medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Detection of Somatic Copy Number Deletion of *CDKN2A* Gene for Clinical Practices Based on Discovery of A Base-Resolution Common Deletion Region

Running title: Detection of somatic *CDKN2A* deletion

Yuan Tian[1#], Rui Xing[2#], Jing Zhou[1#], Juanli Qiao[1], Zhaojun Liu[1], Liankun Gu[1], Baozhen Zhang[1], Jiafu Ji[3], Youyong Lu[2], Dajun Deng[1]

[1] Key Laboratory of Carcinogenesis and Translational Research (MOE/Beijing), Division of Etiology, Peking University Cancer Hospital and Institute, Beijing, 100142, China;

[2] Key Laboratory of Carcinogenesis and Translational Research (MOE/Beijing), Division of Tumor Biology, Peking University Cancer Hospital and Institute, Beijing, 100142, China;

[3] Key Laboratory of Carcinogenesis and Translational Research (MOE/Beijing), Division of Surgical Oncology, Peking University Cancer Hospital and Institute, Beijing, 100142, China

Equal first authors

Corresponding author: Dajun Deng, Key Laboratory of Carcinogenesis and Translational Research (MOE/Beijing), Division of Etiology, Peking University Cancer Hospital and Institute, Fu-Cheng-Lu #52, Haidian District, Beijing, 100142, China, Phone:+8610-88196752 (E-mail: dengdajun@bjmu.edu.cn)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

ABSTRACT

CDKN2A is most frequently inactivated by somatic copy number deletion (SCND) in human cancers. An accurate method to detect *CDKN2A* SCND is absent. We extracted interstial base-resolution deletion/fusion coordinates for *CDKN2A* from published articles and our whole genome sequencing (WGS) datasets, and characterized a 5.1-kb common deletion region (CDR) from the *CDKN2A/P16*^{INK4A} promoter to intron-2 in 90% (83/92) of cancer cell lines and 100% (18/18) of gastric carcinoma (GC) samples by WGS. Then, a highly reproducible and quantitative assay (P16-Light) targeted to the CDR was setup with a detection limit of 20%. P16-Light was further compared with WGS (gold standard) and loss of heterozygosity (LOH) assay by denaturing high performance liquid chromatography. The positive rate of *CDKN2A* SCND by P16-Light was more than two times of that by WGS in GCs from a patients cohort (n=156; 25.6% vs. 10.9%), and was significantly correlated with each other (p<0.028). While *CDKN2A* SCND was significantly associated with GC distant metastasis (p=0.023), such association was not observed for *CDKN2A* LOH (p=0.213). In conclusion, based on our finding that there is a 5.1-kb CDR within *CDKN2A* gene, a vary convenient P16-Light assay was setup to detect *CDKN2A* SCND for clinical applications.

Key words

CDKN2A; common deleted region (CDR); whole genome sequencing (WGS), LOH, gene copy number variations

Key points

- We characterized a base resolution common deletion region (CDR) within *CDKN2A* in human cancers;
- A novel multiplex PCR method was developed to detect somatic copy number deletion of *CDKN2A*;
- The method is very convenient for clinical applications with a high reproducibility and sensitivity.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

INTRODUCTION

Somatic copy number deletion (SCND) of *CDKN2A* gene is a landmark of human cancers (1,2). The frequency of *CDKN2A* SCND detected by single nucleotide polymorphism (SNP) microarray, whole genome sequencing (WGS) or whole exome sequencing (WES) was found to be ranged from 30% to 60% in bladder cancer, melanoma, head and neck cancer, pleural mesothelioma, glioblastoma, and esophageal squamous cell cancer (ESCC) with an average frequency of 13% in pan-cancer datasets in The Cancer Genome Atlas (TCGA) (Supplemantary Figure S1A) (2-6). *CDKN2A* deep-deletion is associated with downregulation of *CDKN2A* gene expression, while *CDKN2A* amplification is associated with upregulation of *CDKN2A* gene expression in Pan-TCGA cancers (Supplemantary Figure S1B).

It is well known that genetic *CDKN2A* inactivation contributes to malignant transformation, cancer metastasis, and therapy sensitivity of cancers to drugs, including CDK4/6 inhibitors and their combination with PD-1 blockades (7-11). However, current gene copy number detection methods, including fluorescence-in-situ hybridization (FISH) and WGS or WES, are not sensitive enough or too costly for routine clinical use. While the amplification of oncogenes (such as *EGFR*, c-*ERBB2,* c-*MYC*, and c-*MET*) are increasingly driving decision-making for precise cancer treatments, clinical applications of SCND of tumor suppressor genes, including *CDKN2A*, are still rare owing to the lack of a feasible detection assay.

METHODS

Patients, tissues, and DNA preparation

Frozen fresh GC and paired SM tissue samples were collected from 140 patients in Beijing Cancer Hospital and 156 patients in the WGS study (12). Clinicopathologic information were also obtained. The 2010 UICC-tumor-node-metastasis (TNM) system was used for the classification of these GCs (13). Genomic DNA was extracted from these samples with a phenol/chloroform method. The Institution Review Board of Peking University Cancer Hospital & Institute approved this study and was carried out in accordance with the principles outlined in the Declaration of Helsinki. Informed consent was obtained from each patient prior to their inclusion in the study.

Optimized quantitative multiplex PCR assay (P16-Light) to detect *CDKN2A* copy number

A number of multiplex primer and probe combinations were designed based on the best multiplex primer probe scores for the CDR in the *CDKN2A* gene and *GAPDH* sequences by Bacon Designer 8 software. Multiplex PCR assays were established according to the Applied Biosystems (ABI) TaqMan universal PCR

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

master mix manual. The performances of these assays for the detection of *CDKN2A* copy numbers were compared with each other. Finally, a multiplex primer and probe combination was selected (Table 1) and their components' concentrations were optimized. Briefly, each multiplex PCR assay was carried out in a total of 20 μL volume that included 5-10 ng of input DNA, 10 μM of forward and reverse primers and probe for *CDKN2A* intron-2, 10 μM forward and reverse primers and probe for *GAPDH*, and 10 μL of 2 x TaqMan Universal Master Mix II with uracil-N-glycosylase (Kit-4440038, ABI, Lithuania). The PCRs were performed in triplicate in a MicroAmp Fast Optical 96-Well Reaction Plate with a barcode (0.1 mL; ABI, China) with an ABI 7500 Fast Real-Time PCR System. The specific conditions of the PCR were as follows: initial incubation for 10 min at 95°C, followed by 40 cycles of 95°C for 20 sec and 58°C for 60 sec. When Ct value for *GAPDH* input for a sample was 34 or less cycles, this sample was considered as *CDKN2A* SCNV informative.

Table 1. Oligo sequences

Gene/STS	Assay	Oligo	Sequence (5'-3')	PCR product size
CDKN2A	P16-Light	F-primer	caggtctgtttcctcatttg	129-bp
		R-primer	ggtcagattagttgagttgtg	
		Probe	FAM-ctggctggaccaacctcagg-BHQ1	
GAPDH	P16-Light	F-primer	gctcacatattctggaggag	135-bp
		R-primer	ggtcattgatggcaacaata	
		Probe	Cy5-tgccttcttgcctcttgtctctt-BHQ2	
D9S941	MSI/DHPLC	Forward	aaaaaaaggaagggaggagg	205-bp
		Reverse	acaccggatgaagtcaccta	
D9S1749	MSI/DHPLC	Forward	aggagagggtacgcttgcaa	120-bp
		Reverse	tacagggtgcgggtgcagataa	
D9S974	MSI/DHPLC	Forward	gagcctggtctggatcataa	190-bp
		Reverse	aagcttacagaaccagacag	
D9S942	MSI/DHPLC	Forward	gcaagattccaaacagta	100-bp
		Reverse	ctcatcctgcggaaaccatt	
D9S1748	MSI/DHPLC	Forward	cacctcagaagtcagtgagt	130-bp
		Reverse	gtgcttgaaatacacctttcc	
D9S171	MSI/DHPLC	Forward	agctaagtgaacctcatctctgtct	158-177-bp
		Reverse	accctagcactgatggtatagtct	

Definitions of *CDKN2A CDR* deletion-positive and amplification-positive

We used the genomic DNA from A549 cells containing no *CDKN2A* allele to dilute genomic DNA from RKO cells containing 2 wild-type *CDKN2A* alleles, and then we set the standard curve according to the relative copy number of the *CDKN2A* gene at different dilution concentrations. The ΔCt value and relative copy number for the *CDKN2A* gene were calculated using the *GAPDH* as the internal reference. When the

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

CDKN2A copy number in the A549-diluted template was consistently lower than the copy number in the RKO control template, and the difference was statistically significant ($p<0.05$), it was judged that the lowest dilution concentration was the detection limit of *CDKN2A* deletion (the difference in *CDKN2A* copy number between the 100% RKO template and 80% RKO template spiked with 20% A549 DNA). When the *CDKN2A* relative copy number in a tissue sample was significantly lower or higher than that of the paired blood sample, the sample was defined as somatic *CDKN2A CDR* deletion-positive or amplification-positive, respectively. The 100% A549, 100% RKO, and 20% A549 + 80% RKO DNA mix controls were analyzed for each experiment.

Call for *CDKN2A* interstitial deletion/fusion in the GC WGS datasets

We used Meerkat (http://compbio.med.harvard.edu/Meerkat/) 23 to predict somatic SVs and their breakpoints in WGS datasets (accession numbers, EGAD00001004811 with 36× of sequencing depth) for gastric adenocarcinoma samples from 168 patients using the suggested parameters (12). This method used soft-clipped and split reads to identify candidate breakpoints, and precise breakpoints were refined by local alignments. *CDKN2A* deletion information was obtained from WGS datasets for 157 GC samples.

Detection of MSI by DHPLC

The microsatellite-specific primers were used to amplify the target microsatellite at the corresponding annealing temperature (Table 1). Input DNA template (5 ng) were used per PCR interaction. The microsatellite PCR products were separated with a DNASep Analytic Column using the DHPLC platform (WAVE™ DNA Fragment Analysis System, Transgenomic, Inc., Omaha, USA; FL-detector) at 50 °C to determine the MSI status of target locus according to the difference in the peak composition and peak area ratio of each microsatellite locus between GC and paired SM samples. Three microsatellites (*D9S974*, *D9S942*, and *D9S1748*) within the *P16* CDR closely flanking fragment were used as markers that represented the LOH/MSI in the CDR (14-16). The criteria for MSI were as follows: a) in the paired test result, if the detecting peak disappeared, it was directly determined that MSI occurs; b) if the peak height/area of the detecting waveform corresponding to the cancer tissue was reduced by >30%, it was determined that MSI occurs; c) when the two detection peaks coincided together, it was defined as not informative (17-19). When MSI at least one of three *CDKN2A*-CDR flanking microsatellites was detected in a sample relative to the normal control, the sample was defined as *CDKN2A* LOH (or gain of heterozygosity, GOH)-positive.

Cell lines and cultures

The human gastric cancer cell line MGC803 (kindly provided by Dr. Yang Ke at Peking University Cancer

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Hospital) and the *P16* allele homogygously deleted cell line A549 (kindly provided by Dr. Zhiqian Zhang of Peking University Cancer Hospital and Institute) were grown in RPMI-1640 medium, and the RKO cell line containing two wild type *P16* alleles was purchased from American Type Culture Collection and grown in DMEM media. The medium was supplemented with 10% (v/v) fetal bovine serum (FBS). These cell lines were tested and authenticated by Beijing JianLian Genes Technology Co., Ltd. before they were used in this study. STR patterns were analyzed using a Goldeneye™ 20A STR Identifiler PCR Amplification kit.

Statistical analysis

Chi-square or Fisher's exact tests were used to compare the proportion of *CDKN2A* SCND or amplification between different groups of tissue samples. Student t-test was used to compare the proportion of *CDKN2A* gene copy number between genomic DNA samples. All statistical tests were two-sided, and *p* value less than 0.05 was considered to be statistically significant.

RESULTS

Characterization of a base-resolution *CDKN2A* common deletion region (CDR) in human cancers

It has been previously reported that a homozygous deletion of approximately 170 kilobase pairs (kb), including the *CDKN2A* locus, can be detected in human cancers by microsatellite instability (MSI) analyses (20). To characterize the base-resolution genomic coordinates of *CDKN2A* deletions in cancers, we extracted sequence information of interstitial *CDKN2A* deletions from all available published articles and our sequencing data (Supplemantary Table S1) (21-31). We found a 5.1-kb CDR (chr9: 21,970,277 - 21,975,386, hg19) that spanned from the $P16^{INK4a}$ promoter to intron-2 in 83 (90%) of 92 reported cancer cell lines or tissue samples containing interstitial *CDKN2A* deletions (Figure 1, blue lines). This CDR sequence is exactly the same as the *CDKN2A* deletion fragment in the HCC193 lung cancer cell line (28). The CDR coordinates were also confirmed in our WGS datasets (average sequencing depth, 36×) of 18 (100%) of 18 GCs (12), in which interstitial *CDKN2A* deletions/fusions were identified (Figure 1, purple lines; Supplemantary Table S1).

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

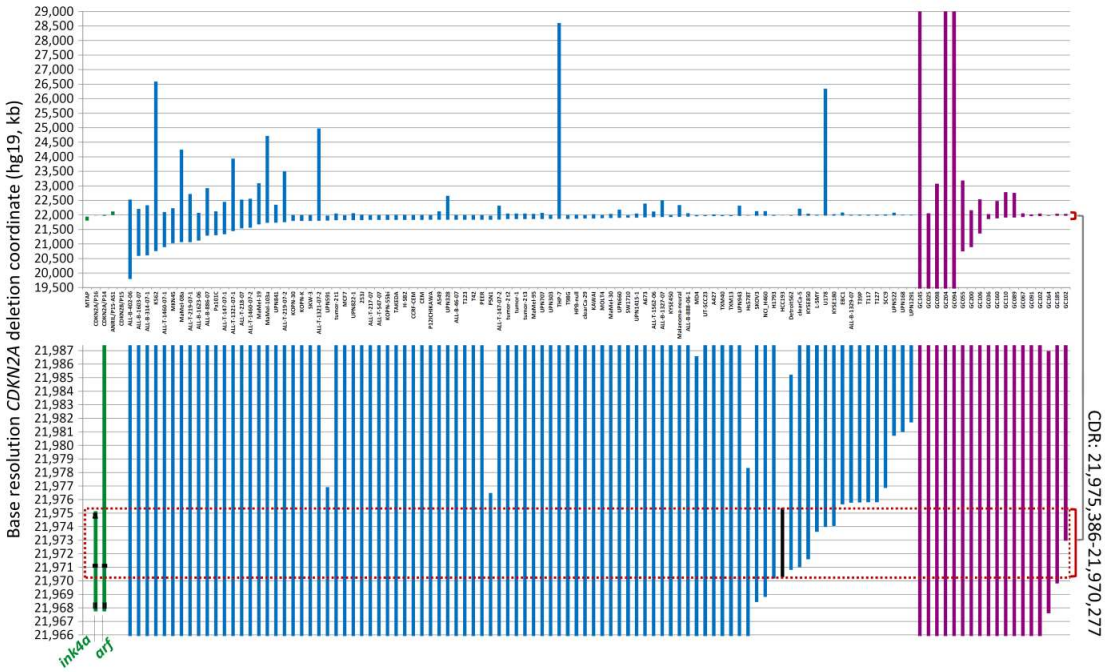

Figure 1. Base resolution of the genomic coordinates of interstitial *CDKN2A* deletion/fusion in human cancers. The top chart displays the coordinates of whole deleted fragments by Sanger sequencing of chromosome 9p21 in *CDKN2A*-deleted cancer cell lines and tissue samples (*n*=92, blue lines), and in gastric carcinomas (GC, *n*=18, purple lines) by whole genome sequencing (WGS). The sample ID is labeled under each column. The bottom chart displays an amplified view of the *CDKN2A* gene, where the 5.1-kb common deletion region (CDR) occurred from the $P16^{INK2a}$ promoter to intron-2 (highlighted with a red dashed line rectangle), which is exactly the same region as the deleted *CDKN2A* fragment in the HCC193 lung cancer cell line (highlighted with a black line). The detailed deletion coordinates for each sample are listed in Supplemantary Table S1. Each line represents a *CDKN2A* deletion fragment. The locations of $P16^{INK4a}$ and $P14^{ARF}$ (green lines) and exon-1α/2/3 (black dots) are also labeled as landmarks.

It is well known that germline *CDKN2A* inactivation can lead to a high predisposition for melanoma and pancreatic cancer (32-34). Interestingly, we found that 14 (93.3%) of 15 *CDKN2A* allelic variants in the Online Mendelian Inheritance in Man (OMIM) database are located within the CDR sequence (Supplemantary Figure S2) (35,36). These phenomena indicate that both inherited and somatic defects of the *CDKN2A CDR* region are cancer drivers.

Both $P16^{INK4a}$ and $P14^{ARF}$ mRNAs are transcribed from the human *CDKN2A* gene at chromosome 9p21, but with different transcription start sites (37); they share the same exon-2 but have different translation reading frames. Because *CDKN2A* exon-2 located within the CDR is the essential exon for coding $P16^{INK4a}$ and $P14^{ARF}$ proteins, the above findings indicate that $P16^{INK4a}$ and $P14^{ARF}$ are co-inactivated in 87% (96/110) of human cancer cell lines and tissues containing *CDKN2A CDR* deletion (Figure 1).

Establishment of a convenient PCR assay (P16-Light) to detect somatic *CDKN2A CDR* deletion

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

The current clinical method FISH for detecting SCND is composed of a set of biotin-labelled probes which should cover at least 50-kb DNA sequence. Thus, FISH is not a suitable method for detecting the copy number of the 5.1-kb *CDKN2A* CDR. To provide a convenient method for routine clinical use, we designed and experimentally evaluated a set of multiplex quantitative PCR assays and finally optimized the *CDKN2A* CDR-specific quantitative multiplex PCR assay called P16-Light for detecting the copy number of a 129-bp amplicon within *CDKN2A/P16*^{INK4a} intron-2 (Figure 2A), which covers 86% (94/110) of known *CDKN2A* deletion fragments (Figure 1).

The copy number of *GAPDH* gene was used as the internal reference (Table 1). Genomic DNA from human A549 cells (with homozygous deletion of *CDKN2A* alleles) and RKO cells (with 2 wild-type *CDKN2A* alleles) were used as *CDKN2A* CDR deletion-positive and deletion-negative controls, respectively. The proportions of *CDKN2A* CDR copy number were linearly correlated with the ratios (0 - 100%) of RKO cell DNA and A549 cell DNA in the input mixtures when the A549 DNA was spiked in at different proportions for the P16-Light analyses (Figure 2B). Furthermore, there was a high reproducibility when DNA with homozygous deletion of *CDKN2A* was present in ≥20% of the cells, which were verified in ten experimental repeats that were performed on different days (Figure 2C). Thus, when the proportion of *CDKN2A* copy number was significantly decreased (or increased) in a sample relative to the paired normal control (t-test, $p<0.05$) in the P16-Light analyses, the sample would be defined as *CDKN2A* SCND positive (or amplification positive).

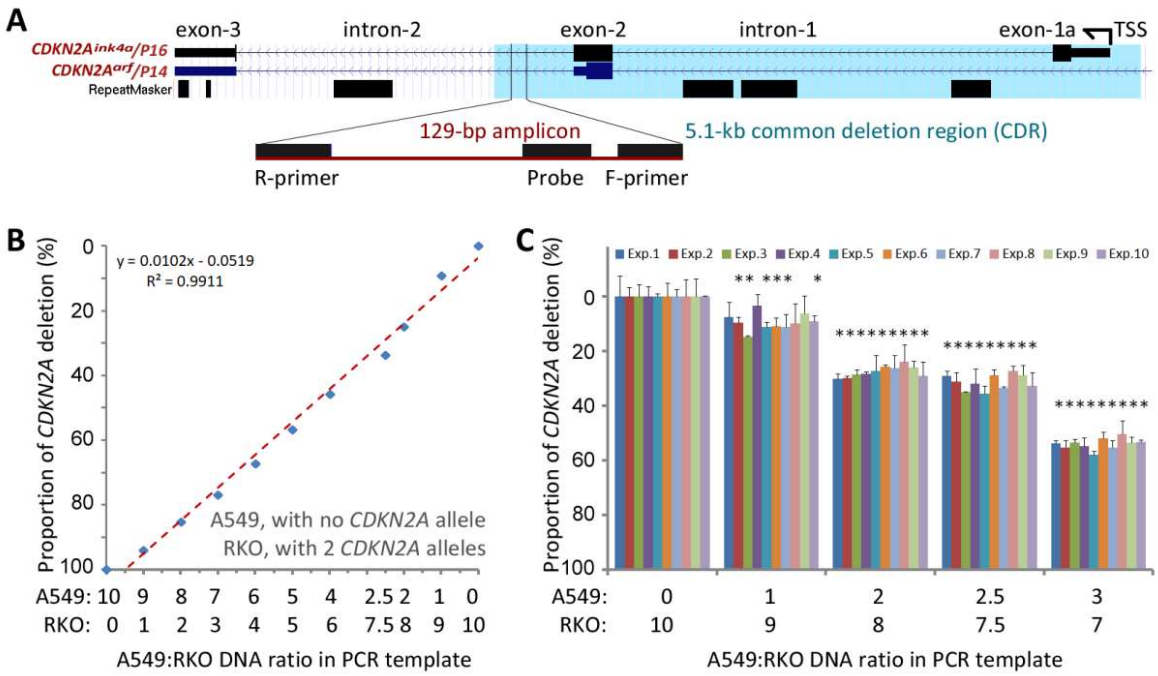

Figure 2. Detection of the copy number of *CDKN2A* intron-2 with quantitative gene-specific multiplex

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

PCR (P16-Light). (**A**) The location of the 129-bp amplicon within the common deletion region (CDR) and its host genes. (**B**) The linear relationship between the proportion of *CDKN2A* copy number deletion and ratios of RKO cells (with two wild-type *CDKN2A* alleles) spiked with A549 cells (with a homozygous *CDKN2A* deletion). (**C**) Stability of the proportion of the *CDKN2A* copy number deletion by P16-Light during ten experiments over different days. The RKO cell DNA templates were spiked with 0, 10%, 20%, 25%, and 30% A549 cell DNA. Each column represents the average proportion of *CDKN2A* copy number deletion in triplicate. Exp. 1 - 10: the results of 10 repeated experiments performed on different days. *: compared to the 100% RKO control in chi-square test, $p<0.05$.

Comparison of P16-Light with WGS

As we described above, information on interstitial copy number deletion/fusion of *CDKN2A* gene was extracted from WGS datasets for 157 of 168 GC patients enrolled in a GC genome study and a total of 18 *CDKN2A* deletion/fusion coordinates at the base-resolution were detected in 17 (10.8%) GCs (Supplemantary Table S1). To compare the performance of P16-Light with WGS, we analyzed the status of somatic copy number variations (SCNVs), including SCND and amplification, of *CDKN2A* in 156 of these GCs with enough amounts of genomic DNA samples with P16-Light using the paired surgical margin tissue (SM) as the diploid reference. *CDKN2A* SCND and amplification were detected in 40 (25.6%) and 34 (21.8%) of these GCs, respectively. The P16-Light analysis was confirmed by the WGS results: the frequency of *CDKN2A* SCND (or amplification) by P16-Light was significantly higher (or lower) in 17 GCs containing interstitial *CDKN2A* deletion/fusion than those 139 GCs without (chi-square test, $p<0.028$; Figure 3). These results suggest that there is a significantlyt higher sensitivity for detecting *CDKN2A* SCND by the quantitative P16-Light assay than the hemi-quantitative WGS.

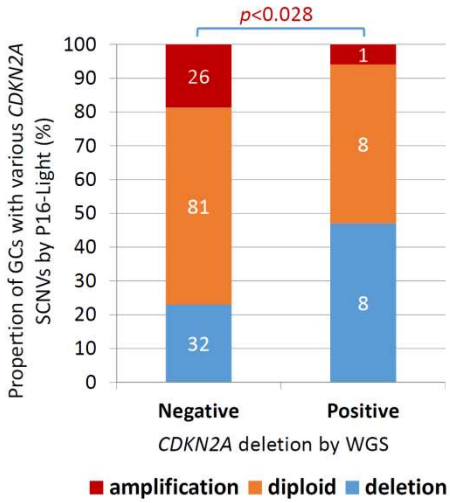

Figure 3. Comparisons of somatic copy number variations (SCNVs) of *CDKN2A* gene in gastric carcinoma samples (n=156) relative to the paired normal control in the P16-Light and WGS (30×) analyses

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Comparison of P16-Light with LOH/MSI assay

Denaturing high performance liquid chromatography (DHPLC) was a rapid sensitive assay, which could be used to detect loss of heterozygosity (LOH) of genes (17-19). To compare P16-Light with a sensitive method, we established a DHPLC assay coupled with post-column double-strand DNA fluorescence-labeling device and a fluorescence detector (38). The fragment separation mode at 50 °C was used to detect microsatellite instability (MSI) of previously reported microsatellite loci within and outside of *CDKN2A* gene: 3 *CDKN2A*-CDR flanking microsatellites (*D9S974*, *D9S942*, and *D9S1748*) and 3 non-*CDKN2A* microsatellites (*D9S941*, *D9S1749*, and *D9S171*) (Figure 4A). The status of MSI of the representative locus *D9S974* in three cell lines (AGS, HeLa, and MGC803) could be stably confirmed with clone-sequencing (Supplemantary Figure S3). Thus, we used the DHPLC assay to detect MSI at these microsatellite loci in GCs from another set of patients (n=140) using the paired SMs as the normal control (Figure 4B). The frequency of *CDKN2A* SCND by P16-Light was 38.6% in these GCs, while the MSI frequency of these loci by DHPLC were 22.9% for *D9S974*, 16.4% for *D9S942*, 17.9% for *D9S1748*, 32.9% for *D9S941*, 21.4% for *D9S1749*, and 10% for *D9S171*, respectively. And the positive rate of *CDKN2A* LOH was 35.7%. Both the frequency of *CDKN2A* LOH and SCND was significantly increased in GCs at advanced stages (p=0.017 and 0.036; Table 2). Notably, while *CDKN2A* SCND was significantly associated with distant metastasis of GCs (p=0.023), such association was not observed for *CDKN2A* LOH (p=0.213).

In addition, the frequency of *CDKN2A* SCND was higher in GCs with MSI of any of these 6 microsatellites or *CDKN2A* LOH than those without MSI (chi-square test, p=0.021 for *D9S941*; Figure 4C). More *CDKN2A* amplification was also detected in GCs with MSI at *D9S1748* than in those without (p<0.001; Figure 4D), consistenting with that MSI represents both copy number deletion and amplification of host genes. These results suggest the advantage of P16-Light over DHPLC-LOH (or GOH) for detection of *CDKN2A* SCNVs as the prognosis biomarker.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

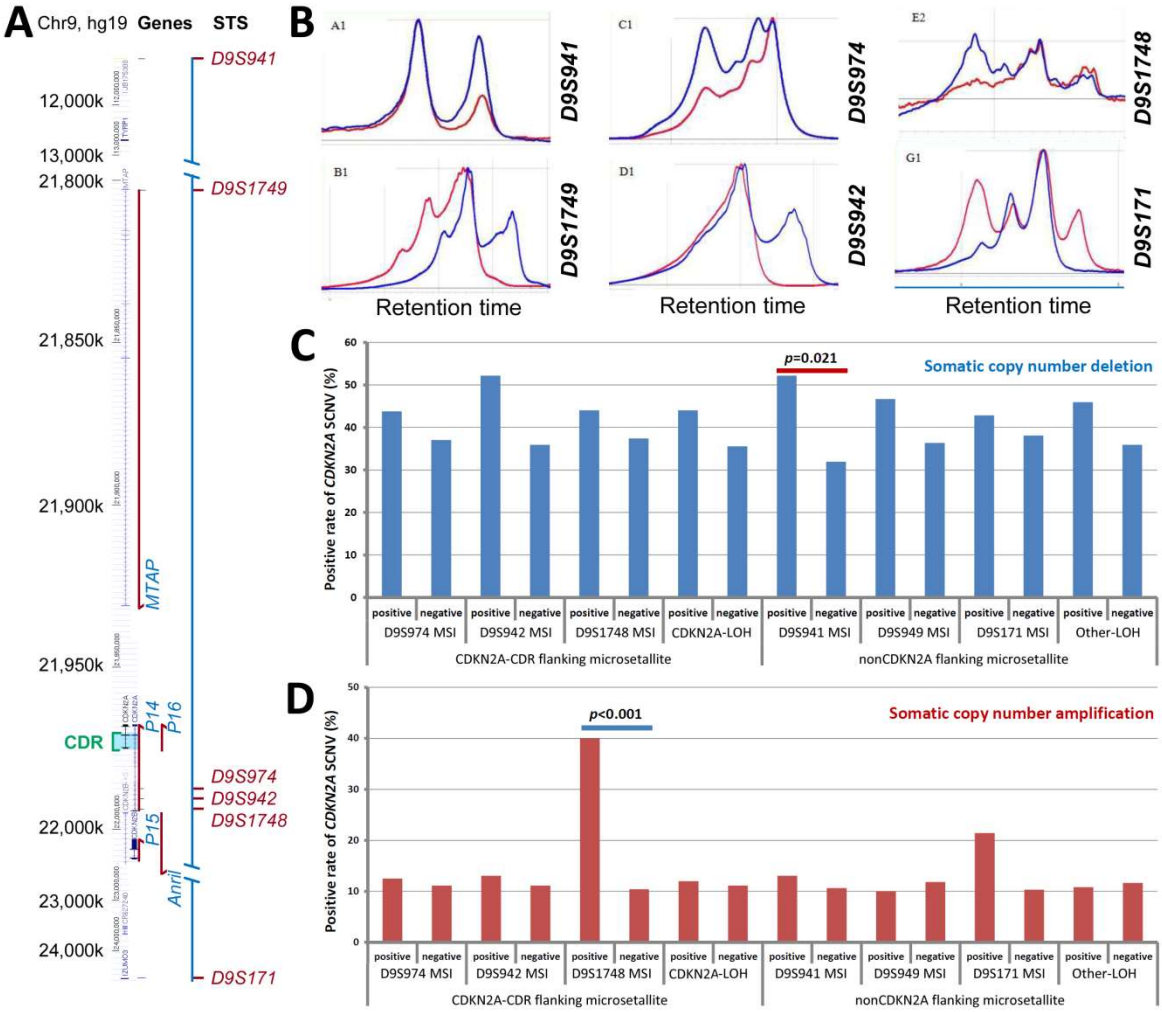

Figure 4. The results of the microsatellite instability (MSI) of the *CDKN2A* gene and neighboring regions in gastric carcinoma samples by three common deletion region (CDR) flanking microsatellites (*D9S974*, *D9S942*, and *D9S1748*) within the *CDKN2A* gene and three non-*CDKN2A* microsatellites (*D9S941*, *DS749*, and *D9S171*) outside of the *CDKN2A* gene. (**A**) The graphic views for the locations of the 5.1-kb CDR and microsatellites at 9p21 are illustrated. Each red arrowed line represents a gene and its transcription direction. (**B**) The informative chromatograms for the PCR products of various microsatellites in gastric cancer and the paired normal tissues from a representative patients with denaturing high-performance liquid chromatography (DHPLC, sizing model) at 50°C. Blue line: gastric carcinoma tissue; red line: paired surgical margin tissue. (**C**) The positive rate of somatic *CDKN2A* copy number deletion in informative GCs with and without LOH of 6 microsatellites and their combination (Anylocus); (**D**) The positive rate of somatic *CDKN2A* copy number amplification in 140 informative GCs with and without LOH of 6 microsatellites and their combination. Statistically significant *p*-values in chi-square tests were also labeled

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Table 2. Association between clinicopathological characteristics and instability at various microsatellite loci within and not within the *CDKN2A* gene by DHPLC and somatic copy number variations of *CDKN2A* gene by P16-Light

		n	***CDKN2A***	Microsatellite instability (MSI) by DHPLC (%)						***CDKN2A* SCNV by P16-Light (%)**		
			LOH (%)	*D9S941*	*D9S1749*	*D9S974*	*D9S942*	*D9S1748*	*D9S171*	**Amplified**	**Diploid**	**Deleted**
Age	<60	70	23 (32.9)	19 (27.1)	14 (20.0)	12 (17.1)	12 (17.1)	13 (18.6)	7 (10.0)	7 (10.0)	38 (54.3)	25 (35.7)
	≥60	70	27 (38.6)	27 (38.6)	16 (22.9)	20 (28.6)	11 (15.7)	12 (17.1)	7 (10.0)	9 (12.9)	32 (45.7)	29 (41.4)
Sex	Male	104	36 (34.6)	35 (33.7)	20 (19.2)	26 (25.0)	16 (15.4)	16 (15.4)	9 (8.7)	13 (12.5)	49 (47.1)	42 (40.4)
	Female	36	14 (38.9	11 (30.6)	10 (27.8)	6 (16.7)	7 (19.4)	9 (25.0)	5 (13.9)	3 (8.3)	21 (58.3)	12 (33.3)
Location[a]	Proximal	71	26 (36.6)	**30 (42.3)**[b]	17 (23.9)	16 (22.5)	12 (16.9)	11 (15.5)	6 (8.5)	7 (9.9)	37 (52.1)	27 (38.0)
	Distal	69	24 (34.8)	**16 (23.3)**	13 (18.8)	16 (23.2)	11 (15.9)	14 (20.3)	8 (11.6)	9 (13.0)	33 (47.8)	22 (31.9)
Different.	Poor	102	35 (34.3)	32 (31.4)	21 (20.6)	21 (20.6)	17 (16.7)	17 (16.7)	13 (12.7)	11 (10.8)	45 (44.1)	46 (45.1)
	Well/mod.	38	15 (39.5)	14 (36.8)	9 (23.7)	11 (28.9)	6 (15.8)	8 (21.1)	1 (2.6)	5 (13.2)	9 (23.7)	24 (63.2)
Vascular	Negative	61	**16 (26.2)**[c]	18 (29.5)	11 (18.0)	10 (16.4)	6 (9.8)	10 (16.4)	7 (11.5)	9 (14.8)	32 (52.5)	20 (32.8)
embolus	Positive	79	**34 (43.0)**	28 (35.4)	19 (24.1)	22 (27.8)	17 (21.5)	15 (19.0)	7 (8.9)	7 (8.9)	38 (48.1)	34 (43.0)
pTNM	I-II	47	**11 (23.4)**[d]	11 (23.4)	7 (14.9)	9 (19.1)	**3 (6.4)**[e]	5 (10.6)	4 (8.5)	4 (8.8)	30 (63.8)	**13 (27.7)**[f]
stage	III	39	**14 (35.9)**	14 (35.9)	10 (25.6)	6 (15.4)	**7 (17.9)**	8 (20.5)	5 (12.8)	4 (10.3)	20 (51.3)	**15 (38.5)**
	IV	54	**25 (46.3)**	21 (38.9)	13 (24.1)	17 (31.5)	**13 (24.1)**	12 (22.2)	5 (9.3)	8 (14.8)	20 (37.0)	**26 (48.1)**
Invasion	T1-2	25	8 (32.0)	8 (32.0)	7 (28.0)	6 (24.0)	3 (12.0)	5 (20.0)	1 (4.0)	1 (4.0)	13 (52.0)	11 (44.0)
	T3	82	27 (32.9)	26 (31.7)	16 (19.5)	17 (20.7)	11 (13.4)	14 (17.1)	10 (12.2)	8 (9.8)	42 (51.2)	30 (36.6)
	T4	33	15 (45.5)	12 (36.4)	7 (21.2)	9 (27.3)	9 (27.3)	6 (18.2)	3 (9.1)	7 (21.2)	15 (45.5)	13 (39.4)
Lymph	Negative	50	14 (28.0)	13 (26.0)	10 (20.0)	9 (18.0)	7 (14.0)	8 (16.0)	5 (10.0)	5 (10.0)	23 (46.0)	22 (44.0)
metast.	Positive	90	36 (40.0)	33 (36.7)	20 (22.2)	23 (25.6)	16 (17.8)	17 (18.9)	9 (10.0)	11 (12.2)	47 (52.2)	32 (35.6)
Distant	Negative	109	36 (33.0)	37 (33.9)	22 (20.2)	25 (22.9)	17 (15.6)	18 (16.5)	13 (11.9)	**13 (11.9)**	**62 (56.9)**	**36 (33.0)**[g]
metast.	Positive	31	14 (45.2)	9 (29.0)	8 (25.8)	7 (22.6)	6 (19.4)	7 (22.6)	1 (3.2)	**3 (9.4)**	**8 (25.8)**	**18 (58.1)**
(Total)		140	50 (35.7)	46 (32.9)	30 (21.4)	32 (22.9)	23 (16.4)	25 (17.9)	14 (10.0)	16 (11.4)	70 (50.0)	54 (38.6)

[a,] Proximal, Cardia/Funfus/Body/Greater curvature; Distal, Antrum/Lesser curvature;

[b/c/d/e/f/g,] Chi-square test, p=0.016/0.040/0.017/0.018/0.036/0.023

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

DISCUSSION

CDKN2A SCND is frequent in many human cancers. However, clinical applications of SCND of *CDKN2A* and other genes are rare. One of the major reseans is the lack of convenient detection approches for routine clinical use. Interstitial deletion/fusion is the main type of *CDKN2A* SCND and the breaking/ coordinates for *CDKN2A* SCNDs are diverse that block the establishment of a feasible assay for *CDKN2A* SCND detection, although many efforts have been afforded (25). In the present study, we characterized for the first time, a 5.1-kb CDR sequence at base resolution within the *CDKN2A* gene in human cancers, and confirmed this finding using WGS datasets for all of 18 gastric cancers containing *CDKN2A* SCND (12). We also established a highly reproducible, convenient, quantitative PCR assay called P16-Light for *CDKN2A* SCNVs detection with the detection limit of 20%.

WGS is generally used as a gold standard to study structural alterations of genomic DNAs, especially for insterstitial gene copy deletion/fusions. Because WGS is a cost assay and its accuracy is depend on sequencing depth, WGS at sequencing depth 36× should be considered as a hemi-quantitative assay. In our calling *CDKN2A* SCND coordiate processes, we found that 18 *CDKN2A* SCND coordiates were identified in 17 (10.8%) of 157 GCs, which is consistent with the frequency (11.4%) of homozygous deletion of *CDKN2A* in 422 GCs in WES or WGS sequencing datasets (Supplemantary Table S2) (39). The positive rate (25.6%) of *CDKN2A* SCND in 156 GCs with enough amounts of genomic DNA samples in the P16-Light analysis was more than two times of that of WGS. And the results of P16-Light analyses were significantly correlated with that of WGS. These phenomena suggest that P16-Light is a more sensitive, more convenient, and less costly assay relative to WGS.

DHPLC-LOH is also a sensitive assay. A set of microsattelite loci are routinely employed as STS markers to detect LOH of a gene. Hence, a set of PCR reactions and DHPLC separation/detection should be performed. Using three *CDKN2A*-flanking microsattelites as the STS marker set to detect LOH of *CDKN2A*, the LOH positive rate of *CDKN2A* was was similar to that of *CDKN2A* SCND. Both LOH and SCND of the gene were significantly associated with GC TNM stages. Because only *CDKN2A* SCND was significantly associated with GC metastasis and the occurence of LOH and SCND of *CDKN2A* were not significantly associated with each other in these GCs, we could not complelely exclude the posssibility that they might be two complementary assays.

The driver function of the *CDKN2A* gene in cancer development is enigmatic. $P16^{ink4a}$ inactivation contributes less than $P19^{arf}$ (the murine counterpart of human $P14^{ARF}$) inactivation to cancer development

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

in mice while *P16*^{INK4a} inactivation contributes more than *P14*^{ARF} inactivation to cancer development in humans (40,41). The exact mechanisms leading to the difference among species is still unclear. Here, we reported that aproximately 87% of genetic *P16*^{INK4a} inactivation is accompanied by *P14*^{ARF} inactivation in human cancer cell lines or tissues. This may account for the species-related functional difference of the *CDKN2A* gene. This explanation is also supported by the report that knocking out both *p16*^{ink4a} and *p19*^{arf} leads to more cancer development than individual inactivation in mice (42).

Relative to the internal control *GAPDH* gene, the copy number of *CDKN2A* could be calculated in the P16-Light assay. A paired normal control sample (such as SM or whole blood DNA) was needed to be used to distinguish SCNVs from germline changes. The sensitivity (or detection limit) for P16-Light was 20%, which might be further increased when a digital PCR platform could be employed in the multiplex PCR reactions. The results of our prospective cohort studies on both regression of esophageal squamous cell dysplasia and post-surgical metastasis of GCs suggest that the sensitivity of 20% is good enough for routine clinical use (prepared for publication).

In conclusion, we have, for the first time, found that there is a 5.1-kb CDR region within the *CDKN2A* gene, and that most *CDKN2A* deletions lead to *P16*^{INK4a} and *P14*^{ARF} coinactivation in human cancers. Using the CDR as a target sequence, we developed a convenient quantitative multiplex PCR assay, the P16-Light to detect *CDKN2A* SCNVs in clinical practice. In addition, we observed a similar CDR region within other tumor suppressor genes such as *ATM*, *FAT1*, *miR31HG*, *PTEN*, and *RB1* in the SNP array-based TCGA SCNV datasets (Supplemantary Figure S4), suggesting that our strategy to detect *CDKN2A* SCNVs may be suitable for the establishment of detection methods for other genes.

Data availability: Detailed information on the base-resolution interstitial common deletion region in *CDKN2A* gene in 110 cancer samples was listed in Supplementary Table S1.

Finding: This work was supported by the Beijing Natural Science Foundation (grant number 7181002 to D.J.D.); Capital's Funds for Health Improvement and Research (grant number 2018-1-1021 to D.J.D.).

Conflict of Insterest Disclosure: The authors have nothing to disclose.

Acknowledgements: We thanks Dr. Sanford Dawsey at NCI, NIH, Bethesda, Maryland for critical comments for the manuscript. We also thanks Miss Gina Mckeown in New York, USA for English language editing.

Author contributions: D.J.D. designed the study, performed bioinformatics analyses, characterized base

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

resolution *CDKN2A CDR* coordinates, and wrote the manuscript draft. J.Z., Y.T., Z.J.L., and L.K.G. developed the P16-Light assay and analyzed *CDKN2A* SCNVs in samples. R.X., Y.Y.L., J.F.J., provided the *CDKN2A CDR* coordinates from whole genome sequencing datasets. All authors approved the final manuscript.

REFERENCES

1. Beroukhim, R., Mermel, C.H., Porter, D., Wei, G., Raychaudhuri, S., Donovan, J., Barretina, J., Boehm, J.S., Dobson, J., Urashima, M. et al. (2010) The landscape of somatic copy-number alteration across human cancers. Nature, 463, 899-905.
2. Cerami, E., Gao, J., Dogrusoz, U., Gross, B.E., Sumer, S.O., Aksoy, B.A., Jacobsen, A., Byrne, C.J., Heuer, M.L., Larsson, E. et al. (2012) The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov, 2, 401-404.
3. Mermel, C.H., Schumacher, S.E., Hill, B., Meyerson, M.L., Beroukhim, R. and Getz, G. (2011) GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. Genome Biol, 12, R41.
4. Gao, J., Aksoy, B.A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S.O., Sun, Y., Jacobsen, A., Sinha, R., Larsson, E. et al. (2013) Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal, 6, pl1.
5. Network, C.G.A.R., University, A.W.G.A., Agency, B.C., Hospital, B.a.W.s., Institute, B., University, B., University, C.W.R., Institute, D.-F.C., University, D., Centre, G.P.C. et al. (2017) Integrated genomic characterization of oesophageal carcinoma. Nature, 541, 169-175.
6. Cui, Y., Chen, H., Xi, R., Cui, H., Zhao, Y., Xu, E., Yan, T., Lu, X., Huang, F., Kong, P. et al. (2020) Whole-genome sequencing of 508 patients identifies key molecular features associated with poor prognosis in esophageal squamous cell carcinoma. Cell Res, 30, 902-913.
7. Serrano, M., Lee, H., Chin, L., Cordon-Cardo, C., Beach, D. and DePinho, R. (1996) Role of the INK4a locus in tumor suppression and cell mortality. Cell, 85, 27-37.
8. Deng, J., Wang, E.S., Jenkins, R.W., Li, S., Dries, R., Yates, K., Chhabra, S., Huang, W., Liu, H., Aref, A.R. et al. (2018) CDK4/6 Inhibition Augments Antitumor Immunity by Enhancing T-cell Activation. Cancer Discov, 8, 216-233.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

9. Jerby-Arnon, L., Shah, P., Cuoco, M.S., Rodman, C., Su, M.J., Melms, J.C., Leeson, R., Kanodia, A., Mei, S., Lin, J.R. et al. (2018) A Cancer Cell Program Promotes T Cell Exclusion and Resistance to Checkpoint Blockade. Cell, 175, 984-997.e924.

10. Zhang, J., Bu, X., Wang, H., Zhu, Y., Geng, Y., Nihira, N.T., Tan, Y., Ci, Y., Wu, F., Dai, X. et al. (2018) Cyclin D-CDK4 kinase destabilizes PD-L1 via cullin 3-SPOP to control cancer immune surveillance. Nature, 553, 91-95.

11. Yu, J., Yan, J., Guo, Q., Chi, Z., Tang, B., Zheng, B., Yin, T., Cheng, Z., Wu, X., Yu, H. et al. (2019) Genetic Aberrations in the CDK4 Pathway Are Associated with Innate Resistance to PD-1 Blockade in Chinese Patients with Non-Cutaneous Melanoma. Clin Cancer Res, 25, 6511-6523.

12. Xing, R., Zhou, Y., Yu, J., Yu, Y., Nie, Y., Luo, W., Yang, C., Xiong, T., Wu, W.K.K., Li, Z. et al. (2019) Whole-genome sequencing reveals novel tandem-duplication hotspots and a prognostic mutational signature in gastric cancer. Nat Commun, 10, 2037.

13. Sobin, L., Gospodarowicz, M., Wittekind, C. (2009) TNM Classification of Malignant Tumours, 7th edn. International Union Against Cancer (UICC). Wiley Press, New York, USA.

14. Kresty, L.A., Mallery, S.R., Knobloch, T.J., Song, H., Lloyd, M., Casto, B.C. and Weghorst, C.M. (2002) Alterations of p16(INK4a) and p14(ARF) in patients with severe oral epithelial dysplasia. Cancer Res, 62, 5295-5300.

15. Sabah, M., Cummins, R., Leader, M. and Kay, E. (2006) Loss of p16 (INK4A) expression is associated with allelic imbalance/loss of heterozygosity of chromosome 9p21 in microdissected malignant peripheral nerve sheath tumors. Appl Immunohistochem Mol Morphol, 14, 97-102.

16. Qin, S.S., Li, Q., Zhou, J., Liu, Z.J., Su, N., Wilson, J., Lu, Z.M. and Deng, D.J. (2014) Homeostatic Maintenance of Allele-Specific p16 Methylation in Cancer Cells Accompanied by Dynamic Focal Methylation and Hydroxymethylation. Plos One, 9.

17. Kleymenova, E. and Walker, C.L. (2001) Determination of loss of heterozygosity in frozen and paraffin embedded tumors by denaturating high-performance liquid chromatography (DHPLC). J Biochem Biophys Methods, 47, 83-90.

18. Pan, K.F., Liu, W., Lu, Y.Y., Zhang, L., Li, Z.P., Lu, W.L., Thibodeau, S.N. and You, W.C. (2003) High throughput detection of microsatellite instability by denaturing high-performance liquid chromatography. Hum Mutat, 22, 388-394.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

19. Ezzeldin, H., Hoffmayer, C., Soong, R., Johnson, M.R., Lee, A., Heslin, M. and Diasio, R. (2004) Simultaneous detection of variable number tandem repeats, single nucleotide polymorphisms, and allelic imbalance in the thymidylate synthase gene enhancer region using denaturing high-performance liquid chromatography. Anal Biochem, 334, 276-283.

20. Cairns, P., Polascik, T.J., Eby, Y., Tokino, K., Califano, J., Merlo, A., Mao, L., Herath, J., Jenkins, R., Westra, W. et al. (1995) Frequency of homozygous deletion at p16/CDKN2 in primary human tumours. Nat Genet, 11, 210-212.

21. Novara, F., Beri, S., Bernardo, M.E., Bellazzi, R., Malovini, A., Ciccone, R., Cometa, A.M., Locatelli, F., Giorda, R. and Zuffardi, O. (2009) Different molecular mechanisms causing 9p21 deletions in acute lymphoblastic leukemia of childhood. Hum Genet, 126, 511-520.

22. Xie, H., Rachakonda, P.S., Heidenreich, B., Nagore, E., Sucker, A., Hemminki, K., Schadendorf, D. and Kumar, R. (2016) Mapping of deletion breakpoints at the CDKN2A locus in melanoma: detection of MTAP-ANRIL fusion transcripts. Oncotarget, 7, 16490-16504.

23. Norris, A.L., Kamiyama, H., Makohon-Moore, A., Pallavajjala, A., Morsberger, L.A., Lee, K., Batista, D., Iacobuzio-Donahue, C.A., Lin, M.T., Klein, A.P. et al. (2015) Transflip mutations produce deletions in pancreatic cancer. Genes Chromosomes Cancer, 54, 472-481.

24. Guney, S., Bertrand, P., Jardin, F., Ruminy, P., Kerckaert, J.P., Tilly, H. and Bastard, C. (2011) Molecular characterization of 9p21 deletions shows a minimal common deleted region removing CDKN2A exon 1 and CDKN2B exon 2 in diffuse large B-cell lymphomas. Genes Chromosomes Cancer, 50, 715-725.

25. Patel, A., Schwab, R., Liu, Y.T. and Bafna, V. (2014) Amplification and thrifty single-molecule sequencing of recurrent somatic structural variations. Genome Res, 24, 318-328.

26. Florl, A.R. and Schulz, W.A. (2003) Peculiar structure and location of 9p21 homozygous deletion breakpoints in human cancer cells. Genes Chromosomes Cancer, 37, 141-148.

27. Kitagawa, Y., Inoue, K., Sasaki, S., Hayashi, Y., Matsuo, Y., Lieber, M.R., Mizoguchi, H., Yokota, J. and Kohno, T. (2002) Prevalent involvement of illegitimate V(D)J recombination in chromosome 9p21 deletions in lymphoid leukemia. J Biol Chem, 277, 46289-46297.

28. Sasaki, S., Kitagawa, Y., Sekido, Y., Minna, J.D., Kuwano, H., Yokota, J. and Kohno, T. (2003) Molecular processes of chromosome 9p21 deletions in human cancers. Oncogene, 22, 3792-3798.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

29. Cayuela, J.M., Gardie, B. and Sigaux, F. (1997) Disruption of the multiple tumor suppressor gene MTS1/p16(INK4a)/CDKN2 by illegitimate V(D)J recombinase activity in T-cell acute lymphoblastic leukemias. Blood, 90, 3720-3726.

30. Pasmant, E., Laurendeau, I., Héron, D., Vidaud, M., Vidaud, D. and Bièche, I. (2007) Characterization of a germ-line deletion, including the entire INK4/ARF locus, in a melanoma-neural system tumor family: identification of ANRIL, an antisense noncoding RNA whose expression coclusters with ARF. Cancer Res, 67, 3963-3969.

31. Raschke, S., Balz, V., Efferth, T., Schulz, W.A. and Florl, A.R. (2005) Homozygous deletions of CDKN2A caused by alternative mechanisms in various human cancer cell lines. Genes Chromosomes Cancer, 42, 58-67.

32. Hussussian, C.J., Struewing, J.P., Goldstein, A.M., Higgins, P.A., Ally, D.S., Sheahan, M.D., Clark, W.H., Tucker, M.A. and Dracopoli, N.C. (1994) Germline p16 mutations in familial melanoma. Nat Genet, 8, 15-21.

33. Freedberg, D.E., Rigas, S.H., Russak, J., Gai, W., Kaplow, M., Osman, I., Turner, F., Randerson-Moor, J.A., Houghton, A., Busam, K. et al. (2008) Frequent p16-independent inactivation of p14ARF in human melanoma. J Natl Cancer Inst, 100, 784-795.

34. Harinck, F., Kluijt, I., van der Stoep, N., Oldenburg, R.A., Wagner, A., Aalfs, C.M., Sijmons, R.H., Poley, J.W., Kuipers, E.J., Fockens, P. et al. (2012) Indication for CDKN2A-mutation analysis in familial pancreatic cancer families without melanomas. J Med Genet, 49, 362-365.

35. Hamosh, A., Scott, A.F., Amberger, J.S., Bocchini, C.A. and McKusick, V.A. (2005) Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Res, 33, D514-517.

36. Amberger, J., Bocchini, C.A., Scott, A.F. and Hamosh, A. (2009) McKusick's Online Mendelian Inheritance in Man (OMIM). Nucleic Acids Res, 37, D793-796.

37. Stone, S., Jiang, P., Dayananth, P., Tavtigian, S.V., Katcher, H., Parry, D., Peters, G. and Kamb, A. (1995) Complex structure and regulation of the P16 (MTS1) locus. Cancer Res, 55, 2988-2994.

38. Luo, D.Y., Zhang, B.Z., Lv, L.B., Xiang, S.Y., Liu, Y.H., Ji, J.F. and Deng, D.J. (2006) Methylation of CpG islands of p16 associated with progression of primary gastric carcinomas. Laboratory Investigation, 86, 591-598.

39. Liu, Y., Sethi, N.S., Hinoue, T., Schneider, B.G., Cherniack, A.D., Sanchez-Vega, F., Seoane, J.A.,

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

Farshidfar, F., Bowlby, R., Islam, M. et al. (2018) Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell, 33, 721-735.e728.

40. Peters, G. (2008) Tumor suppression for ARFicionados: the relative contributions of p16INK4a and p14ARF in melanoma. J Natl Cancer Inst, 100, 757-759.

41. Li, H., Collado, M., Villasante, A., Strati, K., Ortega, S., Cañamero, M., Blasco, M.A. and Serrano, M. (2009) The Ink4/Arf locus is a barrier for iPS cell reprogramming. Nature, 460, 1136-1139.

42. Sharpless, N.E., Ramsey, M.R., Balasubramanian, P., Castrillon, D.H. and DePinho, R.A. (2004) The differential impact of p16(INK4a) or p19(ARF) deficiency on cell growth and tumorigenesis. Oncogene, 23, 379-385.

medRxiv preprint doi: https://doi.org/10.1101/2021.09.16.21263412; this version posted October 17, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
All rights reserved. No reuse allowed without permission.

SUPPLEMENTARY FIGURE LIST

Supplemantary Figure S1. Prevalence of *CDKN2A* deep-deletion and the levels of gene expression in 10967 samples from cancer patients in Pan-TCGA studies. (**A**) Prevalence of *CDKN2A* deep deletion detected by Affymetrix SNP6.0 microarray. The number of total cancer cases and cases with *CDKN2A* deep deletion are listed for each kind of cancer. (**B**) The levels of $P16^{INK4a}$ mRNA determined by RNA sequencing in cancers with various *CDKN2A* genetic changes. The charts for patients (n=10953) in 32 Pan-TCGA studies were adapted from a graphic view at the cBioPortal Cancer Genomics website (www.cbioportal.org).

Supplemantary Figure S2. Distribution pattern of the Online Mendelian Inheritance in Man (OMIM) allelic variants within the *CDKN2A* common deletion region (CDR, highlighted in blue shadow). 12 allelic variants are located in *CDKN2A* exon-2, 2 allelic variants are located in *CDKN2A* exon-1α, and 1 allelic variant is located in *CDKN2A* exon-1β. This chart was adapted from the UCSC Genome Browser on March 10, 2021.

Supplemantary Figure S3. Chromatograms for the microsatellite *D9S974* in human cancer cell lines AGS, HeLa, and MGC803. (**A**) DHPLC chromatograms detected at 50°C; (**B**) the results of clone-sequencing

Supplemantary Figure S4. Approximate locations of the estimated common deletion fragment within tumor suppressor genes *ATM*, *FAT1*, *RB1*, *PTEN*, and *miR31HG* in TCGA pan-cancers (by Affymetrix SNP6.0 microarray).

SUPPLEMENTARY TABLE LIST

Supplemantary Table S1. Base-resolution coordinates for *CDKN2A* deletions in 110 cancer samples

Supplemantary Table S2. *CDKN2A* homozygous deletion in STADs from patients in TCGA cohort