+c.-180G>A, c.-30G>A, c.-42T>A, c.-45G>A, c.-59C>G and c.-252A>T) we estimated allelic imbalance through the analysis of Sanger sequencing electropherograms (Bisio et al., 2010) (Figure S3). When LCLs were not available, we used MCF7 cells and transfected equal amounts of pGL3 promoter reporter plasmids containing wild-type and variant 5'UTR (c.-30G>A and c.-252A>T) to mimic the heterozygous state in the patients. Results showed higher polysomal loading for the c.-14C>T allele while for the remaining variants we detected a slight imbalance in favor of the wild type allele that was more evident for the c.-42A>T variant.

Discussion

Functional activity of germline variations in the CDKN2A 5'UTR and promoter have been investigated leading to the identification of the first causative mutation for melanoma at position -34 (c.-34G>T), which gives rise to an alternative translation initiating codon able to decrease translation from the wild-type AUG, and probably arose from a common founder in the United Kingdom (Harland et al., 2000; Liu et al., 1999). Since then, rare variants at the *CDKN2A* 5'UTR, encountered during routine screening are usually defined as having unknown significance once their frequency in a control population has been determined and cosegregation analysis has been performed in the family, when possible.

We previously developed functional assays to study the clinical significance of 7 heterozygous germline variants at the *CDKN2A* 5'UTR which can be applied in general to the study of 5'UTR sequence changes of uncertain significance that are identified during mutation screening of disease genes. That study led to the identification of a melanoma predisposing mutation at nt -21 altering p16 mRNA translation regulation, and the classification of variants at nt-191 and -33 as common polymorphisms and functionally neutrals. Interpretation for other variants could not be exhaustive.

The aim of this study was to develop further functional assays to clarify the pathogenicity even for those variants which could not be equivocally classified and at the same time give insights into p16 mRNA translation regulation, as the first experiments using bicistronic transfection vectors were suggestive of a mechanism of action of 5'UTR sequence variants relying on the activity of this 5'UTR in cap-independent translation.

To do this we extended the analysis to a number of novel germline variations in the 5'UTR, and re-analyzed those previously described for a total of 17 variants, with a more exhaustive approach which included reporter assays with mono and bicistronic reports, western blot to assess endogenous protein expression, and polysomal profiling to clarify p16 mRNA translation regulation. Relative polysomal loading of mRNAs is

considered a proxy for their translation efficiency (Zaccara et al., 2014). The latter approach enabled us to exploit patients-derived surrogate cell models that harbors the *CDKN2A* 5'UTR variants at the endogenous gene and in the heterozygous state. The possibility to accurately measure the relative allelic representation on polysomes by parallel sequencing and compare it with cytoplasmic mRNAs fractions not associated with polysomes, can provide information on the impact of the 5'UTR variants on mRNA translation. To this aim we developed a protocol to quantify allelic imbalance at high resolution using Ion Torrent PGM and conditions that generated a 300-1000x coverage. While the parallel sequencing approach is more sensitive compared to the quantification of Sanger sequencing electropherograms, it still requires the capacity to perform sucrose gradient fractionation for detecting allelic imbalance at polysomal level, which limits its application to bioptic samples. Although relatively few variants were examined, there was a good agreement between the results with the parallel sequencing assay and the gene reporter assays with the bicistronic reporter, which also has single allele resolution, and also with western blot analysis normalized to total mRNA expression (Figure 3 and Table 2). One limitation of this approach is that p16 steady-state protein levels in LCLs could also be in part influenced by the EBV-immortalization protocol and they reflect the expression of both variant and wild-type alleles. Our recent functional and structural results with the wild type p16 5'UTR confirm that this sequence can act as an internal ribosome entry site (Bisio et al, 2015), which justify the emphasis on bicistronic reporter assays and polysomal profiling to characterize the impact of p16 5'UTR variants.

To assess overall pathogenicity as established by these three or more independent approaches, and determine the functional impact of each variant, we tried to assign a score to each one based on concordance among the different assays.

Variants were classified as neutral (score 0) when no difference compared to wild type in at least 3 assays was observed. This applied to variants c.-14C>T, c.-20A>G, c.-25C>T+c.-180G>A, c.-30G>A, c.-40C>T, c.-45G>A, c.-59C>G, c.-87T>A, c.-252A>T.

When variants displayed less than 60% of wild type activity in one or two assays, or less than 60% of protein expression, or positive imbalance ($\geq$ 10%) (score 1-2), we classified the variants as potential mutations. This applied to variants c.-42T>A and c.-67G>C. The interpretation of c.-42T>A variant is complicated by the compound heterozygosis with the c.-180G>A+ c.-25C>T allele. The separate analysis of the two alleles in gene reporter assays indicates that both alleles are neutral. However, testing mRNAs from the LCL, the c.-42A>T alleles was less represented in polysomes and p16 protein levels were reduced in western blot analysis. Hence, we propose to classify c.-42A>T as potential mutation.

Finally, we classified variants as causal mutations (score 3 or more), when three or more assays showed impairment. This applied to the variants, c.-27del23, c.-56G>T, c.-93-91delAGG, as well as to the c.-21C>T and c.-34G>T, which had already showed impairment.

Compared to our previous analysis, by extending the number of assays and repeating transfection experiments in three metastatic melanoma cell lines, we were able to finally classify variants at nt c.-25+-180 as neutral, the variant at nt.-67 as potential mutation and the variant at nt c.-56 as causal.

In addition we were able to suggest a classification for novel 5'UTR variants to inform genetic counseling of melanoma patients and family relatives carrier (Table 2).

Overall 7 out of 17 variants (41%) were classified as potential or causal mutation, suggesting that any rare variant detected in the 5'UTR of *CDKN2A* during routine screening should be followed with attention and functionally tested. This study to our knowledge is the largest in identifying and pooling novel germline variants in the 5'UTR of *CDKN2A*. Seventeen of these variants were found in 25 probands out of 5889 cases. Only one was found in one control but classified as potential mutation. A much more extensive survey based on the ExAC browser of exome sequencing data of at least 11 thousand alleles from unrelated European non-Finnish individuals, sequenced as part of various disease-specific and population genetic studies, identified the variants c.-14C>T, c.-25A>G and c.-34G>A at frequencies ranging from 0.01 to 0.05% Data from exome sequencing is however limited to the first 40 nucleotides of the *CDKN2A* 5'UTR.

In conclusion, we standardized a panel of reporter assays and allelic analysis from patients-derived SNV heterozygous cells and developed a classification score to classify *CDKN2A* 5'UTR SNVs. Although we tested all the variants with pGL3 promoter and pRuf bicistronic constructs in three melanoma cell lines, we observed that the bicistronic construct, that is tailored on investigating translational control, was more informative and consistent with the impairment observed with other assays (as resulting from the score).The relative impact of the different 5'UTR alleles compared to the wild type was very consistent in the three melanoma cell lines tested. Based on these results, we suggest a pipeline for the functional analysis that could be applied to characterize any new variant identified in the *CDKN2A* 5'UTR, starting from gene reporter assays based on bicistronic vectors in one melanoma cell line (WM266-4, which shows the best transfection efficiency). Results could be integrated with polysomal profiling and p16 protein expression studies, assays which could be used in clinical practice in lymphoblastoid cell lines only after additional standardization and development of miniaturized protocols.

As an additional future development and improvement of this functional analysis workflow a CRISPR/Cas9 knock-in strategy can be considered to introduce desired sequence variants at the endogenous $p16^{INK4a}$ locus in a p16 wild type melanoma or melanocyte-derived cell lines. This approach would open up the possibility to test the impact of the sequence variants in a more relevant cell type and at the level of the endogenous gene level, using both polysomal profiling, inhibitors of cap-independent translation and p16 protein analysis.

Overall, our data indicate that the sequencing of 5'UTR of *CDKN2A* should be included in routine screening of melanoma families and that the sequenced region should be extended from current set-up of exome sequencing to cover the entire 5'UTR region. We found evidence that nearly half of SNVs identified in the *CDKN2A/p16*^{INK4a} 5'UTR in melanoma patients can have a negative impact on the capacity of this noncoding sequence to influence mRNA translation acting as a cellular IRES.

Methods

Study population

After our initial evaluation of *CDKN2A* 5'UTR variants in the Italian population, (Bisio et al., 2010) we surveyed additional melanoma cases and controls from Italy (Genoa and Padua) and individuals belonging to two case control- studies from France and England, accounting to a total of 5889 melanoma cases and 1709 controls. A subset of these patients belonged to study populations that were previously enriched for cases with a possible genetic susceptibility to melanoma, and tested for mutations in *CDKN2A*. Detailed descriptions of these populations have been reported elsewhere. (Bisio et al., 2010; Ghiorzo et al., 2012, 2013; Harland et al., 2000; Kannengiesser et al., 2009; Maubec et al., 2012; Menin et al., 2011; Newton Bishop et al., 2000; Newton-Bishop et al., 2010; Pastorino et al., 2008) Briefly, the Italian sample included a cohort from Genoa and one from Padua; the Genoa cohort included the cases and controls that were previously analyzed for presence of variants in the *CDKN2A* 5'UTR and described (Bisio et al., 2010) as well as 150 additional melanoma cases and 150 additional controls, accounting to 1350 cases and 300 controls overall. The Padua cohort, which was also surveyed for *CDKN2A* 5'UTR variants, included 257 cases and 206 controls. Both cohorts belonged to a case control study, and included both single primary melanoma (SPM) patients, probands from melanoma families (at least two cases of melanoma in the family), patients with multiple primary melanoma (MPM), and patients with SPM and another type of cancer. The English study was comprised of 1585 cases and 499 controls from a case-control study of sporadic melanoma. The French study included 2697 cases: 1017 sporadic MPM patients, 1680 probands from melanoma families (at least two cases of melanoma in the family or one melanoma and a pancreatic cancer/or brain tumor among first degree relatives), and 404 controls. All of the controls were individuals with no history of cancer who were recruited in each case-control study. All diagnoses were confirmed either by histology or medical records.

All of the patients and all of the control individuals signed an informed consent form, provided a blood sample and answered a questionnaire with the aid of a trained interviewer, under local Ethics Committee-approved protocols.

CDKN2A mutation analysis

The *CDKN2A* (NM000077) coding region, including splice junctions, the 5' and 3'UTRs, the intronic sequence described to contain the IVS2-105 A/G mutation, and exon 1B were entirely sequenced, as well as *CDK4* exon 2. The primers and procedures for sequencing have been previously described (Ghiorzo et al., 2007; Harland et al., 2000; Maubec et al., 2012; Menin et al., 2011). The 5'UTR/promoter region was screened up to 150 base pairs upstream of the ATG for all the cases and controls, while a subset (857 cases and 506 controls) was screened up to the 420th base upstream.

Plasmids and transfections

pGL3-promoter and pRuF bicistronic vectors containing the whole wild-type or variant *CDKN2A* 5'UTR were prepared as previously described, (Bisio et al.,2010) with slight modifications (Supplemental Methods). The WM266-4, G361 and SK-Mel-5 human metastatic melanoma cell lines were transfected as previously described (Bisio et al., 2010) using a small volume protocol (Andreotti et al., 2011). The average relative light units and the standard deviations of at least three independent biological repeats were reported. The experiments were repeated at least four times for each cell line.

Relative quantification of luciferase mRNA by RT-qPCR in WM266-4 cells transfected with 5'UTR variants.

Total RNA and genomic DNA were extracted for relative mRNA luciferase quantification as previously described (Bisio et al.,2010). Part of the cells harvested after transfection were used to verify whether the experiment was consistent with the luciferase assay. Potential variations due to transfection efficiency were taken into account by correcting for plasmid copy number using three biological repeats. The experiments were repeated two times and expressed as luciferase mRNA/copy number with the standard deviations of at least three independent biological repeats.

Relative quantification of p16 mRNA and protein in LCLs by RT-qPCR and Western Blot respectively

qPCR measurements of endogenous *CDKN2A/p16*^{INK4a} mRNA from total RNA obtained from melanoma patient-derived LCLs, wild-type or harboring the indicated 5'UTR variants at heterozygous levels was performed as described above GAPDH and B2M served as reference genes. Bars are normalized on wild-type and are the average of three biological replicates.

LCLs were harvested with standard methods and 40μg of protein extracts were loaded into 12% PolyAcrylamide gels. SDS-PAGE were performed as previously described (Ghiorzo et al., 2006), proteins were transferred on nitrocellulose membranes using the semi-dry i-Blot transfer system (InVitrogen, Life Technologies) and specific protein levels detected using primary antibodies diluted in 1% non-fat skim milk in

PBS-T. The endogenous levels of p16 (N-20, Santa Cruz) were tested. Beta-Tubulin was used as reference protein. To reveal immune-reactive bands we used ECL select reagent (Amersham, GE-Health Care, Milan, Italy) acquiring images by ChemiDoc XRS+ documentation system and quantifying bands using the ImageLab software (BioRad, Milan, Italy).

Polysomal RNA extraction

Polysome preparations were obtained from patient-derived LCLs as previoualy described (Bisio et al., 2010; Bisio et al., Oncotarget in press). For each lymphoblast cell sample, cDNAs were prepared starting from RNA preparations corresponding to subpolysomal and polysomal fractions using the AffinityScript cDNA Synthesis Kit (Stratagene).

Polysomal RNA profiling for allelic imbalance

The relative abundance of the two *CDKN2A/p16*^{INK4a} 5'UTR alleles at the level of the endogenous gene was then assessed by double strand DNA Sanger sequencing on two to three independent subpolysomal/polysomal RNA preparations and quantification of the electropherograms was conducted using the Image J software. Average signal intensities for each dideoxynucleotide in the electropherograms were taken into account in the comparison of individual sequencing runs. To further control for technical sources of allele imbalance in the sequencing runs, we also compared the relative signal intensity of the same nucleotide type found at the heterozygous mutation site at specific positions surrounding the position of the mutation. For a few variants (c.-191A>G, c.-93_91delAGG, c.-34A>G, c.-27del23), the relative abundance of the two 5'UTR alleles was evaluated by Next Generation Sequencing using with an Ion PGM™ System (Life Technologies). Following amplicon purification, we quantified the unamplified library by qPCR with the Ion Library Quantitation Kit and determined to avoid other amplification steps. The template preparation was carried out applying the Ion One Touch 200 template kit v2 DL (Life Technologies) as described in the appropriate user guide provided by the manufacturer (Publication Number MAN0006957).

Sequencing of the amplicon libraries was carried out on the Ion Torrent Personal Genome Machine (PGM) system using the Ion Sequencing 200 v2 kit (Life Technologies) on a 314 chip.

Data were analyzed using Torrent Suite™ Software, reviewing, automatically, the quality and accuracy of the sequencing run. In addition, to perform further analyses such as variant calling, we used the Torrent Variant Caller plug-in, which has been optimized to detect single nucleotide polymorphisms (SNPs) and insertions and deletions (indels) from Ion Torrent™ sequencing data. The BAM files generated by Torrent Suite™ Software were loaded on the Integrative Genomics Viewer (IGV) and the allelic imbalance manually reviewed. One variant (c.-27del23), was tested by both Sanger and NGS, obtaining very comparable results.

Funding

This work was supported by: University of Genoa PRA Ateneo2013 to PG, IRCCS AOU San Martino-IST 5 per mille per la Ricerca Corrente to GBS and PG, Italian Association for Cancer Research AIRC (IG#12869 to AI and IG#15460 to PG). A grant support from Institut National du Cancer (INCA) was attributed to B B-de P. for coordination of Melanoma Oncogenetics in France.

Acknowledgments

We thank Dr. Chiara Baldo (Galliera Genetic Bank—Network of Telethon Genetic Biobanks (Project No. GTB12001, Telethon Italy) for providing lymphoblastoid cell lines and Michele Nicastro for assistance in qPCR and WB analysis with those cell lines. We wish to thank Christophe Blondel for technical assistance in CDKN2A genotyping; Marie-Françoise Avril and French clinicians network for patients referral; Gustave Roussy Biobank (BB-0033-00074) in providing DNA resources.

Conflict of Interest Statement

None declared

References

Andreotti, V., Ciribilli, Y., Monti, P., Bisio, A., Lion, M., Jordan, J., Fronza, G., Menichini, P., Resnick, M. a., and Inga, A. (2011). P53 transactivation and the impact of mutations, cofactors and small molecules using a simplified yeast-based screening system. PLoS One *6*.

Aoude, L.G., Wadt, K.A.W., Pritchard, A.L., and Hayward, N.K. (2015). Genetics of familial melanoma: 20 years after CDKN2A. Pigment Cell Melanoma Res. *28*, 148–160.

Bisio, A., Nasti, S., Jordan, J.J., Gargiulo, S., Pastorino, L., Provenzani, A., Quattrone, A., Queirolo, P., Bianchi-Scarrà, G., Ghiorzo, P., et al. (2010). Functional analysis of CDKN2A/p16INK4a 5'-UTR variants predisposing to melanoma. Hum. Mol. Genet. *19*, 1479–1491.

Bisio, A., Latorre, E., Andreotti, V., Paillerets, B.B., Harland, M., Scarra, G.B., Ghiorzo, P., Spitale, R.C., Provenzani, A., Inga, A. (2015). The 5'-untranslated region of p16INK4a melanoma tumor suppressor acts as a

cellular IRES, controlling mRNA translation under hypoxia through YBX1 binding. Oncotarget. Oct 17. doi: 10.18632/oncotarget.5387 [Epub ahead of print]

Bruno, W., Ghiorzo, P., Battistuzzi, L., Ascierto, P.A., Barile, M., Gargiulo, S., Gensini, F., Gliori, S., Guida, M., Lombardo, M., et al. (2009). Clinical genetic testing for familial melanoma in Italy: A cooperative study. J. Am. Acad. Dermatol. *61*, 775–782.

Dutton-Regester, K., Gartner, J.J., Emmanuel, R., Qutob, N., Davies, M. a, Gershenwald, J.E., Robinson, W., Robinson, S., Rosenberg, S. a, Scolyer, R. a, et al. (2014). A highly recurrent RPS27 5'UTR mutation in melanoma. Oncotarget *5*, 2912–2917.

Edwards, S.L., Beesley, J., French, J.D., and Dunning, M. (2013). Beyond GWASs: Illuminating the dark road from association to function. Am. J. Hum. Genet. *93*, 779–797.

Ghiorzo, P., Ciotti, P., Mantelli, M., Heouaine, A., Queirolo, P., Rainero, M.L., Ferrari, C., Santi, P.L., De Marchi, R., Farris, A., et al. (1999). Characterization of Ligurian melanoma families and risk of occurrence of other neoplasia. Int. J. Cancer *83*, 441–448.

Ghiorzo, P., Gargiulo, S., Pastorino, L., Nasti, S., Cusano, R., Bruno, W., Gliori, S., Sertoli, M.R., Burroni, A., Savarino, V., et al. (2006). Impact of E27X, a novel CDKN2A germ line mutation, on p16 and p14ARF expression in Italian melanoma families displaying pancreatic cancer and neuroblastoma. Hum. Mol. Genet. *15*, 2682–2689.

Ghiorzo, P., Gargiulo, S., Nasti, S., Pastorino, L., Battistuzzi, L., Bruno, W., Bonelli, L., Taveggia, P., Pugliese, V., Borgonovo, G., et al. (2007). Predicting the risk of pancreatic cancer: on CDKN2A mutations in the melanoma-pancreatic cancer syndrome in Italy. J. Clin. Oncol. *25*, 5336–5337; author reply 5337–5338.

Ghiorzo, P., Bonelli, L., Pastorino, L., Bruno, W., Barile, M., Andreotti, V., Nasti, S., Battistuzzi, L., Grosso, M., Bianchi-Scarrà, G., et al. (2012). MC1R variation and melanoma risk in relation to host/clinical and environmental factors in CDKN2A positive and negative melanoma patients. Exp. Dermatol. *21*, 718–720.

Ghiorzo, P., Pastorino, L., Queirolo, P., Bruno, W., Tibiletti, M.G., Nasti, S., Andreotti, V., Paillerets, B.B. de, and Bianchi Scarrà, G. (2013). Prevalence of the E318K MITF germline mutation in Italian melanoma patients: Associations with histological subtypes and family cancer history. Pigment Cell Melanoma Res. *26*, 259–262.

Goldstein, A.M., Chan, M., Harland, M., Gillanders, E.M., Hayward, N.K., Avril, M.F., Azizi, E., Bianchi-Scarra, G., Bishop, D.T., Bressac-De Paillerets, B., et al. (2006). High-risk melanoma susceptibility genes and pancreatic cancer, neural system tumors, and uveal melanoma across GenoMEL. Cancer Res. *66*, 9818–9828.

Goldstein, A.M., Chan, M., Harland, M., Hayward, N.K., Demenais, F., Bishop, D.T., Azizi, E., Bergman, W., Bianchi-Scarra, G., Bruno, W., et al. (2007). Features associated with germline CDKN2A mutations: a GenoMEL study of melanoma-prone families from three continents. J. Med. Genet. *44*, 99–106.

Hall, M., and Peters, G. (1996). Genetic alterations of cyclins, cyclin-dependent kinases, and Cdk inhibitors in human cancer. Adv. Cancer Res. *68*, 67–108.

Harland, M., Holland, E.A., Ghiorzo, P., Mantelli, M., Bianchi-Scarrà, G., Goldstein, A.M., Tucker, M.A., Ponder, B.A.J., Mann, G.J., Bishop, D.T., et al. (2000). Mutation screening of the CDKN2A promoter in melanoma families. Genes Chromosom. Cancer *28*, 45–57.

Harland, M., Mistry, S., Bishop, D.T., and Bishop, J.A. (2001). A deep intronic mutation in CDKN2A is associated with disease in a subset of melanoma pedigrees. Hum. Mol. Genet. *10*, 2679–2686.

Harland, M., Taylor, C.F., Bass, S., Churchman, M., Randerson-Moor, J.A., Holland, E.A., Mann, G.J., Bishop, D.T., and Newton Bishop, J.A. (2005). Intronic sequence variants of the CDKN2A gene in melanoma pedigrees. Genes Chromosom. Cancer *43*, 128–136.

Hashemi, J., Platz, A., Ueno, T., Stierner, U., Ringborg, U., and Hansson, J. (2000). CDKN2A germ-line mutations in individuals with multiple cutaneous melanomas. Cancer Res. *60*, 6864–6867.

Horn, S., Figl, A., Rachakonda, P.S., Fischer, C., Sucker, A., Gast, A., Kadel, S., Moll, I., Nagore, E., Hemminki, K., et al. (2013). TERT promoter mutations in familial and sporadic melanoma. Science *339*, 959–961.

Huang, F.W., Hodis, E., Xu, M.J., Kryukov, G. V, Chin, L., and Garraway, L. a (2013). Highly recurrent TERT promoter mutations in human melanoma. Science *339*, 957–959.

Jenkins, N.C., Jung, J., Liu, T., Wilde, M., Holmen, S.L., and Grossman, D. (2013). Familial melanoma-associated mutations in p16 uncouple its tumor-suppressor functions. J. Invest. Dermatol. *133*, 1043–1051.

Kandoth, C., McLellan, M.D., Vandin, F., Ye, K., Niu, B., Lu, C., Xie, M., Zhang, Q., McMichael, J.F., Wyczalkowski, M. a, et al. (2013). Mutational landscape and significance across 12 major cancer types. Nature *502*, 333–339.

Kannengiesser, C., Brookes, S., del Arroyo, A.G., Pham, D., Bombled, J., Barrois, M., Mauffret, O., Avril, M.-F.M., Chompret, A., Lenoir, G.M., et al. (2009). Functional, structural, and genetic evaluation of 20 CDKN2A germ line mutations identified in melanoma-prone families or patients. Hum. Mutat. *30*, 564–574.

Liggett, W.H., and Sidransky, D. (1998). Role of the p16 tumor suppressor gene in cancer. J. Clin. Oncol. *16*, 1197–1206.

Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N., and Hogg, D. (1999). Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma. Nat. Genet. *21*, 128–132.

Maubec, E., Chaudru, V., Mohamdi, H., Blondel, C., Margaritte-Jeannin, P., Forget, S., Corda, E., Boitier, F., Dalle, S., Vabres, P., et al. (2012). Familial melanoma: clinical factors associated with germline CDKN2A mutations according to the number of patients affected by melanoma in a family. J. Am. Acad. Dermatol. *67*, 1257–1264.

McKenzie, H., Becker, T.M., Scurr, L.L., Kefford, R.F., and Rizos, H. (2009). Wild type and melanoma-associated mutant p16(IN4a) proteins do not oligomerize in vivo. Pigment Cell Melanoma Res. *22*, 131–133.

McKenzie, H.A., Fung, C., Becker, T.M., Irvine, M., Mann, G.J., Kefford, R.F., and Rizos, H. (2010). Predicting functional significance of cancer-associated p16INK4a mutations in CDKN2A. Hum. Mutat. *31*, 692–701.

Menin, C., Vecchiato, A., Scaini, M.C., Elefanti, L., Funari, G., De Salvo, G.L., Quaggio, M., Tognazzo, S., Agata, S., Dalla Santa, S., et al. (2011). Contribution of susceptibility gene variants to melanoma risk in families from the Veneto region of Italy. Pigment Cell Melanoma Res. *24*, 728–730.

Newton Bishop, J.A., Wachsmuth, R.C., Harland, M., Bataille, V., Pinney, E., Mack, P., Baglietto, L., Cuzick, J., and Bishop, D.T. (2000). Genotype/phenotype and penetrance studies in melanoma families with germline CDKN2A mutations. J. Invest. Dermatol. *114*, 28–33.

Newton-Bishop, J.A., Chang, Y.M., Iles, M.M., Taylor, J.C., Bakker, B., Chan, M., Leake, S., Karpavicius, B., Haynes, S., Fitzgibbon, E., et al. (2010). Melanocytic nevi, nevus genes, and melanoma risk in a large case-control study in the United Kingdom. Cancer Epidemiol. Biomarkers Prev. *19*, 2043–2054.

Pastorino, L., Bonelli, L., Ghiorzo, P., Queirolo, P., Battistuzzi, L., Balleari, E., Nasti, S., Gargiulo, S., Gliori, S., Savoia, P., et al. (2008). CDKN2A mutations and MC1R variants in Italian patients with single or multiple primary melanoma. Pigment Cell Melanoma Res. *21*, 700–709.

Pollock, P.M., Stark, M.S., Palmer, J.M., Walters, M.K., Aitken, J.F., Martin, N.G., and Hayward, N.K. (2001). Mutation analysis of the CDKN2A promoter in Australian melanoma families. Genes Chromosom. Cancer *32*, 89–94.

Qian, J., Liu, H., Wei, S., Liu, Z., Li, Y., Wang, L.E., Chen, W. V., Amos, C.I., Lee, J.E., Iles, M.M., et al. (2013). Association between putative functional variants in the PSMB9 gene and risk of melanoma - re-analysis of published melanoma genome-wide association studies. Pigment Cell Melanoma Res. *26*, 392–401.

Rocco, J.W., and Sidransky, D. (2001). p16(MTS-1/CDKN2/INK4a) in cancer progression. Exp. Cell Res. *264*, 42–55.

Scaini, M.C., Rossi, E., de Siqueira Torres, P.L.A., Zullato, D., Callegaro, M., Casella, C., Quaggio, M., Agata, S., Malacrida, S., Chiarion-Sileni, V., et al. (2009). Functional impairment of p16INK4A due to CDKN2A p.Gly23Asp missense mutation. Mutat. Res. - Fundam. Mol. Mech. Mutagen. *671*, 26–32.

Scaini, M.C., Minervini, G., Elefanti, L., Ghiorzo, P., Pastorino, L., Tognazzo, S., Agata, S., Quaggio, M., Zullato, D., Bianchi-Scarrà, G., et al. (2014). CDKN2A Unclassified Variants in Familial Malignant Melanoma: Combining Functional and Computational Approaches for Their Assessment. Hum. Mutat. *35*, 828–840.

Veinalde, R., Ozola, A., Azarjana, K., Molven, A., Akslen, L.A., Doniņa, S., Proboka, G., Cēma, I., Baginskis, A., and Pjanova, D. (2013). Analysis of Latvian familial melanoma patients shows novel variants in the noncoding regions of CDKN2A and that the CDK4 mutation R24H is a founder mutation. Melanoma Res. *23*, 221–226.

Zaccara, S., Tebaldi, T., Pederiva, C., Ciribilli, Y., Bisio, a, and Inga, a (2014). P53-Directed Translational Control Can Shape and Expand the Universe of P53 Target Genes. Cell Death Differ. 1–13.

Legends to Figures

Figure 1. Position and inheritance of *CDKN2A/p16^{INK4a}* 5'UTR sequence variants analysed in this study. A) Sequence of the 270nt *CDKN2A/p16^{INK4a}* 5'UTR. The position of all the variants analysed in this study is shown

in bold. The sequence changes and numeric position counting backwards from the established ATG are indicated above the sequence, in bold for novel variants. Deletions are underlined. (B-J) Melanoma pedigrees of patients carrying novel *CDKN2A/p16*^{INK4a} 5'UTR germline variants. Age at diagnosis indicates confirmation of cancer diagnosis and is provided below each symbol in brackets. Cancer type is indicated under each symbol. CMM: cutaneous melanoma; black symbols. PC: pancreatic cancer, HCC: hepatocellular carcinoma; CRC: colorectal cancer; RCC: renal cell cancer.

Figure 2. Impact of the *CDKN2A/p16*^{INK4a} 5'UTR variants in the bicistronic pRuF reporter vector. Gene reporter assays were performed in WM266-4, SK-MEL-5, G361 melanoma cell lines. Cells were transiently transfected with the pRuF bicistronic vector containing the wild type or variant 5'UTR, as indicated, cloned between the Renilla and Firefly cDNAs to investigate the post-transcriptional /translational regulation mechanisms. In the pRuF vector, transcription is driven by an SV-40-derived promoter located upstream of Renilla, while a polyA is present downstream of the Luciferase sequence. Firefly Luciferase activity was measured 27 hours post-transfection and normalized for transfection efficiency. Bars plot the average ratio of Firefly over Renilla activity, expressed as relative percentage in the comparison with the vector containing the wild type 5'UTR. The standard deviation of three biological repeats is presented. The experiments were repeated at least four times. In each cell line the presence of the wild type p16^{INK4a} 5'UTR led to a higher relative activity compared to the empty vector (~15 to 30 fold in WM266-4, SK-MEL-5, G361), confirming an IRES activity, while the variant c.-34G>T was always the least active, and resulted only in 2-3 fold induction of induction compared to the empty vector.

Figure 3. **Relative expression of endogenous p16 mRNA and protein from patients'-derived cell lines.** A) qPCR measurements of endogenous *CDKN2A/p16*^{INK4a} mRNA from total RNA obtained from melanoma patient-derived lymphoblastoid cell lines harboring the indicated 5'UTR variants at heterozygous levels. Bars are normalized on wild-type and are the average of three biological replicates. GAPDH and B2M served as reference genes. B), C), D) Western Blot analysis of endogenous p16^{INK4a} protein levels in the cell lines indicated above. Beta-Tubulin was used as loading control.

Figure 4. p16^{INK4a}allelic imbalance measured by next generation sequencing after polysomal profiling. Cytoplasmic RNA from melanoma patient-derived lymphoblastoid cell lines heterozygous for the indicated p16 5'UTR variants was separated on sucrose gradient and fractions were grouped as subpolysomal (SUB) and

polysomal (POL) based on the profile of UV_{254nm} absorbance. RNA from each fraction was subjected to RNA-sequencing using Ion Torrent PGM to measure p16 mRNA allelic coverage in the heterozygous samples. Light gray histograms surrounding the central plot present the read counts for each allele. Presented in the central panel is the relative imbalance determined as percentage of change in allelic representation in the two fractions (wild-type allele vs variant) in polysomal vs subpolysomal RNA.

Table 1. Frequency of *CDKN2A/p16^{INK4a}* 5'UTR germline variants in melanoma patients and controls from Italy, France and England.

Variant	Reference	Sporadic	Familial	SPM	MPM	Cosegregation	Cases N=5889(%)	Controls N=1709
c.-14C>T[(1)]	Hashemi, 2000	YES	-	-	YES	n.a.	3 (0.05)	-
c.-20A>G	Novel	YES	-	YES	-	n.a.	1 (0.02)	-
c.-21C>T[(1)]	Bisio, 2010	YES	YES	-	YES	YES	3 (0.05)	-
c.-27del23	Novel	-	YES	-	-	n.a.	2 (0.03)	-
c.-30G>A	Novel	YES	-	YES	-	n.a.	2 (0.03)	-
c.-34G>T[(1)]	Liu, 1999	-	YES	-	-	YES	2 (0.03)	-
c.-40C>T	Novel	YES	-	-	YES	n.a.	1 (0.02)	-
c.-42T>A[(2)]	Novel	YES	-	-	YES	n.a.	1 (0.02)	-
c.-25C>T+c.-180G>A[(1)]	Bisio, 2010	YES	-	YES	-	n.a.	3 (0.05)	-
c.-45G>A	Novel	YES	-	YES	-	n.a.	1 (0.02)	-
c.-56G>T[(1)]	Bisio, 2010					NO	1 (0.02)	-
c.-59C>G	Novel	YES	YES	YES	YES	NO	2 (0.03)	-
c.-67G>C[(1)]	Bisio, 2010					NO	1 (0.02)	1(0.06)
c.-87T>A	Novel	-	YES	YES	-	NO	1 (0.02)	-
c.-93-91delAGG	Novel	-	YES	YES	-	YES	1 (0.02)	-
c.-252A>T[(1)]	Pollock, 2001	-	YES	-	-	NO	-	-

[(1)] These variants were previously described. The first reference describing each variant is indicated.

[(2)]The c.-42T>A was found in one patient together with an additional variant c.-25C>T+c.-180G>A.

SPM=single primary melanoma; MPM=multiple primary melanoma; n.a.= not applicable.

Table 2. Comparison of four different assays used to assess the impact of *CDKN2A/ p16^{INK4a}* 5'UTR variants

CDKN2A/p16^{INK4a} 5'UTR Variant	Post transcriptional activity (pGL3-promoter)	Cap independent translation (pRuF)	Translational potential (Polysomal RNA allelic imbalance)	p16 expression (WB %)	Score	Impact
WT	100	100	NO	1	0	Neutral
c.-14C>T	73	82	NO	0.8	0	Neutral
c.-20A>G	92	89	-	1.7	0	Neutral
c.-21C>T	**49**	**47**	**YES**	**0.4**	**4**	***Causal***
c.-27del23	97	**51**	**YES**	**0.6**	**3**	***Causal***
c.-30G>A	132	197	NO	-	0	Neutral
c.-34G>T	**19**	**43**	**YES**	0.7	**3**	***Causal***
c.-40C>T	88	110	-	3.2	0	Neutral
c.-42T>A	93	118	**YES**	0.7	**1**	**Potential**
c. -25C>T+c.-180G>A	74	138	NO		0	Neutral
c.-45G>A	82	86	NO	2.9	**0**	Neutral
c.-56G>T	**57**	**60**	**YES**	-	**3**	***Causal***
c.-59C>G	89	94	NO	2.2	0	Neutral
c.-67G>C	170	**47**	**YES**	-	**2**	**Potential**
c.-87T>A	71	119	-	0.9	0	Neutral
c.-93-91delAGG	75	**54**	**YES**	**0.1**	**3**	***Causal***
c.-252A>T	105	104	NO	-	0	Neutral

Values reported in the second and third are the mean of the values obtained by each variant in the three different melanoma cell lines. A score was assigned to each variant to assign a pathogenicity class and determine the functional impact: neutral (no difference compared to wild type in at least 3 assays, score 0); potential (one or two assays displaying less than 60% of wild type activity, or less than 60% of protein expression, or positive imbalance -≥ 10%, score 1-2); causal (concordance of three or more assays showing impairment, score>3); - (not done). Potential and causal variants are highlighted in bold and italics characters, respectively, together with the assays leading to the assigned score.

Abbreviations

UTR: untranslated region, IRES: Internal Ribosome Entry Site, CDKN2A: cyclin dependent kinase 2A, NGS: next generation sequencing, SNV: single nucleotide variant, SNP: single nucleotide polymorphism, MPM: multiple primary melanoma. WT: wild-type, LCLs: lymphoblastoid cell lines.

Figure 1.

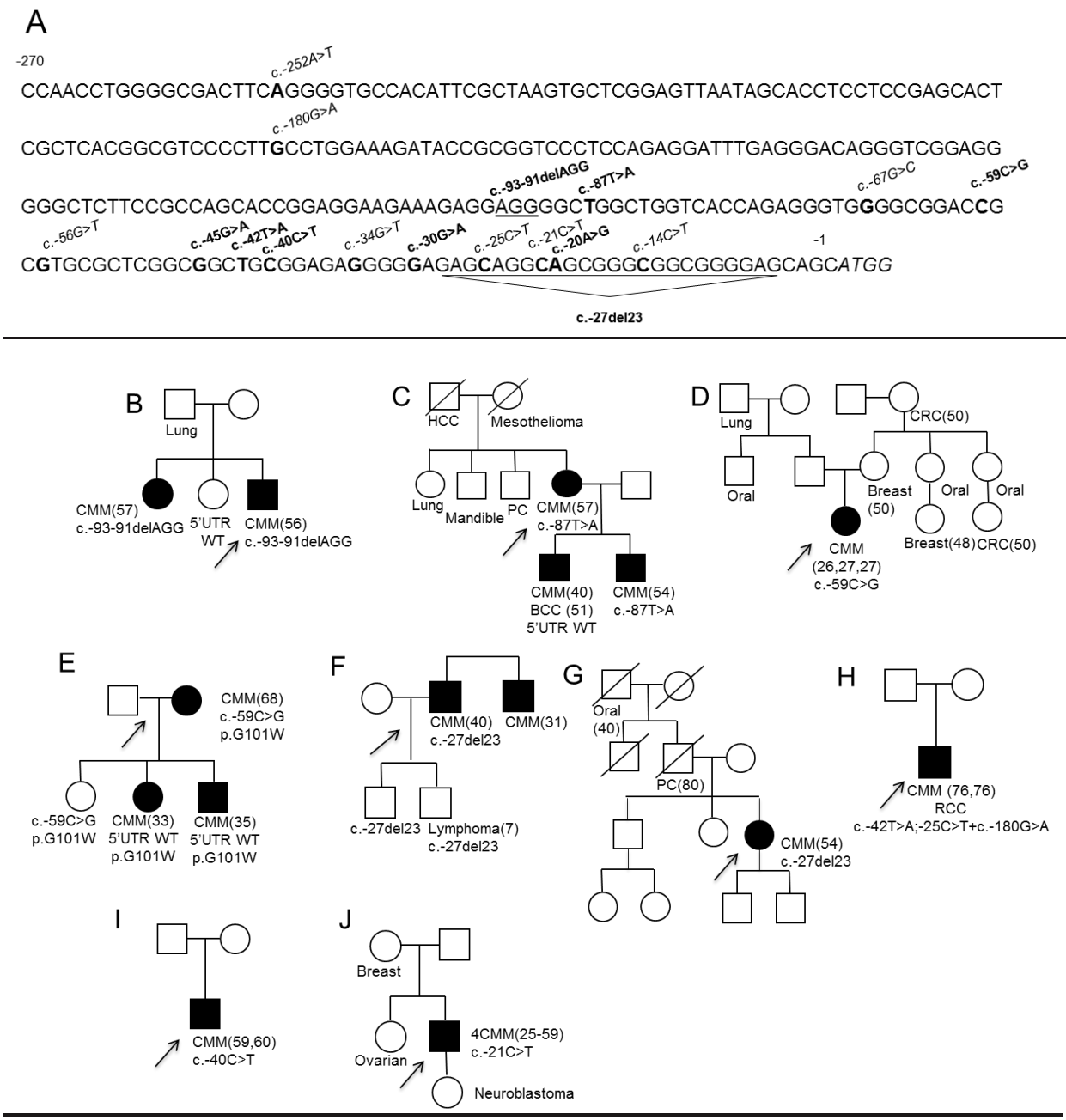


Figure 2.

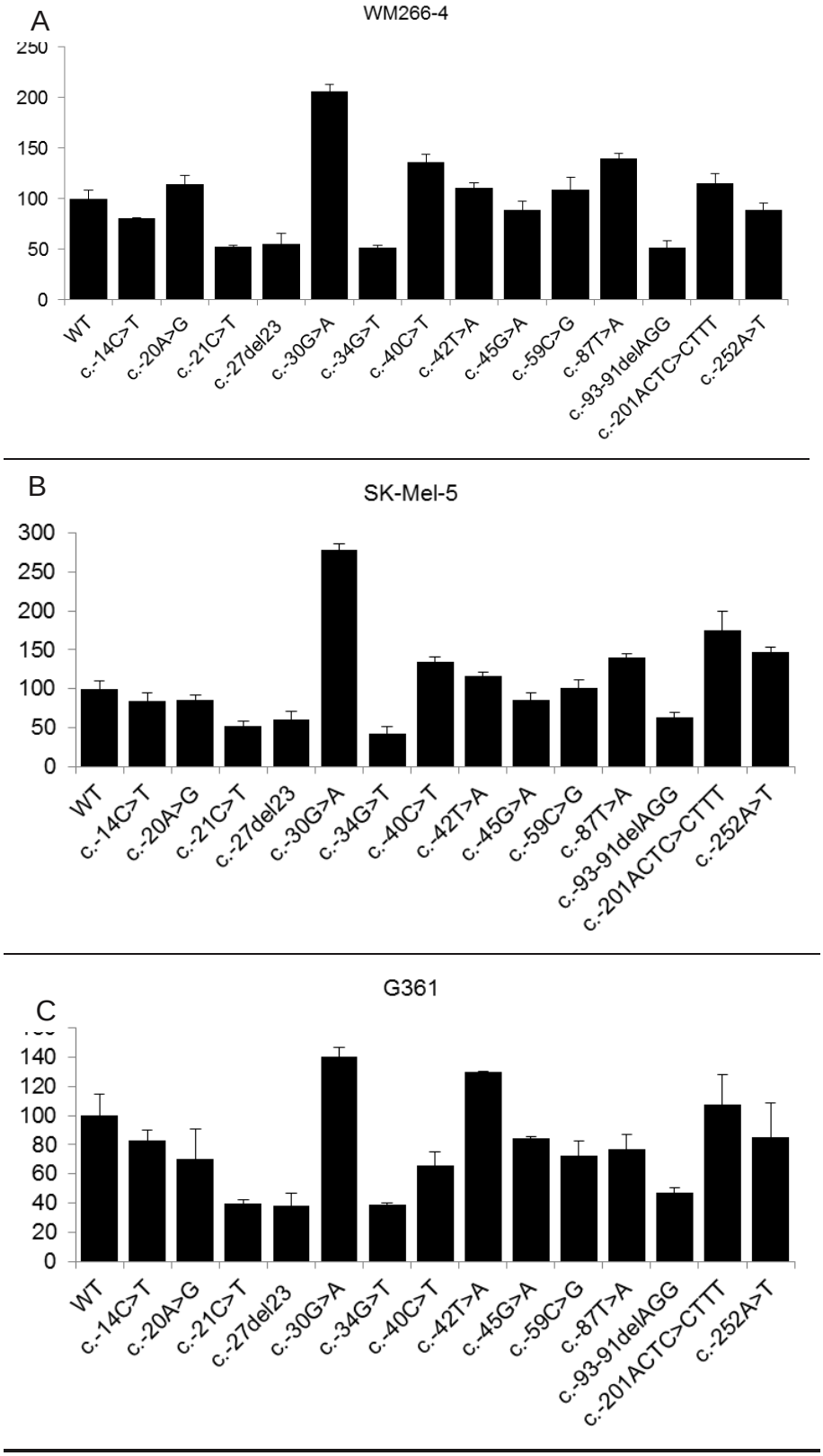


Figure 3.

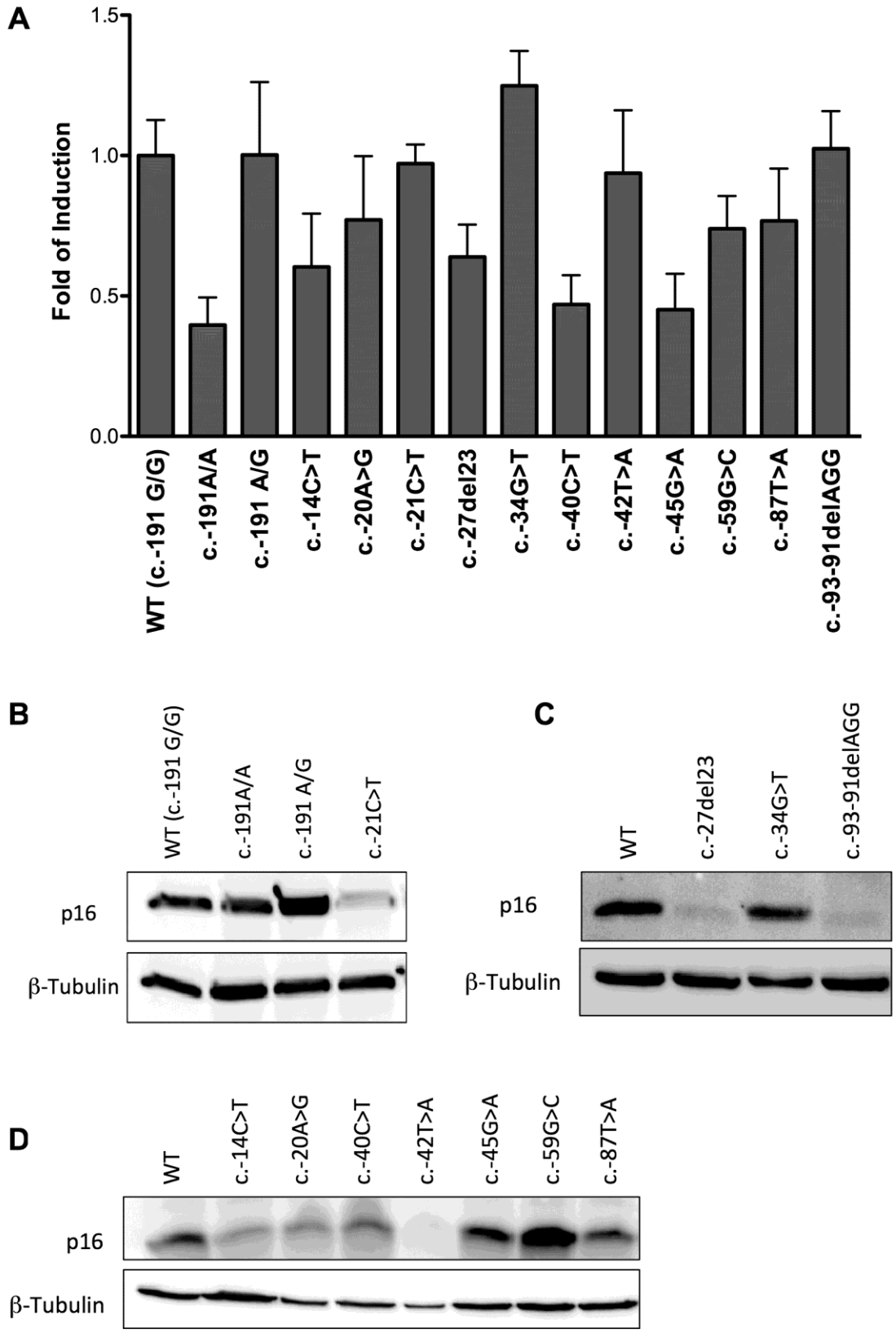
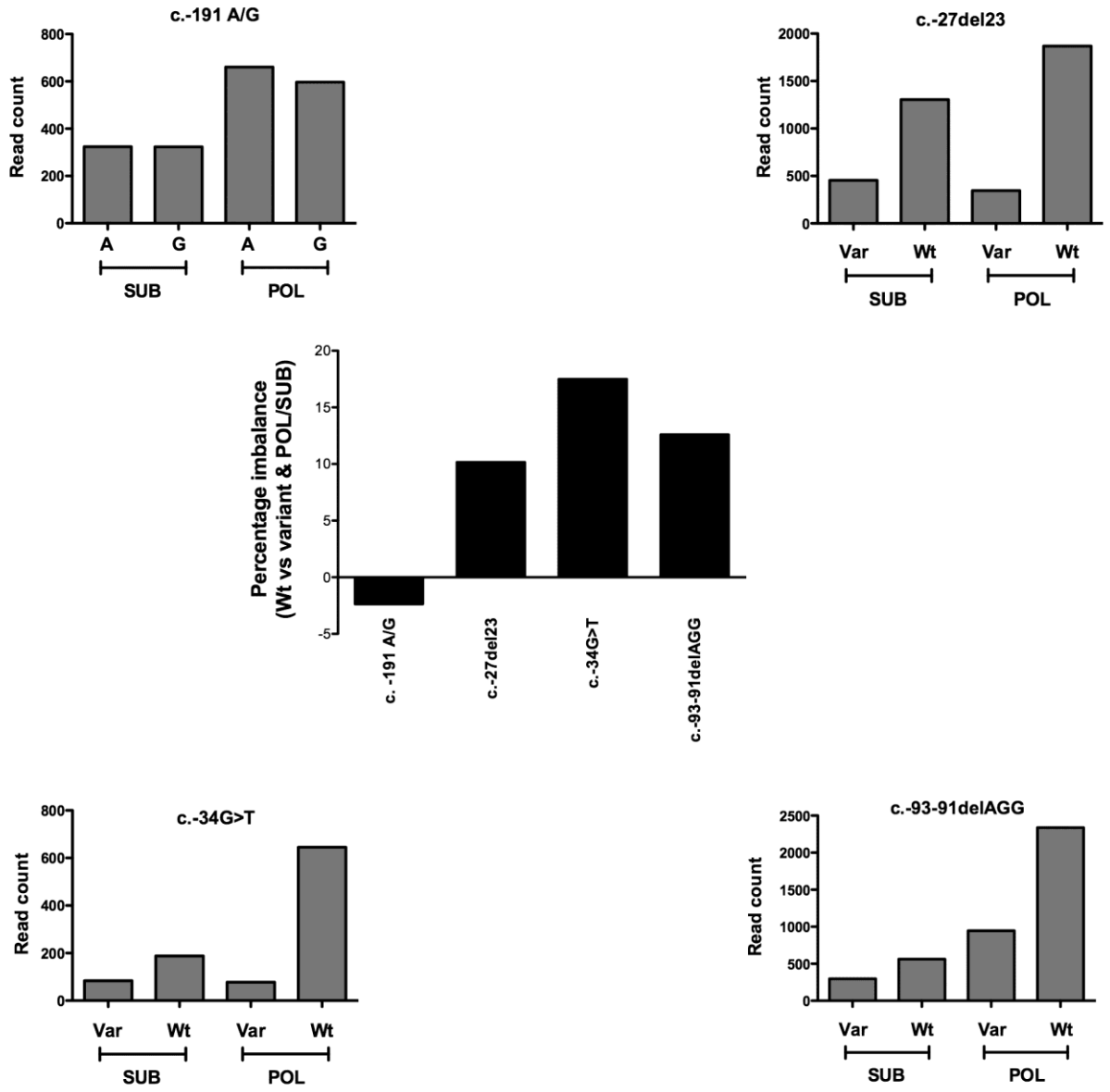

A
Fold of Induction
1.5
1.0
0.5
0.0
WT (c.-191 G/G)
c.-191A/A
c.-191 A/G
c.-14C>T
c.-20A>G
c.-21C>T
c.-27del23
c.-34G>T
c.-40C>T
c.-42T>A
c.-45G>A
c.-59G>C
c.-87T>A
c.-93-91delAGG
B
WT (c.-191 G/G)
c.-191A/A
c.-191 A/G
c.-21C>T
p16
β-Tubulin
C
WT
c.-27del23
c.-34G>T
c.-93-91delAGG
p16
β-Tubulin
D
WT
c.-14C>T
c.-20A>G
c.-40C>T
c.-42T>A
c.-45G>A
c.-59G>C
c.-87T>A
p16
β-Tubulin

Figure 4.


c.-191 A/G
Read count
800
600
400
200
0
A
G
A
G
SUB
POL
c.-27del23
Read count
2000
1500
1000
500
0
Var
Wt
Var
Wt
SUB
POL
Percentage imbalance
(Wt vs variant & POL/SUB)
20
15
10
5
0
-5
c. -191 A/G
c.-27del23
c.-34G>T
c.-93-91delAGG
c.-34G>T
Read count
800
600
400
200
0
Var
Wt
Var
Wt
SUB
POL
c.-93-91delAGG
Read count
2500
2000
1500
1000
500
0
Var
Wt
Var
Wt
SUB
POL

Supplemental Data

Supplemental Methods

Plasmid construction and assessment of activity of c.-191 alleles in the reporter assays

Using a site-directed mutagenesis approach based on a two-round PCR protocol and/or with direct PCR on genomic DNA extracted from the patient, we generated the fragments containing the variants of interest in CDKN2A 5'UTR: c.-14C>T, c.-20A>G, c.-21C>T, c.-25C>T+c.-180G>A, c.-27del23, c.-30G>A, c.-34G>T, c.-40C>T, c.-42T>A, c.-45G>A, c.-59C>G, c.-87T>A, c.-93-91delAGG, c.-252A>T (Table 1).

pGL3-promoter and pRuF bicistronic vectors containing the whole wild-type or variant p16INK4a 5'UTR were prepared as previously described (Bisio et al, 2010), but with a slight modification of the 3' cloning restriction sites, as in the original pRuF vector inserts were cloned using Nde I and both the p16 ATG site and the Firefly luciferase ATG sites were present and separated by 2 nucleotides. To remove the possibility of biased results due to the potential utilization of both ATGs, we modified our plasmid pRuF by site-directed mutagenesis Kit (Stratagene) by replacing the NdeI cloning site with PmlI. We thus reconstructed the reporter plasmid (pRuF) with the wild type 5'UTR and the variants already tested during preliminary analysis, and re-performed transfections, obtaining the same results. An equivalent strategy was put forward for the pGL3-promoter-based reporter assay. All the new variants presented in this work were cloned using the new version of the pRuF and pGL3promoter vector and using EcoR I and Pml I restriction sites. The resulting vectors containing the wild-type or variant CDKN2A 5'UTR were checked by direct DNA sequencing.

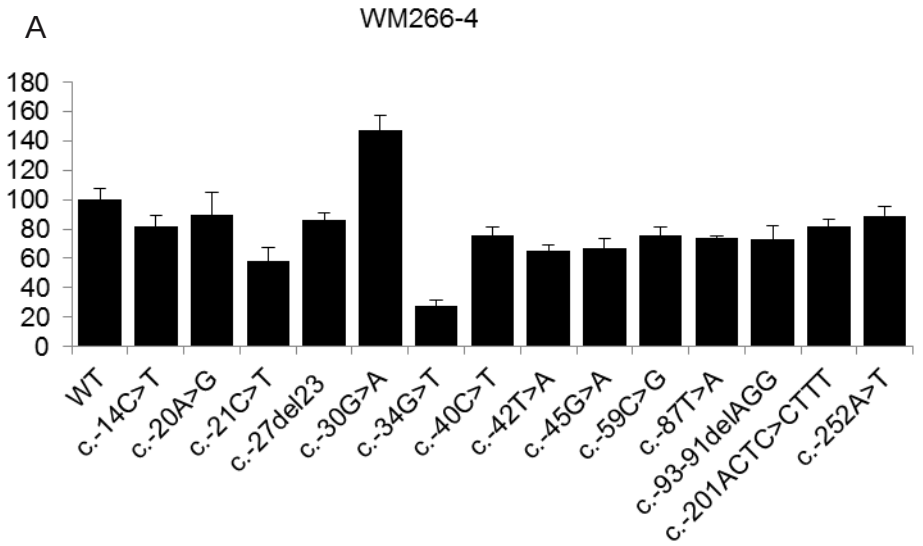
Since several of the variants analyzed were in cis with the common SNP at c.-191A>G and to further confirm that the two c.-191 SNP alleles showed no differences in reporter assays, we examined them alone, in cis with the known melanoma-associated variant (c.-34G>A) or with two of the novel germline variants tested herein (c.-27del23; c.-20C>T) in the WM266-4 melanoma cell line. The c-191 A as well as c-191 G alleles did not appear to act as a modifier of the activity of the identified variants, both considering the relative reporter activity and the actual transcript levels, measured by quantitative PCR, that were not affected by the SNP status (not shown).

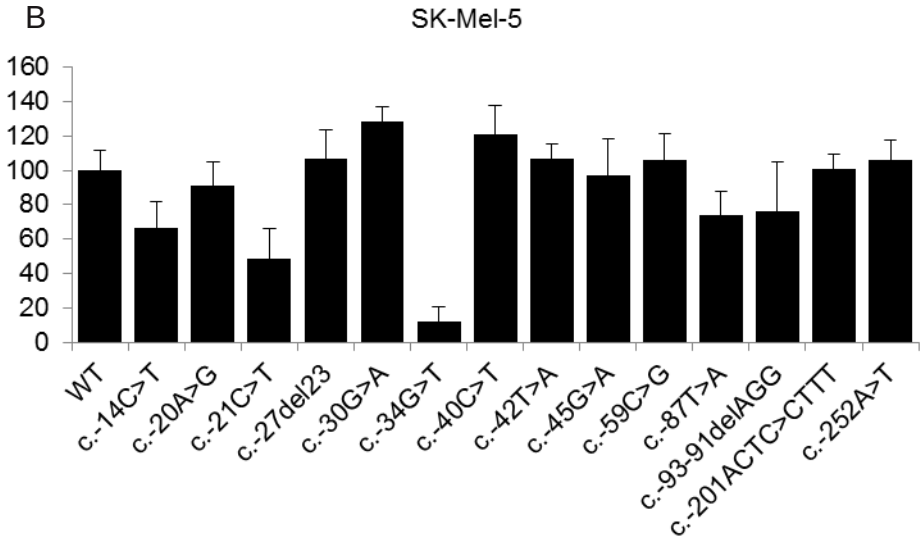
Supplemental Table

Table S1. Impact of the *CDKN2A/p16^{INK4a}* 5'UTR germline variants in pGL3-promoter and pRuF-derived constructs in three melanoma cell lines.

	Relative (%) luciferase reporter activity compared to the wild-type 5'UTR					
	pGL3-promoter			**pRuF**		
p16^{INK4a} 5'UTR variant	WM266-4	SK-mel-5	G361	WM266-4	SK-mel-5	G361
WT	100	100	100	100	100	100
c.-14C>T	81	66	72	81	83	82
c.-20A>G	89	90	96	113	85	70
c.-21C>T	58	48	42	51	52	39
c.-27del23	86	106	100	55	59	38
c.-30G>A	147	128	120	200	250	140
c.-34G>T	28	12	16	51	41	38
c.-40C>T	75	120	70	136	130	65
c.-42T>A	65	106	107	110	115	130
c.-45G>A	67	97	83	88	85	84
c.-59C>G	75	105	88	108	101	72
c.-87T>A	74	74	66	139	140	77
c.-93-91delAGG	74	76	76	51	64	47
c.-252A>T	89	106	120	82	146	85

Supplemental Figures





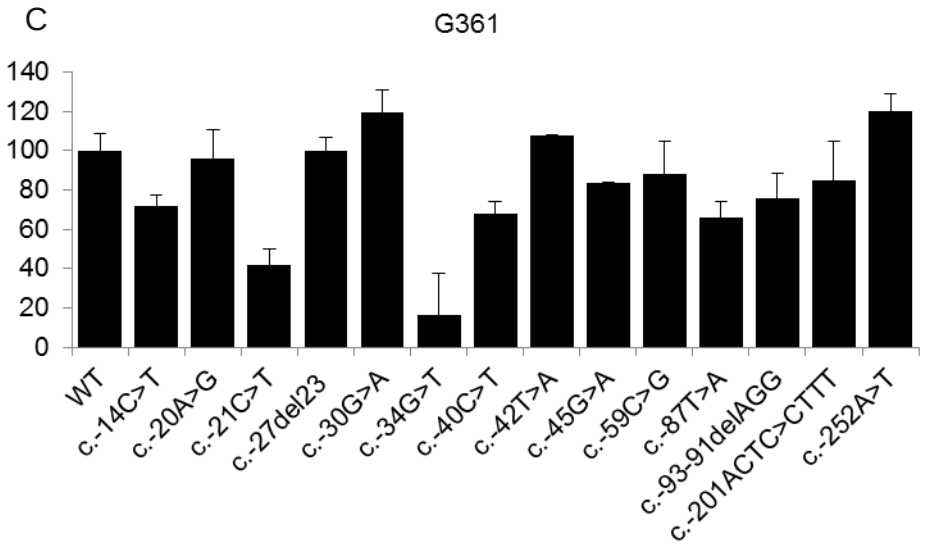


Figure S1. **Impact of the *CDKN2A/p16^{INK4a}* 5'UTR variants in pGL3 promoter constructs.** Gene reporter assays in WM266-4, SK-MEL-5, G361 melanoma cell lines. Cells were transiently co-transfected with the reporter constructs containing the wild type or variant 5'UTR, as indicated, along with the pRL-SV40 control vector. Luciferase activity was measured 27 hours post-transfection and normalized for transfection efficiency and expressed as average fold of induction compared to the empty vector. Bars plot the average activity, normalized for transfection efficiency and expressed as relative percentage in the comparison with the vector containing the wild type 5'UTR.. The standard deviation of three biological repeats is presented. The experiments were repeated at least four times in each cell line. While the impact of the variants was very consistent, the three cell lines were somewhat different in the relative activity of the wild type *CDKN2A/p16^{INK4a}* 5'UTR. In WM266-4 the wild type 5'UTR sequence led to an ~5 fold induction of the Firefly luciferase compared to the empty pRuF vector, while in SK-Mel-5 and G361 the induction were 2 and 5 fold, respectively. In the three cell lines the c.-34G>T variant showed the lowest activity that was however slightly higher compared to the empty vector, with a minimal activity of ~1.5 fold in WM266-4 cells.

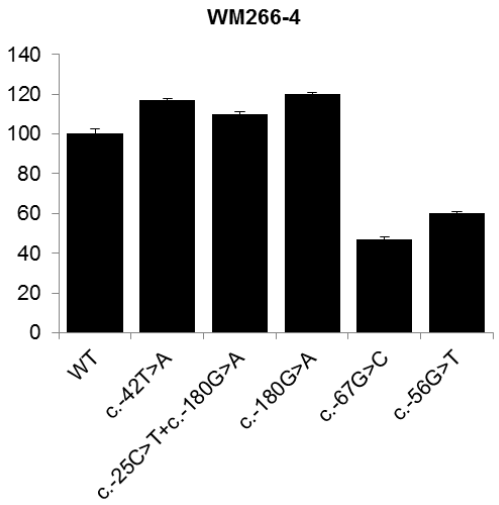


Figure S2. Impact of new and previously described *CDKN2A/p16^{INK4a}* 5'UTR variants in the bicistronic pRuF reporter vector in melanoma cell lines.

The impact of the c.-42T>A and the double variant allele c.-25 C>T+ c.-180G>A, which were detected in trans in a compound heterozygote patient, was analysed in the bicistronic pRuF reporter vector. The c.-180G>A was also tested as single variant; c.-67G>C and c.-56G>T, previously identified but not tested in melanoma-derived cell lines were also examined. Bars plot the average activity, normalized for transfection efficiency and expressed as relative percentage in the comparison with the vector containing the wild type 5'UTR. Results in the WM266-4 melanoma cell line are shown.

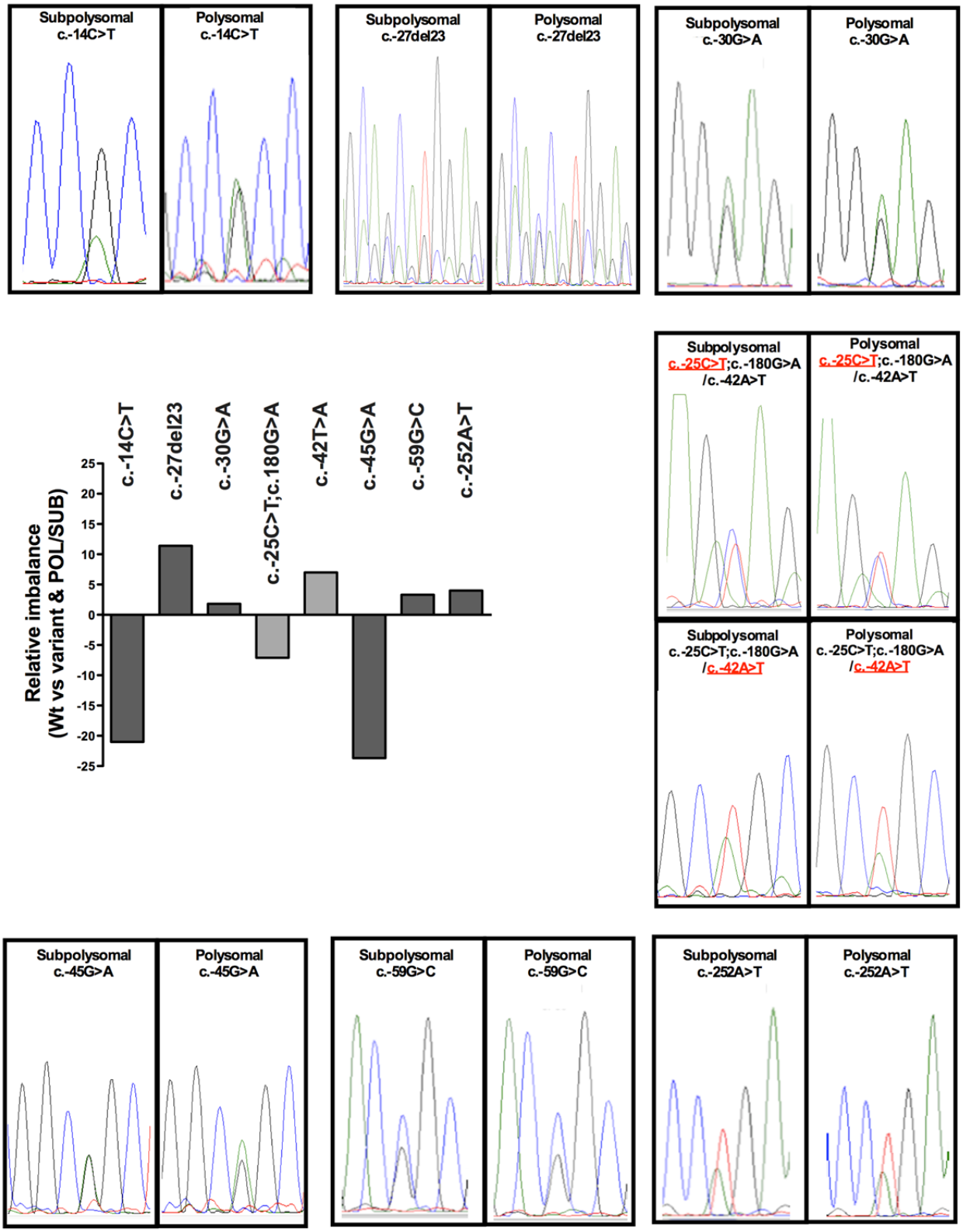


Figure S3. Allelic imbalance between subpolysomal and polysome-associated mRNAs in patients' derived

lymphoblastoid cell lines.

The indicated variants were analyzed using polysomal profiling, RNA extraction from pooled subpolysomal and polysomal fractions, PCR amplification of the entire p16^{INK4a} 5'UTR sequence, double strand Sanger sequencing and quantification of the electropherograms. Presented is a representative sequencing results and a bar graph of the measured imbalance plotted as difference in signal of the wild type over the variant allele in the comparison between subpolysomal and polysomal sequencing. In this rendering, a negative imbalance indicates that the variant allele is more abundant in polysomal fractions. The variants c.-25C>T+c.-180G>A and c.-42T>A are highlighted in light gray as this patient-derived cell line is a p16^{INK4a} compound heterozygote, hence the imbalance of one variant allele is relative to the second allele. In this case, the result indicate that the c.-42T>A is slightly less represented in the polysomes compared to the c.-25C>T+c.-180G>A one, although the separate analysis of each allele in the gene reporter assays did not show a significant defect for each allele compared to the wild type p16^{INK4a} 5'UTR (Figure S1, S2, Table S1).

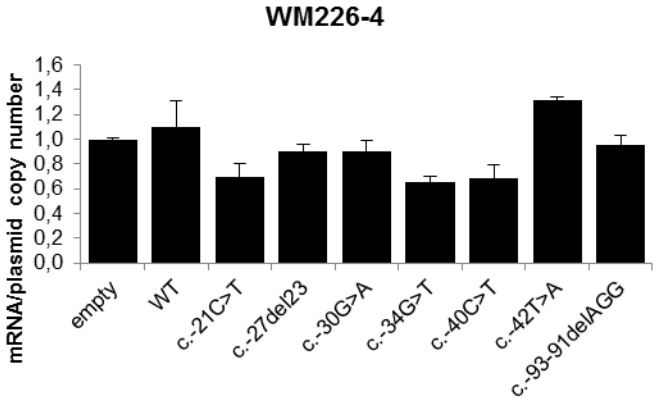


Figure S4. Quantification of luciferase mRNA expression by real-time PCR.

Melanoma cell lines were transiently transfected with pGL3 promoter plasmid containing the wild type or a panel of p16^{INK4a} 5'UTR representative variants. Total RNA and genomic DNA were extracted using the AllPrep DNA/RNA Mini Kit (Qiagen, Milan, Italy) according to the manufacturer's instructions, cDNA was generated starting from 1 µg of RNA by using the AffinityScript cDNA Synthesis Kit (Stratagene, Milan, Italy). Real-time PCR was performed using a cDNA aliquot equivalent to 50 ng of converted RNA or with 10 ng of genomic DNA using a RotorGene 6000 thermal cycler (Corbett Life Science, Ancona, Italy) and the 5PRIME Sybr-Green MasterMix (Eppendorf, Milan, Italy). Relative mRNA quantification was obtained using the ΔCt method, taking into account the efficiency of cDNA synthesis by the quantification of the Beta2Microglobulin (β2M) housekeeping gene from the same cDNA. The average plasmid copy number was quantified using the same PCR

primers on genomic DNA extracted from the same transfection plate. Small differences in relative Firefly mRNA expression levels were observed, which were not consistent with the reporter activity reflecting luciferase protein expression (see for example c.-21C>T, c.-34G>T and c.-40C>T). However an impact of some of these variants on mRNA stability cannot be completely excluded.