




## Research Paper

# A Functional Synonymous Variant in *PDGFRA* Is Associated with Better Survival in Acral Melanoma

Jie Dai<sup>1\*</sup>, Lu Yang<sup>1,2\*</sup>, Tianxiao Xu<sup>1</sup>, Lu Si<sup>1</sup>, Chuanliang Cui<sup>1</sup>, Xinan Sheng<sup>1</sup>, Zhihong Chi<sup>1</sup>, Lili Mao<sup>1</sup>, Bin Lian<sup>1</sup>, Bixia Tang<sup>1</sup>, Xue Bai<sup>1</sup>, Li Zhou<sup>1</sup>, Siming Li<sup>1</sup>, Xuan Wang<sup>1</sup>, Xieqiao Yan<sup>1</sup>, Yan Kong<sup>1</sup>, Jun Guo<sup>1</sup>

1. Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Renal Cancer and Melanoma, Peking University Cancer Hospital and Institute, Beijing 100142, China
2. Department of Radiology, Peking University Shougang Hospital, Beijing 100144, China

\*J. Dai and L. Yang contributed equally to this work.

 Corresponding authors: Jun Guo, M.D., Ph.D., Department of Renal Cancer and Melanoma, Peking University Cancer Hospital and Institute, 52# Fucheng Road, Haidian District, Beijing 100142, China, E-mail: guoj307@126.com; Yan Kong, M.D., Ph.D., Department of Renal Cancer and Melanoma, Peking University Cancer Hospital and Institute, 52# Fucheng Road, Haidian District, Beijing 100142, China, E-mail: k-yan08@163.com

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2019.12.12; Accepted: 2020.01.18; Published: 2020.03.04

## Abstract

**Purpose:** Polymorphisms of genes in the platelet-derived growth factor (PDGF) signaling pathway have been found to predict cutaneous melanoma (CM) survival, but their clinical effects in acral melanoma (AM) patients have not been explored. The aim of this study was to characterize the functional effect of the tag single-nucleotide polymorphism (SNP) rs2228230:C>T and assess its association with clinical outcomes in AM patients.

**Methods:** The effect of rs2228230:C>T on mRNA structures and codon usage values were evaluated using *in silico* analyses. PDGF receptor alpha (*PDGFRA*) expression vectors with the rs2228230:C or rs2228230:T allele were constructed to evaluate the expression and signaling activity of *PDGFRA*. The expression of *PDGFRA* in AM samples was measured using *in situ* RNAscope hybridization and immunohistochemical staining. The association of the rs2228230 genotype with survival was analyzed in two independent AM cohorts.

**Results:** *In silico* analyses indicated that the rs2228230:T allele increases the minimum free energy and reduces synonymous codon usage. The rs2228230:T allele decreased the expression of *PDGFRA* by reducing the stability of its mRNA and protein as well as the signaling activity of the MAPK and PI3K/AKT pathways. *PDGFRA* mRNA and protein expression was significantly reduced in AM tissues with the rs2228230:T allele. The progression-free survival and overall survival of AM patients with the rs2228230:T allele were significantly longer than those of patients with the CC genotype.

**Conclusion:** Our study indicated that rs2228230:T can reduce the expression of *PDGFRA* and downstream signaling activity and is associated with better survival in AM patients.

Key words: acral melanoma, *PDGFRA*, genetic variation, gene expression regulation, survival

## Introduction

Platelet-derived growth factor (PDGF) receptor alpha (*PDGFRA*) encodes a cell surface tyrosine kinase receptor, which binds to three forms of PDGFs and forms a homodimer or heterodimer, thereby mediating many biological processes, including organ development, wound healing, angiogenesis, cell proliferation, and differentiation [1,2]. *PDGFRA* plays a role in tumor progression, and mutations in

*PDGFRA* have been associated with idiopathic hypereosinophilic syndrome [3], gastrointestinal stromal tumors [4], and several other cancers [5]. We and others have demonstrated that mutations, increased copy numbers, and overexpression of *PDGFRA* occur in melanoma patients, and that functional mutations of *PDGFRA* increase MAPK and PI3K/AKT pathway activation, which can be

inhibited by several tyrosine kinase inhibitors such as imatinib, crenolanib, and sunitinib [6-8].

Besides functional mutations, polymorphisms can also influence susceptibility to disease and response to anticancer drug treatment. A genome-wide association study (GWAS) in Australian and European populations found that the single-nucleotide polymorphism (SNP) rs3219090 within *PARP1* is a melanoma susceptibility locus [9], and subsequent studies demonstrated that the rs3219090 risk allele is correlated with higher expression levels of *PARP1* mRNA but longer melanoma-specific survival [10,11]. The polymorphism rs333 (*CCR5*Δ32) within *CCR5* results in premature translation termination, and a study on a large melanoma cohort found that this polymorphism was associated with decreased overall survival (OS) in stage IV melanoma patients treated with immunotherapy [12]. The genetic variants of members of the PDGF signaling pathway have been demonstrated to predict cutaneous melanoma (CM) survival in patients in the United States [13]. Unlike in Caucasians, in whom the most prevalent melanoma subtype is non-acral CM, the major subtype in Asians is acral melanoma (AM), which accounts for 42.8% of melanoma cases in China [14]. However, the clinical relevance of *PDGFRA* polymorphisms in AM patients has not been explored yet.

rs2228230:C>T is the most common and only tag SNP located within the coding sequence of *PDGFRA* (Figure S1) [15]. The tag SNP was mapped to the ATP-binding site and polypeptide substrate-binding site in a conserved protein kinase (PTKc) domain; furthermore, it lies within the exonic splicing enhancer/exonic splicing silencer (ESE/ESS)-binding sites [16], which may disrupt mRNA splicing and affect protein function. The minor T allele of rs2228230 in infants was reported to be associated with a reduced risk of obstructive heart defects after alcohol exposure during the periconceptional period [17]. Meanwhile, the genotype of rs2228230 was found to be significantly associated with corneal astigmatism according to a meta-analysis of five GWASs of 8,513 Asian individuals [18]. The rs2228230:T allele has also been demonstrated to be associated with worse disease-free survival in renal cell carcinoma in a Spanish population [19]. Therefore, in this study, we investigated the functional effect of rs2228230 through *in silico* and *in vitro* studies as well as the association of the rs2228230 genotype with *PDGFRA* expression and clinical outcomes of AM patients in two independent cohorts.

## Materials and Methods

### Study participants

In this retrospective exploratory study, data from patients diagnosed with melanoma who received standard treatment from Peking University Cancer Hospital and Institute between January 2010 and August 2017 were reviewed. A total of 546 AM patients were enrolled in the discovery cohort, and 256 AM patients were enrolled in the replication cohort. A group of 240 CM patients were enrolled in an independent cohort. The last follow-up was carried out in March 2019; the median follow-up times were 33.2 months (range, 0.2–103.8 months), 25.9 months (range, 1.6–94.7 months), and 31.0 months (range, 0.6–91.5 months). All procedures were approved by the Ethics Committee of Peking University Cancer Hospital and Institute and were conducted in adherence to the principles of the Declaration of Helsinki. Informed consent for the use of specimens was obtained from all participants.

### Cell culture

The HEK293 cell line was obtained from the Cell Bank of the Chinese Academy of Sciences and cultured in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum, 100 U/mL penicillin, 100 mg/mL streptomycin, and 2 mM GlutaMAX at 37 °C in a humidified incubator with 5% CO<sub>2</sub>. All reagents used for cell culture were purchased from Gibco (Grand Island, NY, USA). The cell line was authenticated using short tandem repeat (STR) profiling and was confirmed to be negative for mycoplasma.

### DNA extraction and genotyping

Genomic DNA was extracted from formalin-fixed, paraffin-embedded (FFPE) tissue sections using the QIAamp DNA FFPE Tissue Kit (Qiagen, Hilden, Germany). Polymerase chain reaction (PCR)-amplified products were subjected to Sanger sequencing to determine the genotypes of rs2228230 according to our previous report [6]. Chromatogram results were further analyzed by BLAST and manual review.

### RNA extraction and quantitative reverse transcription PCR

Total RNA was extracted using the RNeasy Mini Kit (Qiagen). Two micrograms of RNA was reverse transcribed into cDNA using the ReverTra Ace qPCR RT Kit (Toyobo, Osaka, Japan). Quantitative reverse transcription PCR (qRT-PCR) was performed using GoTaq® qPCR Master Mix (Promega, Madison, WI, USA) on an ABI 7500 FAST Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). *GAPDH*

was used as an internal control, and the relative quantitative levels of *PDGFRA* expression were calculated by the  $2^{-\Delta\Delta C_t}$  method.

### RNA in situ hybridization of *PDGFRA*

*PDGFRA* mRNA in FFPE melanoma samples was measured manually using the RNAscope assay (Advanced Cell Diagnostics, ACD, Hayward, CA, USA). After deparaffinization and incubation with pretreatment reagents, the slides were hybridized with Hs-*PDGFRA*-probes (ACD Probe: 604481) in a HybEZ oven (ACD) at 40 °C for 2 h. Hybridization signals were amplified and detected using the RNAscope 2.5 HD Detection Kit (Red; ACD, 322360). After image capture, digital *in situ*-hybridized slides were analyzed by Fiji (ImageJ) software. Positive signals appeared as red dots. Human peptidylprolyl isomerase B and dihydrodipicolinate reductase were used as positive and negative control probes, respectively, to evaluate RNA quality. The results were scored using RNAscope® Spot Studio Software (ACD) as follows: “0,” no staining or <1 dot per 10 cells; “1,” 1–3 dots per cell; “2,” 4–9 dots per cell or very few clusters; “3,” 10–15 dots per cell or <10% dots in clusters; and “4,” >15 dots per cell or >10% dots in clusters. When the score of the positive control probes was  $\geq 2$ , the tissue section was defined as qualified.

### Immunohistochemical staining

FFPE tissue slides were deparaffinized with xylene and rehydrated through a graded alcohol series. Next, endogenous peroxidase activity was blocked using hydrogen peroxide. After antigen retrieval, FFPE slides were incubated with an anti-*PDGFRA* antibody (dilution 1:250; Santa Cruz Biotechnology, Santa Cruz, CA, USA) overnight in darkness. Horseradish peroxidase (HRP)-conjugated polyclonal rabbit antibody (Dako, Glostrup, Denmark) diluted at 1:400 was used as the secondary antibody. Following antibody incubation, slides were washed twice and then subjected to detection with chromogen aminoethyl carbazole (AEC; ZSGB-BIO, Beijing, China). Specific immunostaining of each sample was scored as 0–3 based on the staining intensity and density by pathologists blinded to the genotype of the samples.

### Plasmid construction and oligonucleotide transfection

Plasmid vectors carrying full-length fragments of *PDGFRA* containing the coding sequence with either the C (pLenti-CMV-*PDGFRA*(C)-3FLAG-PGK-Puro) or T (pLenti-CMV-*PDGFRA*(T)-3FLAG-PGK-Puro) allele of rs2228230 were designed and synthesized by Obio Technology (Shanghai, China).

The control plasmid pLenti-CMV-3FLAG-PGK-Puro vector was purchased from Obio Technology. Transient transfection of plasmids into HEK293 cells was performed using Lipofectamine 3000 (Invitrogen, Carlsbad, CA, USA).

### Protein and mRNA stability assays

After 48 h of transfection, cells were treated with 5  $\mu\text{g}/\text{mL}$  actinomycin D (ActD; APExBIO, Houston, TX, USA) to inhibit RNA synthesis and harvested at 0, 0.5, 1, 2, and 4 h. RNA was extracted and quantified by qRT-PCR to detect mRNA stability. To determine protein stability, cycloheximide (CHX; 20  $\mu\text{g}/\text{mL}$ ; APExBIO) was added to the culture medium to inhibit protein synthesis for 0, 2, 4, and 8 h, and protein was extracted and detected by western blotting. At 48 h after transfection, cells were starved overnight and exposed to 1  $\mu\text{g}/\text{mL}$  MG132 (APExBIO) for 6 h to inhibit protein degradation, and protein was extracted for western blotting.

### Western blotting

Total protein was extracted from cells using RIPA lysis buffer (Beyotime Biotechnology, Shanghai, China) supplemented with protease inhibitors (Roche Diagnostics, Indianapolis, IN, USA). Immunodetection was performed using rabbit anti-*PDGFRA*, anti-phospho-*PDGFRA*, anti-AKT, anti-phospho-AKT, anti-ERK1/2, anti-phospho-ERK1/2 (all 1:1000; CST, Danvers, MA, USA) and mouse anti-GAPDH antibodies (1:5000; Abcam, Cambridge, UK) as primary antibodies. HRP-conjugated anti-mouse or anti-rabbit IgG antibodies (1:3000, CST) were used as secondary antibodies. Protein bands were visualized with Amersham ECL Select Western Blotting Detection Reagent (GE Healthcare, Chicago, IL, USA), and band intensities were quantified using ImageJ64.

### In silico analysis

To identify the rs2228230 genotype in the TCGA Skin Cutaneous Melanoma (SKCM) dataset, level 2 TCGA SKCM genotype data were extracted from NCI Genomic Data Commons Data Portal (GDC Legacy Archive) with a Birdseed confidence threshold of <0.05 [20]. RNAseq expression data were extracted as “RNA Seq V2 RSEM” from the level 3 file “rem.genes.normalized\_result” using the TCGAbiolinks package in R language. The clinical information was extracted from the GDC portal using the TCGAbiolinks package. The effect of rs2228230:C>T on the mRNA structures and thermodynamics of *PDGFRA* were predicted using the RNAfold Web Server. Codon usage values were obtained from the Codon Usage Database, and the relative synonymous codon usage (RSCU) was calculated as  $RSCU = S \times N_c / N_a$ , where S represents the number of synonymous codons



encoding the same amino acid,  $N_c$  is the frequency of the codon in the genome, and  $N_a$  is the relative frequency of the codon for that amino acid [21].

### Statistical analysis

SPSS 21.0 software was used for all statistical analyses. Normally distributed continuous data such as age and Breslow's thickness were evaluated using the two independent sample t-test and described as mean  $\pm$  SD. Categorical data were described as frequencies and percentages. Genotype frequencies and clinical parameters such as sex, subtype, stage, and metastasis were compared using the  $\chi^2$  test or Fisher's exact test. Relative *PDGFRA* mRNA expression and the H-score of *PDGFRA* protein were evaluated by the Mann-Whitney U test. Survival curves were established using the Kaplan-Meier method, and the statistical significance of the results was estimated by log-rank tests. Multivariate analysis was carried out using the Cox proportional hazard regression model. Deviations of allele and genotype frequencies were estimated by the Hardy-Weinberg equilibrium. All statistical tests were two-sided, and significance was considered at  $P < 0.05$ .

## Results

### Regulation of the mRNA secondary structure and expression of *PDGFRA* by the rs2228230 genotype

To clarify the association of rs2228230 genotype with *PDGFRA* expression, we conducted a series of *in silico* and *in vitro* analyses. First, the mRNA secondary structure of full-length *PDGFRA* was predicted using the RNAfold WebServer, which showed that the mRNA secondary structure of *PDGFRA* with the rs2228230:C allele (Figure 1A) differed from that with the rs2228230:T allele (Figure 1B). The minimum free energy (MFE) values of *PDGFRA* with the rs2228230:C allele and rs2228230:T allele were predicted to be -1392.95 kcal/mol and -1478.56 kcal/mol, respectively, indicating that the stability of *PDGFRA* mRNA with the rs2228230:T allele was lower than that with the rs2228230:C allele. Second, we extracted the codon usage values and calculated the RSCU values to verify if the rs2228230:T allele could alter the codon usage of valine. We found that the RSCU value of GTC (rs2228230:C) is higher than that of GTT (rs2228230:T) at both the genome (Figure 1C) and *PDGFRA* gene level (Figure 1D). The RSCU value for GTC in *PDGFRA* is greater than 1, which implies a bias for GTC codon usage in the valine synthesis process. These *in silico* analyses indicated that the rs2228230:T allele might reduce *PDGFRA* expression through decreased mRNA stability and lower codon usage.

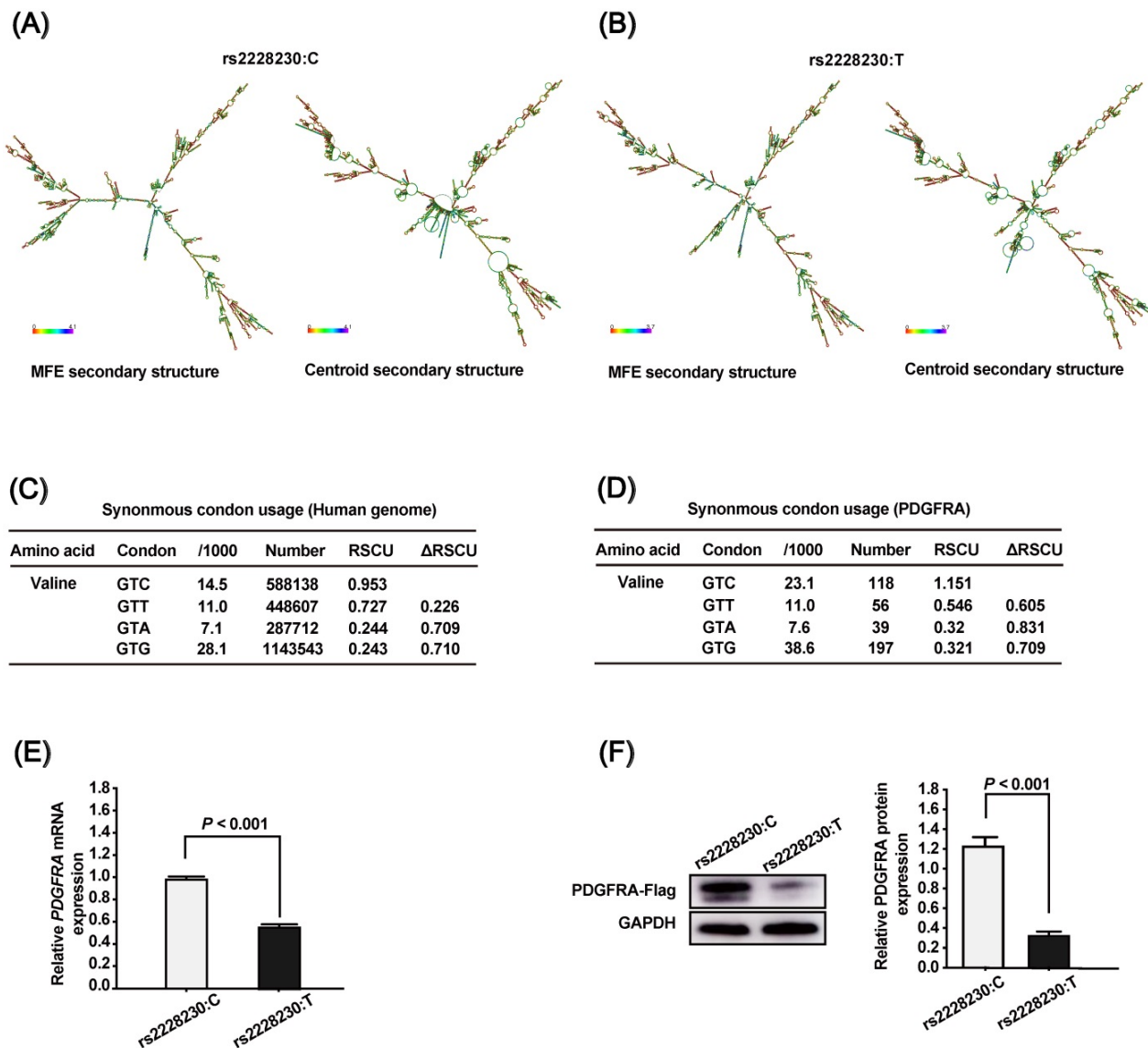
To validate this hypothesis, we constructed *PDGFRA* expression vectors with the rs2228230:C allele or rs2228230:T allele containing a FLAG tag sequence and transfected them into HEK293 cells. After 48 h of transfection, mRNA was extracted and transcribed into cDNA using a reverse primer that can amplify the FLAG tag to avoid the effect of endogenous *PDGFRA* expression. qRT-PCR showed that the relative expression level of *PDGFRA* mRNA was 1.75-fold lower in cells transfected with the rs2228230:T allele than in cells transfected with the rs2228230:C allele ( $P < 0.001$ ; Figure 1E). Protein expression of *PDGFRA* was further examined by western blotting after 48 h of transfection, and the results were consistent with those of qRT-PCR; the expression level of *PDGFRA* was four-fold lower in cells with the rs2228230:T allele ( $P < 0.001$ ; Figure 1F). These results conclusively demonstrated that the genotype of rs2228230 affected the expression of *PDGFRA* at both the mRNA and protein level.

### Influence of the rs2228230 genotype on the mRNA and protein stability of *PDGFRA*

We further investigated whether post-transcriptional or post-translational mechanisms were involved in the reduced expression of *PDGFRA* with the rs2228230:T allele. For the mRNA stability assay, ActD was added to the culture medium to inhibit RNA synthesis. As shown in Figure 2A, the mRNA decay rate of *PDGFRA* was significantly higher in cells with the rs2228230:T allele than in cells with the rs2228230:C allele ( $P < 0.001$ ), indicating that *PDGFRA* mRNA stability was reduced by the rs2228230:T allele.

Cells transfected with the rs2228230:C allele or rs2228230:T allele of *PDGFRA* were treated with the protein synthesis inhibitor CHX for various durations. *PDGFRA* levels decreased with increasing CHX exposure time, and the degradation rate of *PDGFRA* in cells with the rs2228230:T allele was significantly faster than that in cells with the rs2228230:C allele, indicating that the rs2228230:T allele increased proteolytic degradation ( $P < 0.001$ ; Figure 2B-D).

MG132 is a proteasome inhibitor that can block the activity of the 20S proteasome and the proteolytic activity of the 26S proteasome complex [22]. After proteolytic degradation was inhibited, cells with the rs2228230:T allele showed a 43% decrease in *PDGFRA* translation compared to cells with the rs2228230:C allele ( $P < 0.001$ ; Figure 2E, F), indicating that the rs2228230:T allele also decreased the protein synthesis rate of *PDGFRA*. Hence, the rs2228230 genotype impacts the mRNA and protein stability of *PDGFRA*.



**Figure 1.** The genotype of rs2228230 affects the mRNA secondary structure and expression of PDGFRA. The mRNA secondary structures of full-length PDGFRA with the (A) rs2228230:C allele or (B) rs2228230:T allele were predicted using the RNAfold WebServer. Frequencies of the valine codons per thousand and absolute numbers in (C) the genome and (D) PDGFRA gene. The relative synonymous codon usage (RSCU) was calculated as  $RSCU = S \times Nc/Na$ , where S represents the number of synonymous codons encoding the same amino acid, Nc is the frequency of the codon in the genome, and Na is the relative frequency of the codon for that amino acid. PDGFRA expression vectors with the rs2228230:C allele or rs2228230:T allele containing a FLAG tag sequence were constructed and transfected into HEK293 cells. After 48 h of transfection, (E) the mRNA expression of PDGFRA was evaluated by qRT-PCR, and (F) protein expression was detected by western blotting.

### Influence of the rs2228230 genotype on the signaling activity of PDGFRA

As the above investigations demonstrated that the genotype of rs2228230 could influence PDGFRA expression, we hypothesized that it might also affect PDGFRA function. To validate this hypothesis, we analyzed the activation of downstream signaling, including the MAPK and PI3K/AKT pathways. Cells transfected with different genotypes of PDGFRA were serum-starved overnight, followed by serum stimulation for 60 min. The basic expression and phosphorylation levels of all molecules were similar after serum starvation (Figure 2G, H); after serum stimulation, the expression levels of total and phosphorylated PDGFRA, phospho-AKT, and

phospho-ERK were significantly decreased in cells with the rs2228230:T allele compared to those with the rs2228230:C allele, while the expression levels of total AKT and ERK were not different in the two groups. Thus, the genotype of rs2228230 could influence the signaling activity of PDGFRA.

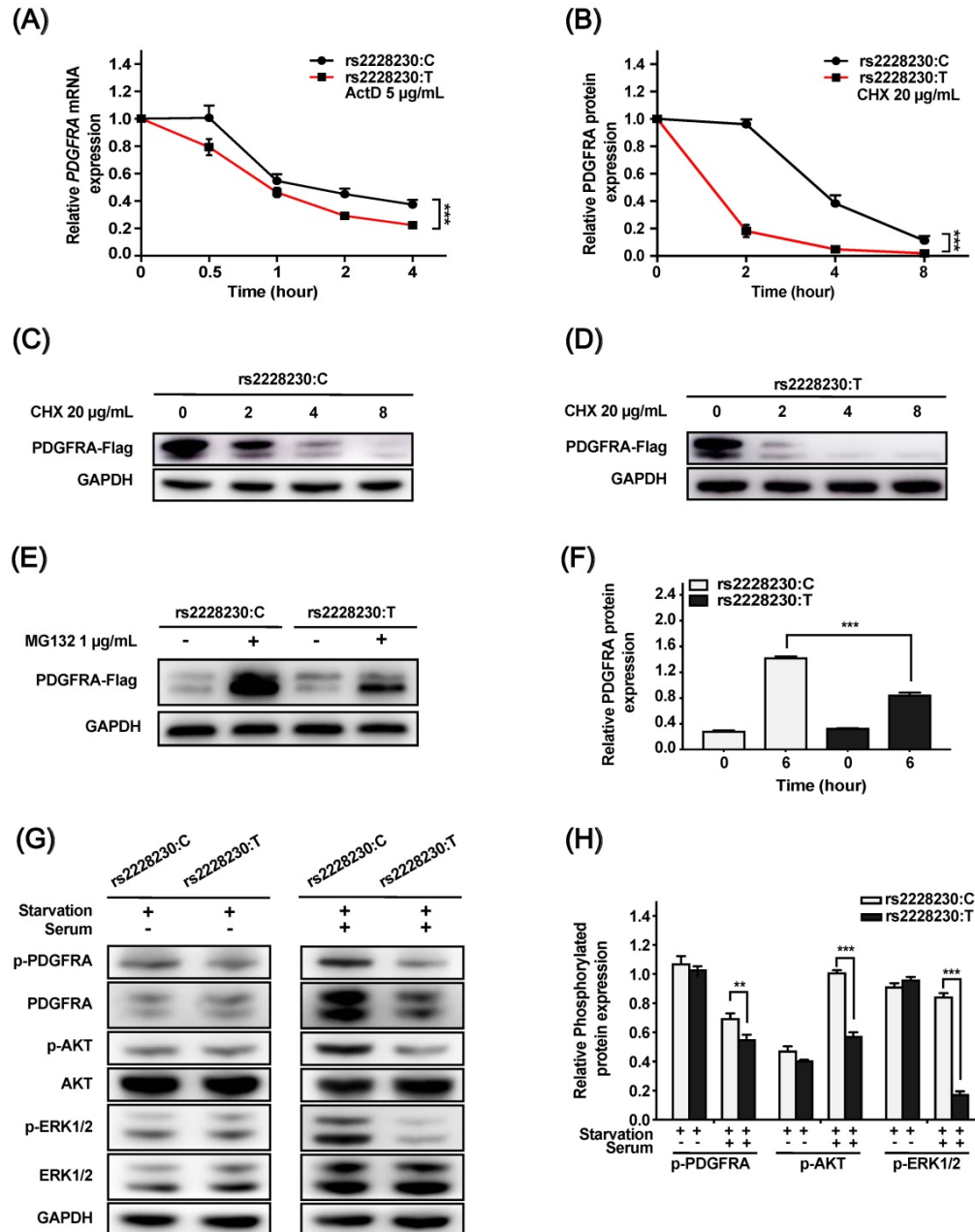
### Association of the rs2228230 genotype with the mRNA and protein expression of PDGFRA in AM

The association of the rs2228230 genotype with expression of PDGFRA in AM was validated in FFPE samples, as no commercial AM cell lines were available. The mRNA expression of PDGFRA was analyzed by semi-quantitative RNAscope *in situ*

hybridization in 104 FFPE AM samples (Figure 3A). The mean expression level of *PDGFRA* mRNA in the CC group was 27.101 (range: 0.000–97.353), while the mean value was 5.068 (range: 0.000–59.542) in the CT+TT group ( $P < 0.001$ ; Figure 3B, C), demonstrating that the *PDGFRA* mRNA expression level was significantly reduced in melanoma tissues with the rs2228230:T genotype.

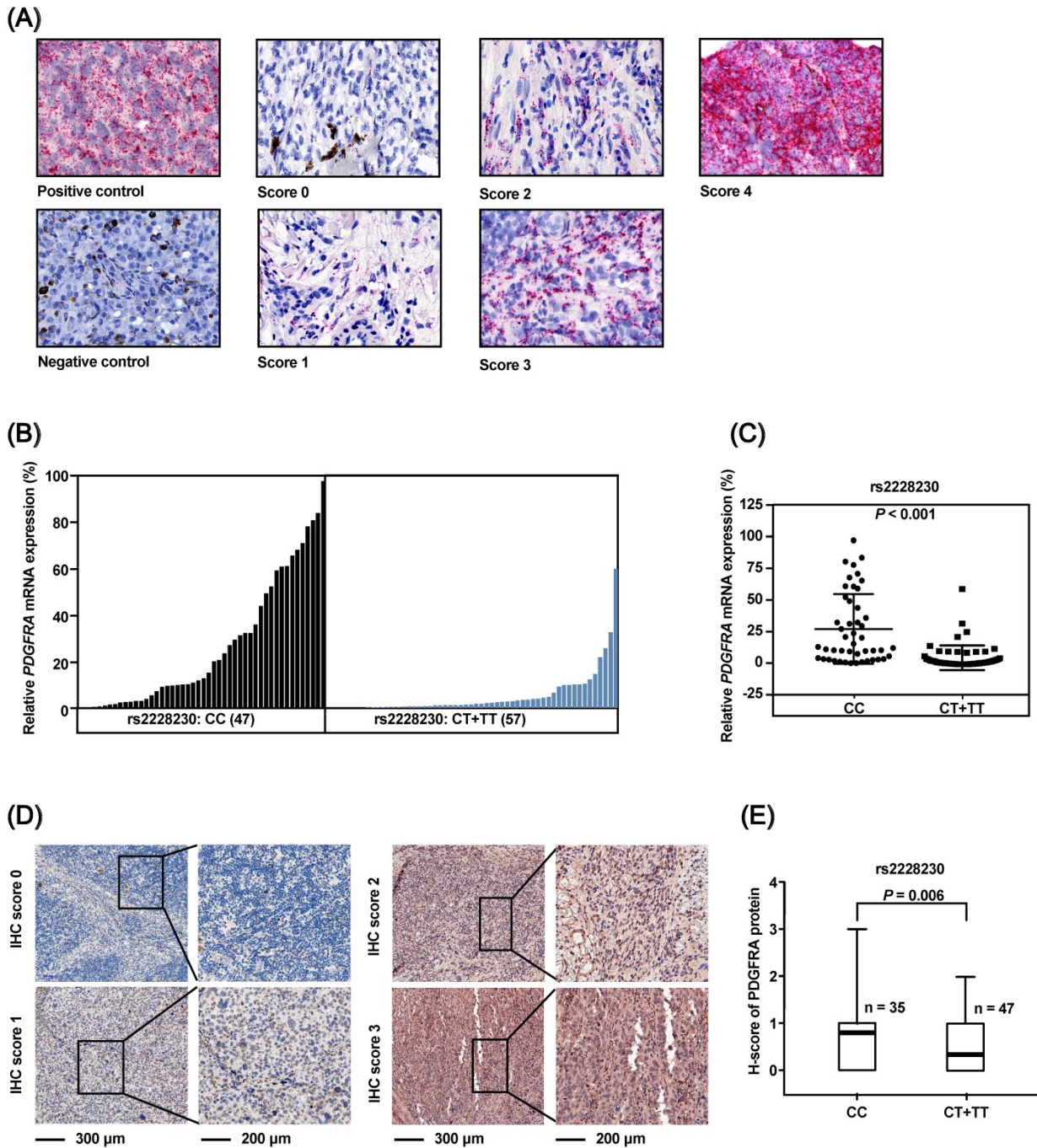
The association of rs2228230 with *PDGFRA* protein expression was analyzed by immunohistochemical (IHC) staining in 82 AM tissues. The positive

rate of *PDGFRA* in patients with the rs2228230:T allele (25.5%, 12/47) was significantly lower than in patients with the CC genotype (54.3%, 19/35;  $P = 0.008$ ). IHC staining was scored from 0 to 3 according to staining intensity and density (Figure 3D), and the results showed that the *PDGFRA* expression level in patients with the CT+TT genotype was significantly lower than that in patients with the CC genotype ( $P = 0.006$ ; Figure 3E). These data indicated that the rs2228230:T allele was associated with decreased expression of *PDGFRA* at both the mRNA and protein level in AM.



**Figure 2.** The genotype of rs2228230 alters the stability and signaling activity of *PDGFRA*. (A) HEK293 cells were transfected with *PDGFRA* expression vectors containing the rs2228230:C allele or rs2228230:T allele. At 48 h after transfection, cells were treated with actinomycin D (ActD) for different time periods to inhibit RNA synthesis, and *PDGFRA* expression was examined by qRT-PCR to evaluate mRNA stability. (B–D) Cells were treated with cycloheximide (CHX) to inhibit protein synthesis for different time periods, and protein was extracted for western blotting to detect the proteolytic degradation of *PDGFRA*. Cells were serum-starved overnight and treated with MG132 for 6 h, and (E) *PDGFRA* expression was examined by western blotting and (F) quantified to evaluate the protein synthesis rate. Cells were serum-starved overnight, followed by serum stimulation for 60 min. Next, the expression and phosphorylation of *PDGFRA*, as well as key molecules of MAPK and PI3K/AKT signaling, were (G) detected and (H) quantified. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .





**Figure 3.** The *PDGFRA* rs2228230 genotype is correlated with *PDGFRA* mRNA and protein expression level in AM. (A) Representative *in situ* RNAscope hybridization images of *PDGFRA* in FFPE specimens. Each red dot represents a single transcript of *PDGFRA* or the positive control. The relative expression of *PDGFRA* mRNA was normalized to positive control dot counts and is shown in a (B) waterfall plot and a (C) scatter plot. (D) Representative IHC staining images of *PDGFRA* in AM. Red staining of the cytoplasm and nucleus was considered positive. The staining of each sample was scored as 0–3 according to the staining intensity and density by pathologists blinded to the genotype of the samples. (E) Correlation between *PDGFRA* expression and rs2228230 genotype. Significance was assessed by the Mann–Whitney U test.

### Correlation of rs2228230 genotype with patient characteristics and clinical outcome in AM

The association of the rs2228230 genotype with prognosis in AM was further retrospectively analyzed in a discovery cohort and a replication cohort. The basic clinicopathological characteristics of the patients are summarized in Table S1. The distribution of rs2228230 was in concordance with the Hardy–

Weinberg equilibrium in both cohorts, and the minor allele frequency of rs2228230 was as expected in Asians according to the NCBI database of genetic variation dbSNP (Table S2). The correlation of the *PDGFRA* rs2228230 genotype with clinical characteristics of AM was analyzed, and no factor was found to be significantly correlated in either cohort (Table S3).

**Table 1.** Association of *PDGFRA* rs2228230 genotype with PFS and OS in AM

	Clinical character	Group	PFS				OS			
			Univariate HR (95% CI)	P value	Multivariate HR (95% CI)	P value	Univariate HR (95% CI)	P value	Multivariate HR (95% CI)	P value
Discovery cohort	Sex	Male vs. Female	1.190 (0.965-1.468)	0.103			1.382 (0.988-1.934)	0.059		
	Age	≥60 vs. <60	0.941 (0.753-1.175)	0.590			0.785 (0.543-1.133)	0.195		
	Thickness	≥4 vs. <4	0.993 (0.802-1.229)	0.948			1.239 (0.866-1.734)	0.211		
	TNM	III/IV vs. I/II	2.854 (2.296-3.549)	<0.001*	2.926 (2.352-3.640)	<0.001*	3.532 (2.443-5.017)	<0.001*	3.581 (2.477-5.178)	<0.001*
	rs2228230	CT/TT vs. CC	0.756 (0.595-0.960)	0.022*	0.696 (0.547-0.884)	0.003*	0.660 (0.447-0.976)	0.037*	0.630 (0.426-0.932)	0.021*
	<i>CKIT</i>	Mut vs. WT	1.023 (0.707-1.479)	0.906			0.868 (0.469-1.608)	0.654		
	<i>BRAF</i>	Mut vs. WT	1.148 (0.887-1.486)	0.293			0.856 (0.547-1.340)	0.497		
	<i>NRAS</i>	Mut vs. WT	1.281 (0.956-1.716)	0.097			1.319 (0.829-2.099)	0.236		
	<i>PDGFRA</i>	Mut vs. WT	0.892 (0.369-2.158)	0.801			1.265 (0.403-3.973)	0.687		
	Replication cohort	Sex	Male vs. Female	1.004 (0.753-1.338)	0.979			1.093 (0.754-1.586)	0.638	
Age		≥60 vs. <60	0.931 (0.690-1.256)	0.639			1.022 (0.694-1.505)	0.913		
Thickness		≥4 vs. <4	0.844 (0.621-1.146)	0.276			1.066 (0.710-1.600)	0.757		
TNM		III/IV vs. I/II	1.762 (1.268-2.448)	0.001*	1.883 (1.353-2.622)	<0.001*	2.172 (1.361-3.467)	0.001*	2.340 (1.463-3.743)	<0.001*
rs2228230		CT/TT vs. CC	0.625 (0.447-0.874)	0.006*	0.622 (0.443-0.873)	0.006*	0.593 (0.377-0.932)	0.023*	0.614 (0.388-0.972)	0.037*
<i>CKIT</i>		Mut vs. WT	1.068 (0.657-1.737)	0.790			1.140 (0.593-2.190)	0.694		
<i>BRAF</i>		Mut vs. WT	1.114 (0.809-1.618)	0.446			1.092 (0.680-1.756)	0.715		
<i>NRAS</i>		Mut vs. WT	1.806 (1.205-2.706)	0.004*	1.749 (1.162-2.634)	0.007*	2.541 (1.583-4.080)	<0.001*	2.432 (1.504-3.935)	<0.001*
<i>PDGFRA</i>		Mut vs. WT	0.566 (0.140-2.287)	0.424			0.455 (0.063-3.261)	0.433		
Combined cohort		Sex	Male vs. Female	1.105 (0.933-1.307)	0.247			1.193 (0.932-1.528)	0.162	
	Age	≥60 vs. <60	0.935 (0.782-1.117)	0.458			0.892 (0.685-1.163)	0.399		
	Thickness	≥4 vs. <4	1.004 (0.849-1.188)	0.960			1.357 (1.058-1.741)	0.016*	1.184 (0.921-1.521)	0.187
	TNM	III/IV vs. I/II	2.465 (2.061-2.949)	<0.001*	2.536 (2.118-3.036)	<0.001*	3.238 (2.429-4.316)	<0.001*	3.232 (2.420-4.315)	<0.001*
	rs2228230	CT/TT vs. CC	0.707 (0.582-0.859)	<0.001*	0.656 (0.540-0.798)	<0.001*	0.615 (0.458-0.827)	0.001*	0.586 (0.436-0.788)	<0.001*
	<i>CKIT</i>	Mut vs. WT	1.042 (0.777-1.398)	0.781			0.998 (0.639-1.561)	0.995		
	<i>BRAF</i>	Mut vs. WT	1.154 (0.938-1.419)	0.175			0.948 (0.685-1.313)	0.748		
	<i>NRAS</i>	Mut vs. WT	1.426 (1.126-1.805)	0.003*	1.346 (1.063-1.705)	0.014*	1.676 (1.206-2.328)	0.002*	1.557 (1.119-2.167)	0.009
	<i>PDGFRA</i>	Mut vs. WT	0.783 (0.371-1.650)	0.520			0.847 (0.315-2.276)	0.742		

\*,  $P < 0.05$ . AM, acral melanoma; PFS, progression-free survival; OS, overall survival; HR, hazard ratio; TNM, tumor-node-metastasis stage; Mut, mutation; WT, wild type.

In the discovery cohort, Kaplan–Meier survival analysis revealed that both progression-free survival (PFS;  $P = 0.022$ ; Figure 4A) and OS ( $P = 0.037$ ; Figure 4B) of AM patients with the rs2228230:T allele were significantly longer compared to those of patients with the CC genotype. Multivariate Cox regression analysis results are shown in Table 1; the rs2228230:T allele was significantly associated with longer PFS (hazard ratio, HR: 0.696; 95% CI: 0.547–0.884;  $P = 0.003$ ) and OS (HR: 0.630; 95% CI: 0.426–0.932;  $P = 0.021$ ).

In the replication cohort, PFS (HR: 0.622; 95% CI: 0.443–0.873;  $P = 0.006$ ) and OS (HR: 0.614; 95% CI: 0.388–0.972;  $P = 0.037$ ) of patients with the rs2228230:T allele were also longer than those of patients with the CC genotype (Figure 4C, D). The association of the rs2228230 genotype with prognosis was further analyzed in the combined cohort, and the protective effect of the rs2228230:T allele remained significant (Figure 4E, F); AM patients with the rs2228230:T allele had longer PFS (HR: 0.656; 95% CI: 0.540–0.798;  $P < 0.001$ ) and OS (HR: 0.586; 95% CI: 0.436–0.788;  $P < 0.001$ ) compared to patients with the CC genotype.

### Association of the rs2228230 genotype with clinical outcome in CM patients

The prognostic significance of rs2228230 was verified in CM. The baseline patient characteristics are

summarized in Table S4, and the distribution of rs2228230 was in concordance with the Hardy–Weinberg equilibrium (Table S5). Both Kaplan–Meier survival analysis (Figure 5A, B) and univariate Cox analysis (Table S6) showed that the genotype of rs2228230 was not related to PFS or OS in CM. The association of the rs2228230 genotype with prognosis and *PDGFRA* expression was also analyzed in the TCGA SKCM dataset, in which most of the patients were Caucasian. As in our CM cohort, the genotype of rs2228230 was not related to PFS or OS (Figure 5C, D). Moreover, the mRNA expression of *PDGFRA* was not different between the different genotypes (Figure S2). These results demonstrated that the effect of rs2228230 on *PDGFRA* expression and survival in CM differed according to original site.

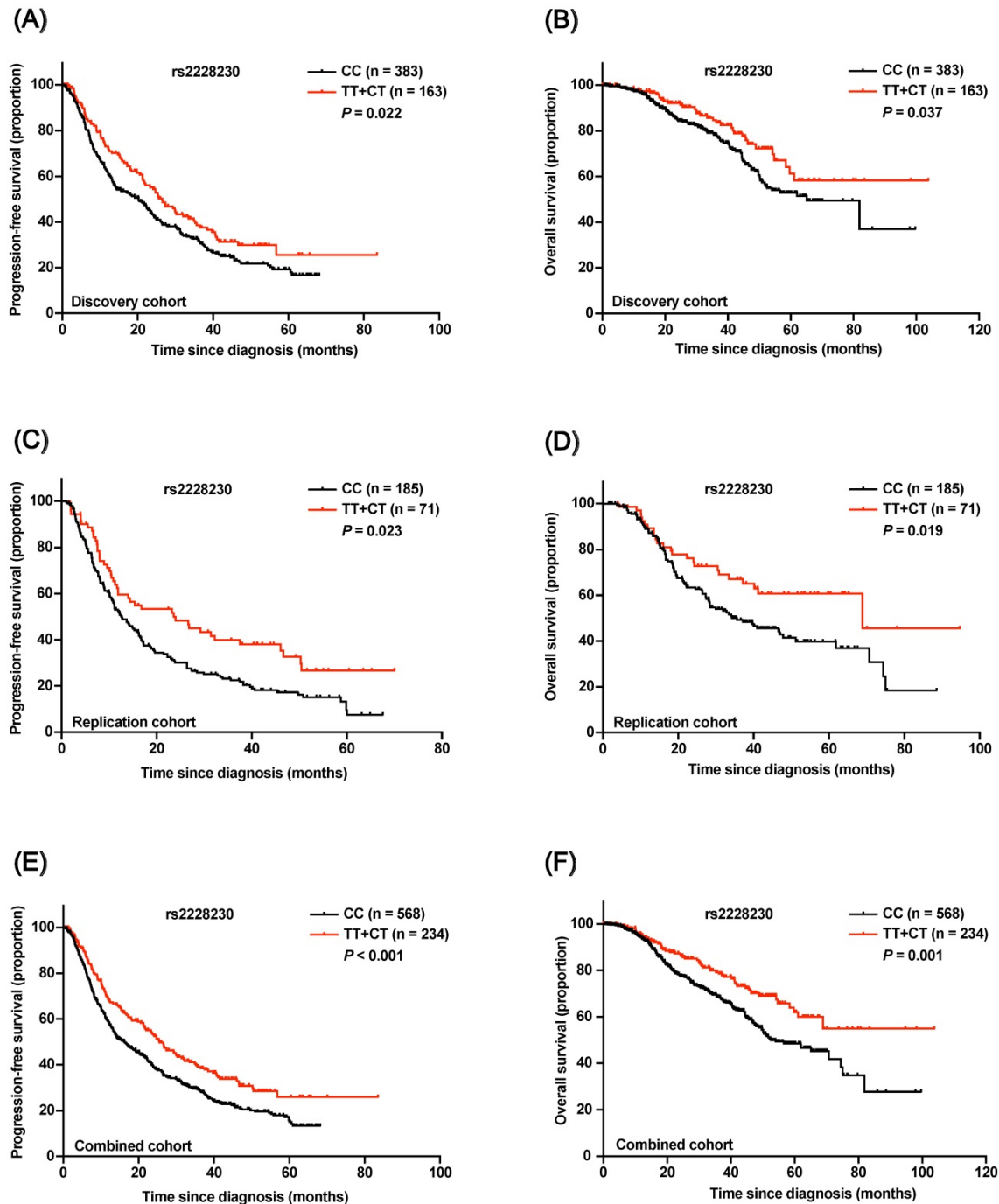
## Discussion

This study was conducted to investigate the functional effect of the tag SNP rs2228230 and its clinical significance in AM. rs2228230 is a synonymous SNP (sSNP), which is a specific type of polymorphism that does not change the encoded amino acid. sSNPs were recently demonstrated to alter mRNA and protein expression levels, structures, and functions by different mechanisms [23]. For example, the sSNP rs1045642 located in the coding sequence of *ABCB1* can reduce its mRNA expression

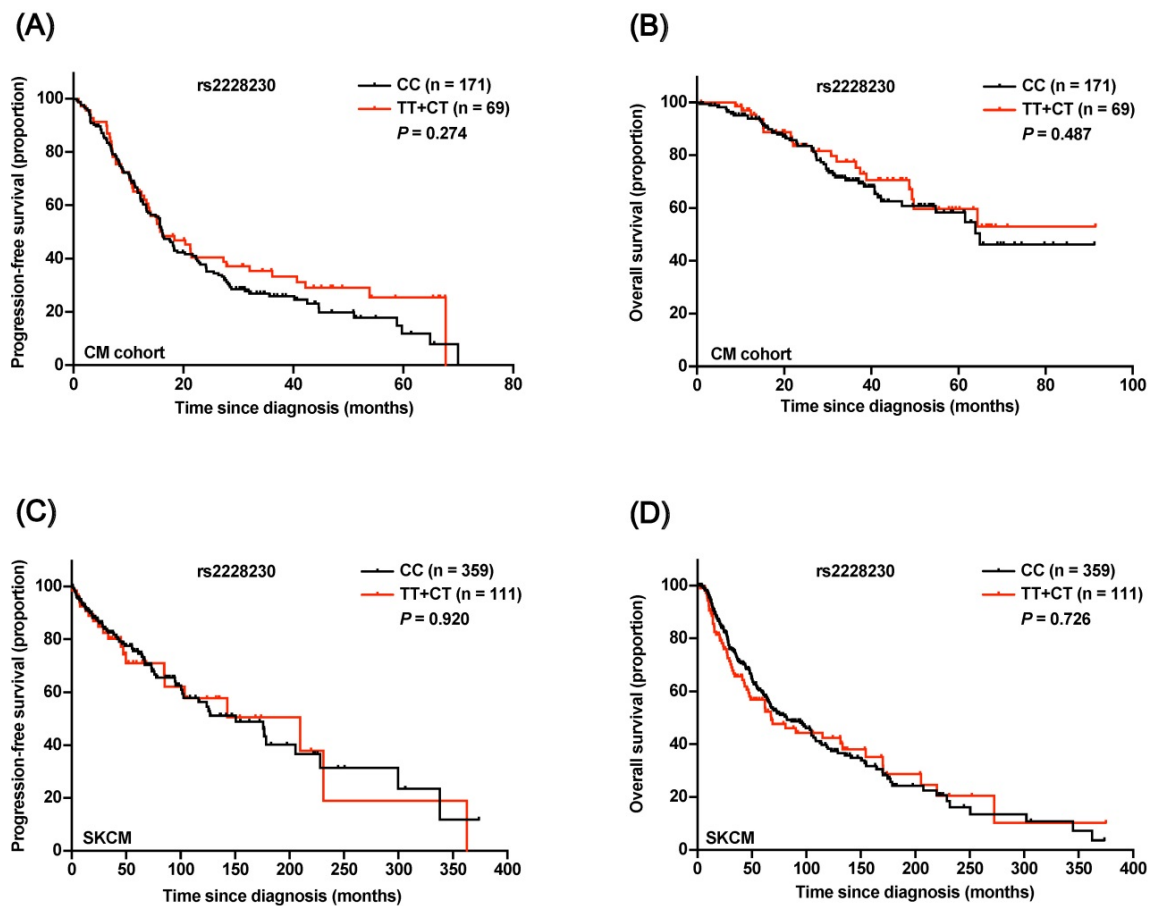


through biased codon usage, resulting in functional alteration of the substrate [24,25]. This *ABCB1* sNP has been reported to be associated with tacrolimus response in ulcerative colitis [26], imatinib response in chronic myeloid leukemia [27], and paclitaxel sensitivity in different cancer cells [28]. Our *in silico* and *in vitro* studies performed to clarify the effect of the rs2228230 genotype on *PDGFRA* expression demonstrated that the mRNA and protein expression decreased in cells transfected with the rs2228230:T

allele. mRNA secondary structure analysis revealed that the rs2228230:T allele changed the structure of *PDGFRA* and increased the MFE. A higher MFE implies lower mRNA stability, which might increase mRNA degradation. The *in silico* prediction was verified by an mRNA stability assay. After RNA synthesis was inhibited by ActD, *PDGFRA* mRNA decayed faster in cells with the rs2228230:T allele than in cells with the rs2228230:C allele.



**Figure 4.** Association of the rs2228230 genotype with prognosis in AM patients. Kaplan–Meier curve of (A) PFS and (B) OS according to the rs2228230 genotype in the discovery cohort. Kaplan–Meier curve of (C) PFS and (D) OS in the replication cohort. Kaplan–Meier plot of (E) PFS and (F) OS according to the rs2228230 genotype in the combined cohort.



**Figure 5.** Association of the rs2228230 genotype with prognosis in CM patients. Kaplan–Meier curve of (A) PFS and (B) OS according to the rs2228230 genotype in 240 CM patients. Kaplan–Meier curve of (C) PFS and (D) OS in the TCGA SKCM dataset.

The genetic code is degenerate, but synonymous codons are not used equally; the codon usage bias might alter the rate of translation elongation and therefore affect protein conformation and expression [29-32]. According to the codon usage values of four valine codons at the genome and *PDGFRA* gene level, we confirmed that codon usage was biased toward the major allele C, indicating that the translation elongation of *PDGFRA* would be slower in cells with the rs2228230:T allele. The subsequent experiment demonstrated that after proteolytic degradation was inhibited, the cells with the rs2228230:T allele showed a decreased rate of *PDGFRA* protein synthesis. Furthermore, the degradation rate of *PDGFRA* was significantly faster in cells with the rs2228230:T allele after protein synthesis was inhibited, indicating that the rs2228230:T allele also increased proteolytic degradation. Thus, the rs2228230 genotype could affect *PDGFRA* expression by regulating mRNA stability, protein synthesis, and degradation rates. Furthermore, rs2228230:T not only reduced *PDGFRA* expression but also decreased the activation of the MAPK and PI3K/AKT pathways.

Other molecular mechanisms by which sSNPs within coding sequences can affect protein expression

and function have been reported. First, coding-region sSNPs can affect splicing accuracy or efficiency by altering the binding of splicing regulatory proteins to ESS or ESE in exons [33,34]. Compared to the major allele, the minor T allele of rs2228230 reduces a serine- and arginine-rich splicing factor 2 (SRSF2)-binding site and increases a serine- and arginine-rich splicing factor 6 (SRSF6)-binding site [35,36]. Both SRSF2 and SRSF6 are arginine-rich proteins; however, SRSF2 is located only in the nucleus, whereas SRSF6 can shuttle between the nucleus and cytoplasm [37], indicating that besides altering mRNA stability, rs2228230 might also affect mRNA splicing. Second, sSNPs in coding sequences can also affect microRNA-binding sites. For example, the minor allele of rs10065172 reduces the binding capacity to miR-196 and therefore dysregulates the expression of *IRGM* in Crohn's disease [38]. Third, tRNA concentrations vary among organs, and differences in tRNA levels are a major determinant of translation elongation speed [39]. It has been demonstrated that the sSNP T2562G within *CFTR* alters the local translation speed and therefore changes *CFTR* stability and function by introducing a low-abundance tRNA in a bronchial epithelial tissue-specific manner [40].

Moreover, sSNPs might influence the protein folding process and result in misfolded protein and decreased expression thereof [41].

In this study, we found that the minor allele T of rs2228230 was associated with reduced mRNA and protein expression of *PDGFRA* in AM samples, and the following study in two independent cohorts demonstrated that AM patients with the rs2228230:T allele had longer PFS and OS compared to patients with the CC genotype, indicating that it may be used as a predictor of improved outcome. However, the rs2228230 genotype was not found to be correlated with survival in our CM cohort or the TCGA SKCM dataset. These results indicated that the prognostic value of rs2228230:T varies based on original site, which might be because UV radiation-induced genetic aberration is more extensive in CM, and other molecular variations could affect *PDGFRA* expression and signaling activation. Patterns of linkage disequilibrium differ across races/ethnicities [42,43], and the expression of *PDGFRA* might be epigenetically regulated [44]. Therefore, the prognostic value of rs2228230 in other populations must be further validated in prospective studies.

In conclusion, our study showed that the minor T allele of rs2228230 can reduce the expression and function of *PDGFRA* by altering the stability and synthesis of its mRNA and protein, and that it is associated with better survival in AM patients. These results highlight the potential of the rs2228230 genotype to serve as a prognostic marker in AM.

## Abbreviations

ActD: actinomycin D; AM: acral melanoma; CHX: cycloheximide; CM: cutaneous melanoma; ESE: exonic splicing enhancer; ESS: exonic splicing silencer; FFPE: formalin-fixed, paraffin-embedded; GWAS: genome-wide association study; HR: hazard ratio; HRP: horseradish peroxidase; IHC: immunohistochemical; MFE: minimum free energy; OS: overall survival; PDGF: platelet-derived growth factor; *PDGFRA*: PDGF receptor alpha; PFS: progression-free survival; qRT-PCR: quantitative reverse transcription polymerase chain reaction; RSCU: relative synonymous codon usage; SKCM: Skin Cutaneous Melanoma; SNP: single-nucleotide polymorphism; SRSF2: serine- and arginine-rich splicing factor 2; SRSF6: serine- and arginine-rich splicing factor 6; sSNP: synonymous SNP.

## Supplementary Material

Supplementary figures and tables.

<http://www.jcancer.org/v11p2945s1.pdf>

## Acknowledgments

We thank Qianzhaoxinye Bio Technology Pty Ltd (Beijing, China) for help with TCGA data analysis. This work was supported by grants from the Beijing Talents Foundation (2016000021223ZK18), National Natural Science Foundation of China (81672696, 81402264, 81972562), Beijing Natural Science Foundation (7154187), and Beijing Municipal Administration of Hospitals' Ascent Plan (DFL20181101).

## Competing Interests

The authors have declared that no competing interest exists.

## References

- Papadopoulos N, Lennartsson J. The PDGF/PDGF-R pathway as a drug target. *Mol Aspects Med.* 2018;62:75-88.
- Kazlauskas A. PDGFs and their receptors. *Gene.* 2017;614:1-7.
- Cools J, DeAngelo DJ, Gotlib J, et al. A tyrosine kinase created by fusion of the *PDGFRA* and *FIP1L1* genes as a therapeutic target of imatinib in idiopathic hypereosinophilic syndrome. *N Engl J Med.* 2003;348:1201-14.
- Heinrich MC, Corless CL, Duensing A, et al. *PDGFRA* activating mutations in gastrointestinal stromal tumors. *Science.* 2003;299:708-10.
- Puputti M, Tynninen O, Sihto H, et al. Amplification of *KIT*, *PDGFRA*, *VEGFR2*, and *EGFR* in gliomas. *Mol Cancer Res.* 2006;4:927-34.
- Dai J, Kong Y, Si L, et al. Large-scale analysis of *PDGFRA* mutations in melanomas and evaluation of their sensitivity to tyrosine kinase inhibitors imatinib and crenolanib. *Clin Cancer Res.* 2013;19:6935-42.
- Barnhill RL, Xiao M, Graves D, Antoniadis HN. Expression of platelet-derived growth factor (*PDGF*)-A, *PDGF*-B and the *PDGF*-alpha receptor, but not the *PDGF*-beta receptor, in human malignant melanoma in vivo. *Br J Dermatol.* 1996;135:898-904.
- Pfarr N, Penzel R, Klauschen F, et al. Copy number changes of clinically actionable genes in melanoma, non-small cell lung cancer and colorectal cancer-A survey across 822 routine diagnostic cases. *Genes Chromosomes Cancer.* 2016;55:821-33.
- Macgregor S, Montgomery GW, Liu JZ, et al. Genome-wide association study identifies a new melanoma susceptibility locus at 1q21.3. *Nat Genet.* 2011;43:1114-8.
- Choi J, Xu M, Makowski MM, et al. A common intronic variant of *PARP1* confers melanoma risk and mediates melanocyte growth via regulation of *MITF*. *Nat Genet.* 2017;49:1326-1335.
- Law MH, Rowe CJ, Montgomery GW, Hayward NK, Macgregor S, Khosrotehrani K. *PARP1* polymorphisms play opposing roles in melanoma occurrence and survival. *Int J Cancer.* 2015;136:2488-9.
- Ugurel S, Schrama D, Keller G, et al. Impact of the *CCR5* gene polymorphism on the survival of metastatic melanoma patients receiving immunotherapy. *Cancer Immunol Immunother.* 2008;57:685-91.
- Li H, Wang Y, Liu H, et al. Genetic variants of *PDGF* signaling pathway genes predict cutaneous melanoma survival. *Oncotarget.* 2017;8:74595-74606.
- Bai X, Kong Y, Chi Z, et al. *MAPK* Pathway and *TERT* Promoter Gene Mutation Pattern and Its Prognostic Value in Melanoma Patients: A Retrospective Study of 2,793 Cases. *Clin Cancer Res.* 2017;23:6120-6127.
- Xu Z, Taylor JA. SNPinfo: integrating GWAS and candidate gene information into functional SNP selection for genetic association studies. *Nucleic Acids Res.* 2009;37:W600-5.
- Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. *Genome Res.* 2002;12:996-1006.
- Tang X, Eberhart JK, Cleves MA, et al. *PDGFRA* gene, maternal binge drinking and obstructive heart defects. *Sci Rep.* 2018;8:11083.
- Fan Q, Zhou X, Khor CC, et al. Genome-wide meta-analysis of five Asian cohorts identifies *PDGFRA* as a susceptibility locus for corneal astigmatism. *PLoS Genet.* 2011;7:e1002402.
- Garrigós C, Espinosa M, Salinas A, et al. Single nucleotide polymorphisms as prognostic and predictive biomarkers in renal cell carcinoma. *Oncotarget.* 2017;8:106551-106564.
- Zhang T, Choi J, Kovacs MA, et al. Cell-type-specific eQTL of primary melanocytes facilitates identification of melanoma susceptibility genes. *Genome Res.* 2018;28:1621-1635.
- Sharp PM, Li WH. The codon Adaptation Index--a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res.* 1987;15:1281-95.
- Guo N, Peng Z. MG132, a proteasome inhibitor, induces apoptosis in tumor cells. *Asia Pac J Clin Oncol.* 2013;9:6-11.
- Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet.* 2011;12:683-91.



24. Wang D, Johnson AD, Papp AC, Kroetz DL, Sadée W. Multidrug resistance polypeptide 1 (MDR1, ABCB1) variant 3435C>T affects mRNA stability. *Pharmacogenet Genomics*. 2005;15:693-704.
25. Kimchi-Sarfaty C, Oh JM, Kim IW, et al. A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science*. 2007;315:525-8.
26. Herlinger KR, Koc H, Winter S, et al. ABCB1 single-nucleotide polymorphisms determine tacrolimus response in patients with ulcerative colitis. *Clin Pharmacol Ther*. 2011;89:422-8.
27. Ni LN, Li JY, Miao KR, et al. Multidrug resistance gene (MDR1) polymorphisms correlate with imatinib response in chronic myeloid leukemia. *Med Oncol*. 2011;28:265-9.
28. Kwon WS, Rha SY, Jeung HC, et al. G-T haplotype (2677G>T/A and 3435C>T) of ABCB1 gene polymorphisms is associated with ethnic differences to paclitaxel sensitivity in cancer cells with different gene expression pattern. *Cancer Lett*. 2009;277:155-63.
29. Chamary JV, Parmley JL, Hurst LD. Hearing silence: non-neutral evolution at synonymous sites in mammals. *Nat Rev Genet*. 2006;7:98-108.
30. Zhou M, Guo J, Cha J, et al. Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature*. 2013;495:111-5.
31. Kirchner S, Cai Z, Rauscher R, et al. Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol*. 2017;15:e2000779.
32. Xie T, Ding D. The relationship between synonymous codon usage and protein structure. *FEBS Lett*. 1998;434:93-6.
33. Cartegni L, Chew SL, Krainer AR. Listening to silence and understanding nonsense: exonic mutations that affect splicing. *Nat Rev Genet*. 2002;3:285-98.
34. Supek F, Miñana B, Valcárcel J, Gabaldón T, Lehner B. Synonymous mutations frequently act as driver mutations in human cancers. *Cell*. 2014;156:1324-1335.
35. Cartegni L, Wang J, Zhu Z, Zhang MQ, Krainer AR. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res*. 2003;31:3568-71.
36. Smith PJ, Zhang C, Wang J, Chew SL, Zhang MQ, Krainer AR. An increased specificity score matrix for the prediction of SF2/ASF-specific exonic splicing enhancers. *Hum Mol Genet*. 2006;15:2490-508.
37. Jeong S. SR proteins: binders, regulators, and connectors of RNA. *Mol Cells*. 2017;40:1-9.
38. Brest P, Lapaquette P, Souidi M, et al. A synonymous variant in IRGM alters a binding site for miR-196 and causes deregulation of IRGM-dependent xenophagy in Crohn's disease. *Nat Genet*. 2011;43:242-5.
39. Sauna ZE, Kimchi-Sarfaty C. Understanding the contribution of synonymous mutations to human disease. *Nat Rev Genet*. 2011;12:683-91.
40. Kirchner S, Cai Z, Rauscher R, et al. Alteration of protein function by a silent polymorphism linked to tRNA abundance. *PLoS Biol*. 2017;15:e2000779.
41. Nicola AV, Chen W, Helenius A. Co-translational folding of an alphavirus capsid protein in the cytosol of living cells. *Nat Cell Biol*. 1999;1:341-5.
42. Foster MW, Sharp RR. Race, ethnicity, and genomics: social classifications as proxies of biological heterogeneity. *Genome Res*. 2002;12:844-50.
43. Li H, Teo YY, Tan EK. Patterns of linkage disequilibrium of LRRK2 across different races: implications for genetic association studies. *PLoS One*. 2013;8:e75041.
44. Ruiz-Camp J, Quantius J, Lignelli E, et al. Targeting miR-34a/Pdgfra interactions partially corrects alveologenesis in experimental bronchopulmonary dysplasia. *EMBO Mol Med*. 2019;11:e9448.