

Research Paper

KRAS mutations in blood circulating cell-free DNA: a pancreatic cancer case-control study

Florence Le Calvez-Kelm¹, Matthieu Foll¹, Magdalena B. Wozniak¹, Tiffany M. Delhomme¹, Geoffroy Durand¹, Priscilia Chopard¹, Maroulio Pertesi¹, Eleonora Fabianova², Zora Adamcakova², Ivana Holcatova³, Lenka Foretova⁴, Vladimir Janout^{5,6}, Maxime P. Vallee¹, Sabina Rinaldi¹, Paul Brennan¹, James D. McKay¹, Graham B. Byrnes¹, Ghislaine Scelo¹

¹International Agency for Research on Cancer (IARC), Lyon, France

²Regional Authority of Public Health, Banska Bystrica, Slovakia

³Charles University of Prague, First Faculty of Medicine, Institute of Hygiene and Epidemiology, Prague, Czech Republic

⁴Masaryk Memorial Cancer Institute and Medical Faculty of Masaryk University, Brno, Czech Republic

⁵Department of Preventive Medicine, Faculty of Medicine, Palacky University, Olomouc, Czech Republic

⁶Faculty of Medicine, University of Ostrava, Czech Republic

Correspondence to: Florence Le Calvez-Kelm, **email:** lecalvez@iarc.fr

Keywords: cell-free DNA, KRAS mutations, plasma, pancreatic cancer detection

Received: June 29, 2016

Accepted: September 19, 2016

Published: October 01, 2016

ABSTRACT

The utility of *KRAS* mutations in plasma circulating cell-free DNA (cfDNA) samples as non-invasive biomarkers for the detection of pancreatic cancer has never been evaluated in a large case-control series. We applied a *KRAS* amplicon-based deep sequencing strategy combined with analytical pipeline specifically designed for the detection of low-abundance mutations to screen plasma samples of 437 pancreatic cancer cases, 141 chronic pancreatitis subjects, and 394 healthy controls. We detected mutations in 21.1% (N=92) of cases, of whom 82 (89.1%) carried at least one mutation at hotspot codons 12, 13 or 61, with mutant allelic fractions from 0.08% to 79%. Advanced stages were associated with an increased proportion of detection, with *KRAS* cfDNA mutations detected in 10.3%, 17.5% and 33.3% of cases with local, regional and systemic stages, respectively. We also detected *KRAS* cfDNA mutations in 3.7% (N=14) of healthy controls and in 4.3% (N=6) of subjects with chronic pancreatitis, but at significantly lower allelic fractions than in cases. Combining cfDNA *KRAS* mutations and CA19-9 plasma levels on a limited set of case-control samples did not improve the overall performance of the biomarkers as compared to CA19-9 alone. Whether the limited sensitivity and specificity observed in our series of *KRAS* mutations in plasma cfDNA as biomarkers for pancreatic cancer detection are attributable to methodological limitations or to the biology of cfDNA should be further assessed in large case-control series.

INTRODUCTION

The latest estimates show that more than 330,000 cases of pancreatic cancer are diagnosed yearly worldwide, and approximately the same number of deaths are attributed to the disease (GLOBOCAN 2012 website: <http://globocan.iarc.fr/>, accessed on 9 Feb 2015). Disease survival is among the poorest of all cancers with 5-year survival at only 6 % in Europe and ~79 % of patients

dying within a year following diagnosis [1, 2]. Improved survival is observed in patients that undergo surgical resection, but this therapeutic option is limited to cases with localized tumors [3]. Early detection has therefore the potential to reduce the mortality associated with pancreatic cancer. Endoscopic ultrasound has shown good sensibility and specificity to detect precancerous and cancerous lesions but this invasive technique has limited use for early detection in asymptomatic individuals [4]. Blood level of

the antigen CA 19-9 is the only validated tumor marker for pancreatic cancer with overall sensitivity of 79% (70-90%) and specificity of 82% (68%-91%) [5, 6]. However, non-specific expression in other benign or malignant diseases and absence of expression in Lewis (a-b-) blood phenotypes (~10-15% of the population) limit the use of this biomarker as a diagnostic test [7].

Cell-free DNA fragments (cfDNA) are released into the bloodstream and other body fluids as part of natural cell apoptosis, necrosis and active secretion. Gene mutations in cfDNA fragments have been found to be tumor-specific leading to the concept of circulating tumor DNA (ctDNA) and their potential utility as highly specific non-invasive biomarkers has raised in the recent years [8]. Pancreatic ductal adenocarcinoma (PDAC) accounts for more than 90% of all pancreatic cancer cases [9] and activating hotspot mutations in the *KRAS* gene are present in the majority of them, representing the most frequent [10] but also the earliest genetic alteration that drives pancreatic neoplasia [11–13]. Of the 596 PDAC cases sequenced within the International Cancer Genome Consortium (ICGC) project (<https://icgc.org/>, as of 23 Feb 2016), 534 (90%) harbored at least one *KRAS* mutation: 83%, 5.5% and 1.5% at codons 12, 61 and 13, respectively. *KRAS* mutations (often restricted to codon 12) have previously been detected in blood (plasma or serum) samples from patients with pancreatic cancer [14–26], showing large variations in the proportion of detected cases (27% to 93%) probably because of inter-laboratory variability, limited sample sizes, and variable sensitivities of the assays. Ultra-deep sequencing technologies allows the identification of low-abundance somatic variants and were shown to be applicable to ctDNA [26–31], but has so far been applied to sample series of limited size and lacking control groups. Here, we investigated whether deep sequencing of *KRAS* codons 12, 13 and 61 in cfDNA from plasma samples from a large series of more than 400 pancreatic cancer cases and 500 controls could represent a comprehensive assay for sensitive and specific detection of pancreatic cancer.

RESULTS

Subject characteristics, sequencing performance and inclusion criteria for analysis

Samples were included when cfDNA total yield was at least 4ng and when sequencing reads were above 1000 on average for all codons. In total, 96 samples (100%) from a pilot set and 903 samples (93.4%; 397 pancreatic cancer cases (94.2%); 132 chronic pancreatitis (91.0%) and 374 controls (93.3%)) from a validation set met the inclusion criteria. Table 1 provides the characteristics of cases and controls, as well as the average of cfDNA yields by status. Analysis of variance was used to compare (log-transformed) cfDNA concentrations by subject characteristics listed in

table 1 and showed significant difference by status (with higher yields in pancreatic cancer cases versus controls; t-test $p < 0.0001$), stage (higher yields in missing stages versus reported stages: $p < 0.0001$), and center (higher yields in Prague and Olomouc when compared to other centers: $p < 0.0001$). Other variables had no significant influence on cfDNA yield (Fisher test $p > 0.05$).

The average mean depth of reads after filtering on mapping quality were, for the pilot and validation sets, respectively: 3992 (SD= 1123) and 2888 (SD=1259) at *KRAS* codon 12 c.34, and 2492 (SD=710) and 3765 (SD=1762) at codon 61 c.181.

Determination of the allelic fraction threshold for the detection of the *KRAS* p.G12V variant

The number of reads obtained from sequencing of 2ng of two independent serial dilutions (in duplicates) of *KRAS* c.35G>T; p.G12V mutated DNA was between 991 and 4205 with an average read depth of 2693 (Supplementary Table S1, Supplementary Data). There was a good correlation between expected and observed mutant allelic fractions ($r^2=0.948$; Supplementary Figure S1). Needlestack analysis was performed independently on the 2 sets of data (Figure 1). Phred scale q-values (QVAL) determined by the Negative binomial regression show that the *KRAS* p.G12V mutation could be reliably detected down to a minor allele frequency of 0.2% when read depth was approximately of 2500 reads QVAL>30 for 3 of the replicates at 0.2%) (Supplementary Table S1, Supplementary Data).

Performance of *KRAS* mutations in cfDNA samples in complement to CA19-9 plasma levels as non-invasive pancreatic cancer biomarker

Applying a threshold of QVAL >30 to the sample set of the pilot series, *KRAS* mutations at hotspot codons reported in PDAC were identified in cfDNA plasma samples in 7 of 40 cases (sensitivity 17.5%) with PDAC and in 1 cfDNA of 27 patients with pancreatic benign neoplasms. None were detected in healthy controls, or in patients with chronic pancreatitis (overall specificity of 98.2%; of 100% against healthy controls) (Tables 2 and 3). All *KRAS* mutations were located at codon 12 (See Supplementary Data Supplementary Table S2 for the complete list of samples harboring cfDNA *KRAS* mutations). Investigating the presence of *KRAS* mutations at other screened codons (from *KRAS* codons 4 to 16 and from codons 51 to 69) and reported mutated for any cancer sites in the COSMIC database identified (i) 2 additional PDAC cases with cfDNA *KRAS* mutations (1 case with p.K5R and 1 case with p.K5R and p.G10R; leading to an overall sensitivity of 22.5%) and (ii) 1 additional mutation in a patient with benign neoplasm of the pancreas (p.A11P). All mutations except one had allelic fraction below 3% (Supplementary Table S2, Supplementary Data).

Table 1: Description of the study population

Characteristics	Pilot series (N=96)						Validation series (N=903)							
	Pancreatic cancer cases		Healthy controls		Chronic pancreatitis		Pancreatic benign neoplasms		Pancreatic cancer cases		Healthy controls		Chronic pancreatitis	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%
Total	40		20		9		27		397		374		132	
Sex														
Male	22	55.0	11	55.0	4	44.4	0	0.0	220	55.4	217	58.0	92	69.7
Female	18	45.0	9	45.0	5	55.6	0	0.0	177	44.6	157	42.0	40	30.3
Missing	0	0.0	0	0.0	0	0.0	27	100.0	0	0.0	0	0.0	0	0.0
Age at blood draw (mean, sd)	64.8 (10.6)		66.2 (8.7)		62.8 (8.2)		Missing		62.2 (10.2)		60.6 (11.9)		55.6 (12.9)	
BMI at blood draw (mean, sd)	24.7 (3.7)		27.4 (4.0)		23.2 (3.8)		Missing		25.1 (4.5)		28.2 (4.3)		24.4 (4.2)	
Recruiting country														
Czech Republic	40	100.0	20	100.0	9	100.0	27	100.0	298	75.1	248	66.3	47	35.6
Slovakia	0	0.0	0	0.0	0	0.0	0	0.0	99	24.9	126	33.7	85	64.4
Tobacco smoking														
Never	20	50.0	9	45.0	3	33.3	0	0.0	167	42.1	175	46.8	45	34.1
Ex-smoker	10	25.0	6	30.0	4	44.4	0	0.0	123	31.0	113	30.2	24	18.2
Current smoker	10	25.0	5	25.0	2	22.2	0	0.0	107	27.0	86	23.0	63	47.7
Missing	0	0.0	0	0.0	0	0.0	27	100.0	0	0.0	0	0.0	0	0.0
Alcohol drinking														
Never	25	62.5	12	60.0	4	44.4	0	0.0	212	53.4	176	47.1	49	37.1
Ex-drinker	6	15.0	3	15.0	4	44.4	0	0.0	95	23.9	36	9.6	48	36.4
Current drinker	9	22.5	5	25.0	1	11.1	0	0.0	87	21.9	162	43.3	35	26.5
Missing	0	0.0	0	0.0	0	0.0	27	100.0	3	0.8	0	0.0	0	0.0
Tumor stage at diagnosis														
Local	6	15.0	-	-	-	-	-	-	33	8.3	-	-	-	-
Regional	17	42.5	-	-	-	-	-	-	126	31.7	-	-	-	-
Systemic	16	40.0	-	-	-	-	-	-	119	30.0	-	-	-	-
Unknown	1	2.5	-	-	-	-	-	-	119	30.0	-	-	-	-
Tumor histological type														
Ductal adenocarcinoma	40	100.0	-	-	-	-	-	-	243	61.2	-	-	-	-
Other ductal carcinoma	0	0.0	-	-	-	-	-	-	19	4.8	-	-	-	-
Endocrine	0	0.0	-	-	-	-	-	-	14	3.5	-	-	-	-
Missing/Unknown	0	0.0	-	-	-	-	-	-	121	30.5	-	-	-	-
Log10 cfDNA concentration, ng/mL plasma (mean, sd)	1.7 (0.5)		1.7 (0.5)		1.8 (0.3)		1.9 (0.7)		2.0 (0.7)		1.7 (0.6)		1.8 (0.7)	

The sensitivity and the overall specificity of plasma CA19-9 levels for detecting PDAC was 90.0% and 64.8% respectively (Table 3). Combining these so that the test was declared positive if a *KRAS* mutation was found at any COSMIC reported position or if the CA19-9 plasma level was positive enabled the detection of 2 additional PDAC cases (38/40) that were negative for CA19-9 plasma level but positive for cfDNA *KRAS* mutation, increasing the sensitivity to 95% (Tables 2 and 3). Comparisons of AUCs of the combined assays versus CA19-9 levels alone showed small increases, approximately 0.02 for each of the three comparisons (cancer cases vs. healthy controls; cancer cases vs. all other conditions; cancer cases vs. benign pancreatic conditions) and were non significant ($p>0.17$ for all comparisons).

Validation of the proportions of detectable cfDNA *KRAS* mutations in pancreatic cancer cases and controls

We extended the cfDNA *KRAS* mutation screening to the validation case-control series (N=903) (Table 1). Of the 397 patients with pancreatic cancer, 75 (18.9%)

carried at least one cfDNA *KRAS* mutation at PDAC hotspot codons, a sensitivity close to that reported for the pilot series (17.5%). We also detected at least one *KRAS* mutations at PDAC hotspot codons in the plasma of 4/132 (3.0%) patients with chronic pancreatitis and of 9/374 (2.4%) healthy controls whereas none were detected in those subjects of the pilot series. Enlarging the search for *KRAS* mutations to other screened codons increased the sensitivity to 20.9% (83 patients with pancreatic cancer carrying at least one mutations in their cfDNA), but decreased the specificity with the detection of cfDNA *KRAS* mutations in 6/132 (4.5%) patients with chronic pancreatitis and in 14/374 (3.7%) healthy controls (Table 4).

Of note, we identified 3 subjects (2 cases and 1 control) with the silent base substitution c.24A>G p.V8V (at 46.38%, 11.46% and 46.98% allelic fractions respectively) which we considered as a rare SNP (rs147406419) as it was reported with an allelic frequency between 0.02% (Exome Variant Server ESP6500sv2) and 0.04% (Exome Aggregation Consortium ExAC) and classified as probably non-pathogenic impact by CLINSIG. This variant was ignored for the rest of the

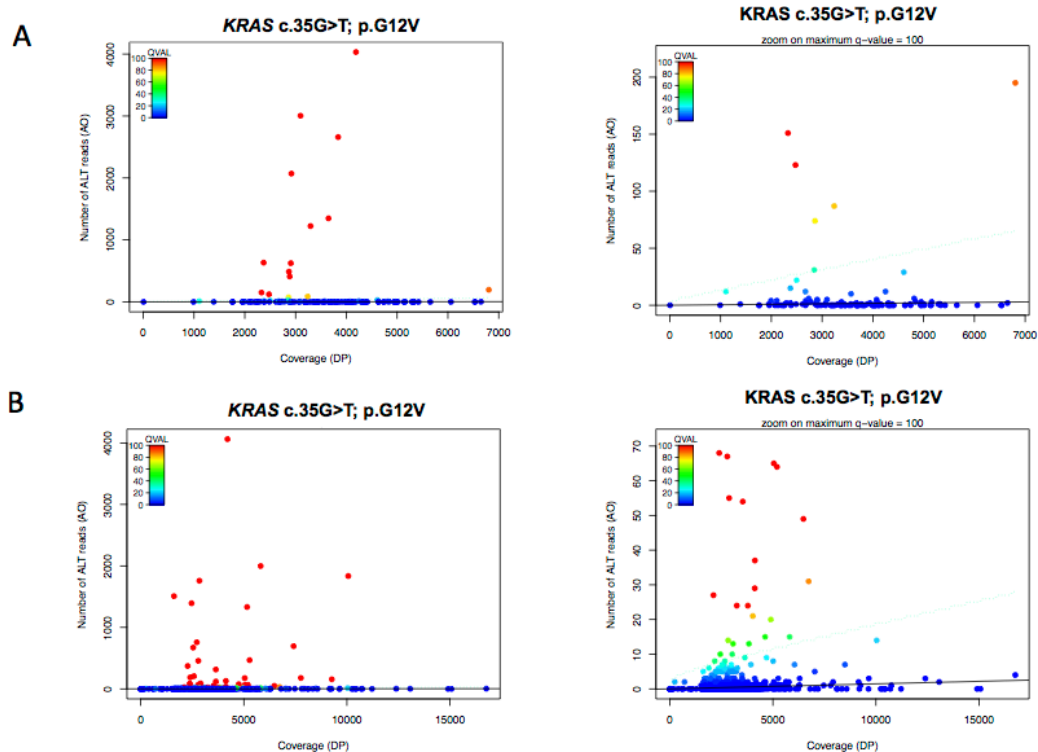


Figure 1: Mutation detection of *KRAS* c.35G>T; p.G12V in serial dilution and cfDNA samples using the Needlestack approach. Negative-binomial regression plot at *KRAS* c.35G>T; p.G12V displaying the total number of reads (coverage, DP) and the number of reads matching the candidate variant (AO). Black solid line: Estimated error rate (e) at the c.35 position for this G>T base change. Blue dashed line: Detection limit at q -values $<10^{-3}$; >30 in Phred scale (QVAL). Dots above the blue dashed line: Outliers of the regression ($QVAL \geq 30$), declared as mutant *KRAS* samples (c.35G>T; p.G12V). Dots below the blue dashed line: Inliers ($QVAL < 30$) declared unmutated at this position for specified base change. **A.** Serial dilution of SW480 cell-lines in duplicates (N=28) and cfDNA from the pilot series (N=96) sequenced on a Ion Torrent PGM 316 Chip ($e = 4.2 \times 10^{-4}$); **B.** Serial dilution of SW480 cell-lines in duplicates (N=28) and cfDNA from the validation series (N=903) sequenced on Ion Torrent PGM 318 chips ($e = 1.4 \times 10^{-4}$).

Table 2: KRAS mutations and CA19-9 plasma levels in the pilot series (N=94)

	cfDNA KRAS mutation at hotspot codons (12, 13, 61) reported in PDAC		cfDNA KRAS mutation at any screened codons reported in any cancer sites	
	N	%	N	%
Plasma CA19-9 positive level (≥ 37Ku/l)				
PDAC case, N=36	5	13.9	7	19.4
Healthy controls, N=3	0	0.0	0	0.0
Benign pancreatic neoplasm, N=11	1	9.1	1	9.1
Chronic pancreatitis, N=5	0	0.0	0	0.0
Plasma CA19-9 negative level (< 37Ku/l)				
PDAC case, N=4	2	50.0	2	50.0
Healthy controls, N=17	0	0.0	0	0.0
Benign pancreatic neoplasm, N=14	0	0.0	1	7.1
Chronic pancreatitis, N=4	0	0.0	0	0.0
Total				
PDAC case, N=40	7	17.5	9	22.5
Healthy controls, N=20	0	0.0	0	0.0
Benign pancreatic neoplasm, N=25*	1	4.0	2	8.0
Chronic pancreatitis, N=9	0	0.0	0	0.0

*Two benign neoplasms were excluded from this analysis because CA19-9 plasma level measurements could not be performed.

Table 3: Performance of NGS-based assay for the detection of cfDNA KRAS mutations, CA19-9 plasma level and combined assays (40 PDAC, 20 healthy controls, 9 chronic pancreatitis subjects, and 25 benign neoplasm subjects)

	Sensitivity	Overall Specificity*	Specificity against healthy controls
cfDNA KRAS mutation			
at PDAC hotspot codons (12, 13, 61)	17.5%	98.2%	100.0%
at any screened codons reported in any cancer sites	22.5%	96.4%	100.0%
CA19-9 plasma level (≥ 37Ku/l)	90.0%	64.8%	85.0%
Combined cfDNA KRAS mutation and CA19-9 plasma level			
at PDAC hotspot codons (12, 13, 61)	95.0%	64.8%	85.0%
at any screened codons reported in any cancer sites	95.0%	63.0% ^a	85.0%

*against non-PDAC and controls

^aDecreased specificity due to the detection of c.31G>C; p.A11P KRAS mutation in a patient with benign neoplasm negative for the plasma CA19-9 assay

analysis. Further restricting the analysis to missense KRAS mutations decreased false positive rates to 3.8% (5/132) and 3.2% (12/374) respectively (Table 4). The

complete list of KRAS mutations identified in cfDNA of the validation series and corresponding allelic fractions is available in supplementary data (Supplementary Data,

Table 4: Proportion of subjects with *KRAS* mutations in their plasma cfDNA

	Pancreatic cancer cases						Chronic pancreatitis						Healthy controls						
	All		Pilot		Validation		All		Pilot		Validation		All		Pilot		Validation		
	N=437		N=40		N=397		N=141		N=9		N=132		N=394		N=20		N=374		
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	
Subjects with <i>KRAS</i> mutations in cell-free DNA	92	21.1	9	22.5	83	20.9	6	4.3	0	0.0	6	4.5	14	3.6	0	0.0	14	3.7	
Numbers of mutation																			
Single	89	20.4	8	20.0	81	20.4	4	2.8	0	0.0	4	3.0	11	2.8	0	0.0	11	2.9	
Multiple	3	0.7	1	2.5	2	0.5	2	1.4	0	0.0	2	1.5	3	0.8	0	0.0	3	0.8	
Location																			
Mutation(s) at PDAC hotpot codon(s) 12, 13 or 61	81	18.5	7	17.5	74	18.6	4	2.8	0	0.0	4	3.0	8	2.0	0	0.0	8	2.1	
Mutation(s) at other codon(s)*	10	2.3	2	5.0	8	2.0	2	1.4	0	0.0	2	1.5	5	1.3	0	0.0	5	1.3	
Mutations at hotpot codons 12, 13 or 61 and others*	1	0.2	0	0.0	1	0.3	0	0.0	0	0.0	0	0.0	1	0.3	0	0.0	1	0.3	
Type																			
Missense	92	21.1	9	22.5	83	20.9	5	3.5	0	0.0	5	3.8	12	3.0	0	0.0	12	3.2	
Silent	0	0.0	0	0.0	0	0.0	1	0.7	0	0.0	1	0.8	2	0.5	0	0.0	2	0.5	

* Codons reported mutated in COSMIC (all cancer sites)

We identified 3 silent base substitutions c.24A>G p.V8V (2 in cases and one in controls), which we considered as a rare SNP (rs147406419) as it was reported with an allelic frequency between 0.0002 (Exome Variant Server ESP6500sv2) and 0.0004 (Exome Aggregation Consortium ExAC) and classified as probably non-pathogenic impact by CLINSIG (Table S3, Supplementary Data). The 3 base substitutions are consequently not included in this table.

Supplementary Table S3). The lowest allelic fraction detected in the cfDNA samples was 0.08% in a plasma case (sample CA93) at *KRAS* p.G13R (Supplementary data, Supplementary Figure S2).

As for somatic *KRAS* mutations reported in PDAC (COSMIC and ICGC data) and chronic pancreatitis (COSMIC data), the majority of cfDNA *KRAS* mutations identified in the combined pilot and validation series were located at codon 12 (76.3 % in pancreatic cancer cases; 77.8% in chronic pancreatitis and 47.4% in healthy controls; Figure 2A). Similar proportions of *KRAS* mutations at codons 61 and 13 were observed in cfDNA of pancreatic cancer cases (7.2% and 3.1% respectively) as compared to PDAC ICGC tumors (6.1% and 1.7% respectively). However, while less than 1% of *KRAS* mutations reported in ICGC/ COSMIC data are located at other codons, 13% (13/97),

22% (2/9), and 31% (6/19) of such mutations were detected in the plasma samples of cancer cases, chronic pancreatitis, and controls, respectively (Figure 2A). The frequencies of the most predominant mutation types reported for PDAC in ICGC, i.e p.G12D, p.G12V, p.G12R, p.G12C followed by p.Q61H, p.Q61R and p.Q61L paralleled the frequencies of the cfDNA *KRAS* mutations in cases (Figure 2B) reflecting the probable tumor origin of the cfDNA *KRAS* mutations. In addition, one cancer case and one control harbored p.Q61P and p.Q61E in their cfDNA, respectively, two non-PDAC COSMIC missense substitutions previously reported in various cancer tissues (Figure 2 and Supplementary Data Supplementary Table S4).

We did not observe striking differences by histological groups. Amongst the 283 PDAC cases, 59 (20.8%) were detected with a cfDNA *KRAS* mutation,

all but three (p.M67L, p.M72V, and p.Q61P) reported as predominant PDAC mutations. Four “other ductal carcinoma” cases out of 19 (21.1%) were also detected with a single cfDNA mutation, all four reported as hotspot PDAC mutations. Amongst the 16 endocrine cases, 3 mutations were detected in 3 cases (18.7%), two of them not reported as hotspot PDAC mutations (p.G60D and p.A59G). The two cases with multiple mutations were found in the pancreatic cancer cases of unknown histological type, where 29 mutations were detected in 27/121 (22.3%) cases. Of these 29 mutations, five (p.A59E, p.E62D, p.Q61R, p.Q70P, and p.Y64D) were not reported as predominant PDAC mutations.

Advanced stages were significantly associated with an increased proportion of detection (*KRAS* cfDNA mutations were detected in 10.3% of cases diagnosed with local stage, 17.5% with regional stage, and 33.3% with systemic stage; chi-squared $p=0.0009$) (Table 5). Among detected cases, there was a non-significant trend of increased allelic fractions with stage

(log10 of fractions were 0.1270, 0.1349, and 0.3047) on average, for local, regional and systemic disease, respectively; linear regression t-test $p=0.3278$). Allelic fractions correlated significantly with status (Table 6), pancreatic cancer cases carrying cfDNA *KRAS* mutations at higher allelic fractions than patients with chronic pancreatitis (t-test on log10(allelic fractions) $p=0.0259$) and healthy controls ($p=0.0008$). Healthy controls and chronic pancreatitis subjects had similar allelic fractions ($p=0.8218$). Of note, 3 PDAC cases were found to carry *KRAS* mutations in their plasma samples at allelic fractions higher than 50% reflecting gain of mutant *KRAS* copies. Other factors associated with allelic fractions were: histological type (with “other ductal carcinoma” cases having higher allelic fractions than PDAC ($p=0.0016$), endocrine ($p=0.0078$), and unknown/missing types ($p=0.0004$); sex (males having higher allelic fractions than females in healthy controls, $p=0.0069$); and age (borderline trend showing higher allelic fractions in older controls, $p=0.0548$).

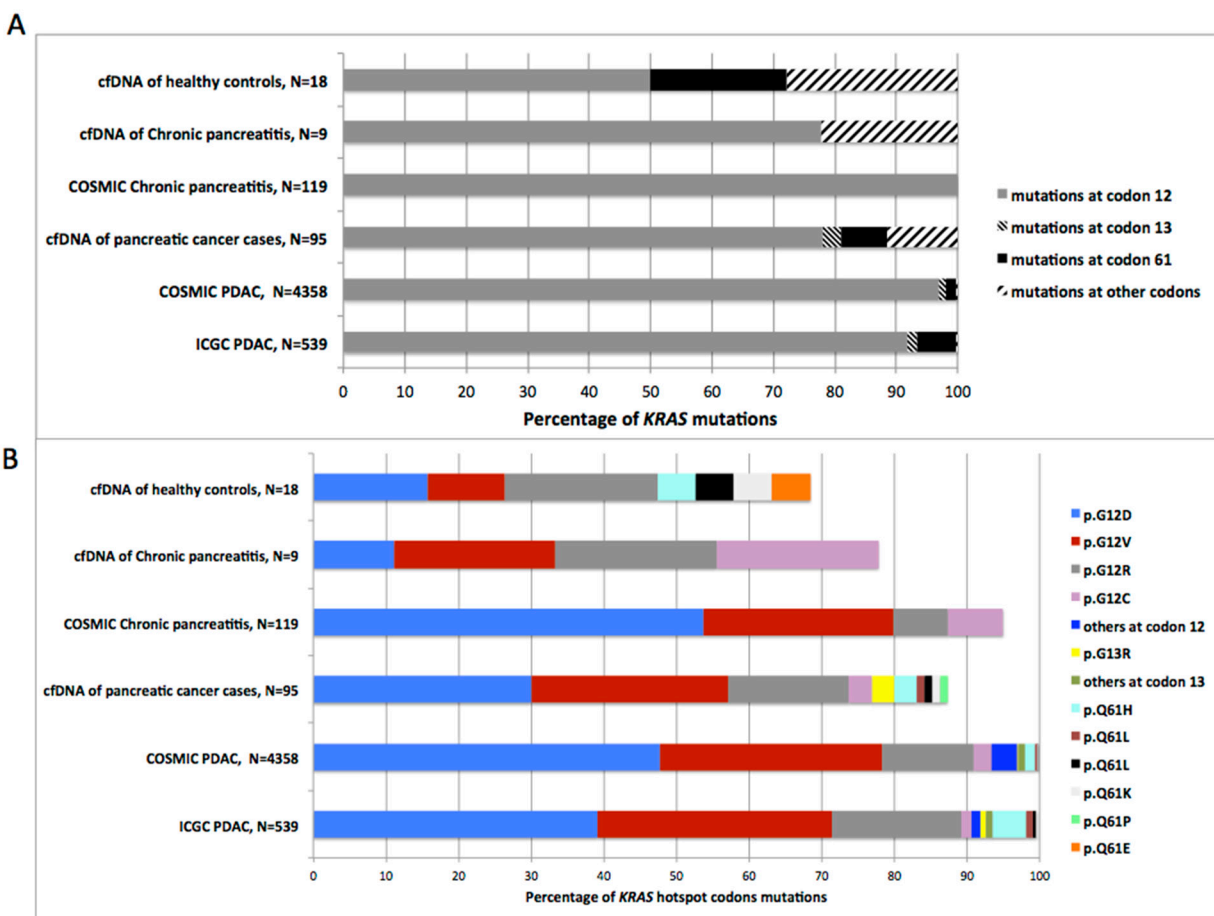


Figure 2: Distribution of *KRAS* mutations detected in plasma samples from pancreatic cases, chronic pancreatitis and healthy controls compared to somatic *KRAS* mutations reported in ICGC and COSMIC database. A. Comparison of cfDNA *KRAS* mutation location; **B.** Comparison of *KRAS* mutation spectrum at hotspot codons (12, 13 and 61). N= Number of *KRAS* mutations.

Table 5: Proportion of pancreatic cancer cases with *KRAS* mutations in their plasma cfDNA, by stage

Stage	Pilot series			Validation series			All		
	Total	cfDNA <i>KRAS</i> mutation		Total	cfDNA <i>KRAS</i> mutation		Total	cfDNA <i>KRAS</i> mutation	
	N	N	%	N	N	%	N	N	%
Local	6	1	16.7	33	3	9.1	39	4	10.3
Regional	17	1	5.9	126	24	19.0	143	25	17.5
Systemic	16	7	43.8	119	38	31.9	135	45	33.3
Unknown	1	0	0.0	119	18	15.1	120	18	15.0
All	40	9	22.5	397	83	20.9	437	92	21.1

KRAS Mutations identified at hotspot codons 12, 13 and 61 and at other codons reported mutated in COSMIC; the silent base substitution c.24A>G p.V8V was excluded from analysis.

Table 6: Proportion of subjects with cfDNA *KRAS* mutations at various allelic fractions

^a AF (%)	Pancreatic cancer cases						Chronic pancreatitis						Healthy controls					
	All, N=93		Pilot, N=9		Validation, N=84		All, N=6		Pilot, N=0		Validation, N=6		All, N=14		Pilot, N=0		Validation, N=14	
	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
<0.2	4	4.3	0	0.0	4	4.8	1	16.7	0	0.0	1	16.7	4	28.6	0	0.0	4	28.6
[0.2-1]	35	37.6	3	33.3	32	38.1	4	66.7	0	0.0	4	66.7	7	50.0	0	0.0	7	50.0
[1.01-10]	40	43.0	5	55.6	35	41.7	0	0.0	0	0.0	0	0.0	2	14.3	0	0.0	2	14.3
[10.01-50]	11	11.8	1	11.1	10	11.9	1	16.7	0	0.0	1	16.7	1	7.1	0	0.0	1	7.1
[50.01-79]	3	3.2	0	0.0	3	3.6	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0	0	0.0

^aAF : Allelic Fraction

KRAS p.V8V excluded

For samples with multiple variants, the mutation with the highest allelic frequency was taken into account

DISCUSSION

To our knowledge, our study is the largest screening of *KRAS* mutations in plasma samples of pancreatic cancer cases, other pathological pancreatic conditions and healthy controls allowing for the comprehensive assessment of the sensitivity and specificity of *KRAS* mutations as non-invasive biomarkers for the detection of pancreatic cancer. Using only 2ng/amplicon (4ng total) of cfDNA and amplicon sizes below the size of the most prominent peak (166 bp) of the recently reported narrow range distribution of cfDNA fragments size [32], our NGS-based *KRAS* mutation screening assay combined with our developed Needlestack variant caller algorithm proved to be a sensitive approach to detect low-allelic fraction *KRAS* mutations down to 0.08%; a detection limit comparable to other amplicon-based NGS sequencing methods [27, 30, 31, 33].

We demonstrated that cfDNA *KRAS* mutations were detectable at the time of diagnosis in the plasma of 20%

of pancreatic cancer cases at PDAC hotspot codons (12, 13 and 61); a sensitivity which is more consistent with some studies (between 27 to 36%) [14, 16, 20, 25] than others (between 47 to 81%) [15, 17–19]. As previously reported, the majority of these alterations were located at the hotspot codon 12, the spectrum was concordant with the distribution of *KRAS* tumor mutation types from ICGC data [34–36], suggesting that *KRAS* mutations in the circulating DNA mainly originate from tumor cells. Interestingly, although it has been shown that 90% of patients with PDAC carry primary *KRAS* mutations at codons 12, 13 or 61, we identified 9 cfDNA variants outside of the predominantly mutated codons, not reported in the ICGC PDAC database but reported in the COSMIC database for other types of cancer, allowing for an increased sensitivity of 22.5%. Those non-hotspot cfDNA *KRAS* mutations identified in pancreatic cancer cases may reflect the heterogeneity of the tumors or the alterations of genetically different metastatic lesions. In agreement with previous reports, we also demonstrated

that the proportion of cases with detectable cfDNA *KRAS* mutations tended to increase with more advanced stages and that *KRAS* allelic fractions were higher in cases than controls or in patients with chronic pancreatitis [23, 26]. Using a ddPCR assay focusing on the four most common PDAC mutations (G12D, G12V, G12R, G13D) Takai and colleagues identified cfDNA *KRAS* mutations in PDAC patients with distant organ metastasis in higher proportion than us (58.9% and 33.3% respectively). However, both studies report similar proportion of detected cases in non-metastatic and localized disease; 8.3% of patients with resectable PDAC (stages I and II) in Takai study and 10.3% of patients with localized pancreatic cancer in our study [23]. While a recent study using ddPCR demonstrated a higher sensitivity (43%; 22 patients) for the detection of *KRAS* mutation in plasma samples of patients with localized PDAC, 10 patients harbored a mutation at an allelic fraction $\leq 0.08\%$ [22]. As 0.08% represents the lowest allele fraction that we could detect with our NGS-based approach and Needlestack algorithm, it is likely that some true low-allelic fraction mutants were too close to the sequencing noise signals to be detected at QVAL > 30. A combined strategy of pre-screening by NGS-amplicon followed by ddPCR of suggestive but inconclusive samples for specific mutations (for example samples with $10 < \text{QVAL} < 30$) could circumvent some limitations by discriminating true positive low-level allelic fractions mutants from inconclusive or false negative NGS samples, providing that the amount of cfDNA obtained is not a limiting factor. Preanalytical parameters regarding blood processing are also known to affect cfDNA concentrations [37]. A limitation of our study is that we did not test whether removing cellular debris with a high speed centrifugation of plasma samples prior cfDNA isolation could improve the sensitivity. However, the low quantities of cfDNA we could extract from the plasma samples on average indicate that contamination by cellular DNA was minimal. It is possible that a proportion of *KRAS* mutant pancreatic cancer do not release *KRAS* mutant cfDNA in the bloodstream, in which case the main limiting factor would be the biology of the tumor rather than the technology. Whether those differences in the release process of cfDNA between patients are due to differences in tumor micro-environment, vascularization, molecular characteristics and/or clonality remains to be discovered [38, 39].

Our study highlights that at our level of detection, a non-negligible proportion of controls are detected. Sausen and colleagues report 99.9% specificity of their assay against matched tumor DNA but they have not evaluated the specificity of their method against plasma of healthy controls. This becomes of capital importance when ultra-low detection limit is required as the proportion of positive calls in non-cancer individuals is likely to increase significantly. The assessment of the biological specificity of mutations in cfDNA as a non-invasive

biomarker is either inexistent or limited in size. This may be partly explained by the fact that somatic mutations are believed to occur at negligible frequencies in normal cell populations [40], and thus expected to derive exclusively from the tumor burden. Yet, using a technique of limited sensitivity, Gormally et al. reported the presence of *KRAS* (1%) and *TP53* (3.2%) mutations in plasma of individuals who had remained clinically cancer-free for more than five years [41]. Very recently two studies revealed low-abundant *TP53* somatic mutations in body fluids of non-cancer individuals [42,43]. In addition, while limited in sample size, two studies described circulating *KRAS* mutations in 5% (2/37) [14] and 13% (4/31) [17] of patients with chronic pancreatitis. In our series, we detected 3.7% (N=14) *KRAS* positive individuals in the healthy controls (N=9 at hotspot codons) and 4.3% (N=6) in subjects with chronic pancreatitis, three of them at PDAC hotspot codon with an allelic fraction >1%. Given the prevalence of *KRAS* mutated cancers (predominantly pancreas, colon and lung) in the population, we cannot exclude that a small proportion of these individuals were non-diagnosed *KRAS* mutated cancer cases. Cell-free DNA fragments released into the blood circulation represent a molecular footprint of the entire genome, potentially including somatic mutations that occur at a mosaic state e.g affecting a limited number of tissues and cells. Syndromes caused by mosaic mutations in the Ras/MAPK signaling pathway (Mosaic RASopathies) have been described as a rather frequent congenital disorder that results in special skin phenotypes, whose epidermal and sebaceous disorders have been recently attributed, among other mutations, to oncogenic mosaic *KRAS* mutations [44]. The relatively high incidence of the most frequent mosaic RASopathy; sebaceous nevi (1 in 1,000 births) suggest that *KRAS* mutations present at a mosaic state in humans may not be a rare phenomenon [45]. Moreover, mosaic RASopathies are predominantly reported as skin disorders because of the accessibility of the lesions but the frequency of those syndromes could be underestimated as mosaic RASopathies of internal organs have been poorly investigated. While there are no accurate estimates of the prevalence and pathogenicity of mosaic *KRAS* mutations in human, it is possible that a proportion of cancer-free individuals with detectable low allelic fractions mutations in circulating DNA could reflect somatic mosaicism.

In conclusion, at a detection limit of 0.08% allelic fraction, our amplicon-based *KRAS* mutations sequencing assay applied to a large case-control series of plasma samples showed a limited sensitivity of 21.1% for the detection of pancreatic cancer and was not as specific as anticipated.

We detected 34% of advanced stages and 10% of early stages, suggesting that the limitation in sensitivity is at least partially attributable to the biology of the pancreatic malignancies. Whether reaching a lower threshold of detection for cfDNA mutations could increase

the discriminatory performance of the test remains to be assessed. We evaluated whether the combination of the detection of circulating *KRAS* mutations and the plasma CA19-9 levels could improve the detection of pancreatic cancer. We confirm a good sensitivity (90%) but a poor specificity for the CA19-9 plasma levels (64.8%). Combining cfDNA *KRAS* mutations and CA19-9 levels improved the sensitivity to 95% but the overall performance of the combined biomarkers did not significantly improve as compared to CA19-9 alone. However, combining cfDNA *KRAS* mutations could potentially contribute to expanded panels of non-invasive biomarkers involving different tumorigenesis processes and/or different mechanisms of release in the bloodstream, such as protein-based [46], exosome-based [47], methylation-based [48] or RNA-based markers [49], for the risk assessment of the disease.

MATERIALS AND METHODS

Study population, sample selection and ethics statement

Samples were selected from a multi-center case-control study conducted in Czech Republic and Slovakia and described in detail elsewhere [50, 51] (Supplementary data).

We conducted this study in two phases, a pilot series where we screened for *KRAS* mutations and measured CA19-9 plasma levels in plasma samples of 96 subjects and a validation series where we extended our initial *KRAS* mutation screening to plasma samples of 967 subjects. For the pilot series, we selected subjects with available plasma and pancreatic tissue (tumor or juice) samples, hence limiting our series to subjects recruited in Czech Republic. We selected all such cases with a histologically-confirmed PDAC diagnosis (N=40) and the 9 subjects diagnosed with chronic pancreatitis (N=9). In addition, we randomly selected 20 healthy controls among 916 with available plasma samples, frequency matched for the 40 PDAC cases on sex, age, tobacco and alcohol consumption. Finally, we selected 27 subjects recruited into the study as pancreatic cancer in first instance, but who subsequently were re-classified as benign neoplasms of the pancreas. For the validation study, we selected all remaining cases with histologically/cytologically confirmed pancreatic cancer (N=421); chronic pancreatitis subjects (N=145); as well as 401 healthy controls among 896, frequency matched for the cancer and chronic pancreatitis subjects on center, sex and age. For pancreatic cancer cases, stage grouping was defined as local, regional, and systemic cancers, based on TNM staging (AJCC 6th edition) when available, and estimation by the clinician when formal TNM staging was not available or not complete.

The study protocol was approved by the institutional review boards of the International Agency for Research

on Cancer and all collaborating centers/institutions, and written informed consent was obtained for all participating subjects.

Isolation of plasma cell-free DNA (cfDNA) and quantification

Peripheral blood from patients was collected in EDTA Vacutainer tubes (Becton Dickinson). Blood samples were processed within 12 h of collection by centrifugation at 2,000g for 10 min and stored frozen in 2mL cryotubes. Circulating DNA (cfDNA) was isolated from 0.6-2.0mL (pilot series; average: 1.4mL) and from 0.3-1.0mL (validation series; average: 0.9mL) plasma with the QIAamp Circulating Nucleic Acid Kit (Qiagen), following manufacturer's instructions [52]. The concentration of purified cfDNA was determined using the Quant-iT™ PicoGreenR dsDNA Assay (Molecular Probes, Invitrogen) PicoGreen® a dilution series of a standard lambda DNA and a Fluoroskan Ascent FL instrument (Thermo Fisher Scientific).

KRAS amplification, library construction and deep sequencing with Ion Torrent PGM

As the size of the cfDNA fragments in cancer patients was recently reported to follow a narrowed-range, unimodal distribution reaching a peak at 166bp [32], primers were designed to amplify exons 2 and 3 so that the amplicon size is < 130bp (79bp and 129bp respectively), covering from codons 4 to 16 (hg19: ch12: 25,398,271 - ch12: 25,398,309) and from codons 51 to 69 (hg19: ch12: 25,380,228 - ch12: 25,380,307), totalling 119 bp excluding primer regions. Forward and reverse primer sequences were 5'-GCCTGCTGAAAATGACTGAA-3' and 5'-AGCTGTATCGTCAAGGCACT-3' for the amplification of partial *KRAS* exon 2 and 5'-GCAAGT AGTAATTGATGGAGAAACC-3' and 5'-TTTATGGCA AATACACAAAGAAAG-3' for the partial amplification of *KRAS* exon 3. Independent PCR amplifications of the 2 exons were performed using 2ng of cfDNA, 5X AccuStart Buffer, 200 nM forward and reverse primers and 0.04 U/mL of AccuStart HiFi Taq Polymerase (Quanta BioSciences) with the following conditions: 2 min at 94°C, 50 cycles of 30s at 94°C, 30s at 58°C and 40s at 72°C and a final elongation of 5 min at 72°C. Approximately 20% of the PCR products were quantified by Qubit™ dsDNA HS Assay Kit and (Invitrogen) and Qubit® 2.0 fluorometer and 20 ng of exon 2 and 3 were pooled together, purified with Serapure magnetic beads at a final concentration of 2.5X and 28% of isopropanol. Library preparation was done using the NEBNext NEB Next® Fast DNA Library Prep Set for Ion Torrent™ kit (New England Biolabs) with some modifications, where each volume of reagents was reduced by a factor 4. Briefly, 12.5µl of the 20µl purified products were end-repaired in 15µl, and added

to 8.6 μ l of ligation reaction mix, 0.7 μ l of the Ion P1 Adapter and 0.7 μ l of each Ion Barcode for the ligation step. The barcoded products were purified using Serapure magnetic beads at final concentration of 1.8X, amplified in 25 μ l and quantified using Qubit quantification system. 40 ng of amplified barcoded products were pooled into a single tube and the cleanup and size selection of pooled libraries (230~250 bp) was performed in a 2% agarose gel and MinElute Gel Extraction Kit (Qiagen). The pool of purified barcoded libraries was quantified using the Qubit quantification system and the assessment of the library quality (molarity and size analysis) was done using the Agilent® High Sensitivity DNA Kit and the Agilent Technologies 2100 Bioanalyzer™ (Agilent Technologies). The pool of purified barcoded libraries was diluted to 280 millions of molecules in 25 μ l and sequenced with the IonTorrent™ PGM sequencer (Thermo Fisher Scientific) at deep coverage using the Ion OneTouch 200 Template Kit v2 DL and Ion PGM Sequencing 200 Kit v2 with the 316 or 318 chips (Thermo Fisher Scientific), following manufacturer's instructions. Library preparation and sequencing conditions were adapted from previous protocols [43].

Detection Threshold

Genomic DNA from the cell-line SW480 harboring a hemizygous *KRAS* p.G12V (c.35G>T) mutation was serially diluted into genomic DNA of a human wild-type lymphoblastoid cell-line in order to assess the accuracy and the detection threshold of the Ion Torrent Sequencing for the measurement of the mutant allelic fraction. Mutant abundances were as follows: 100%, 50%, 20%, 10%, 5%, 2%, 1%, 0.5%, 0.2%, 0.1%, 0.05%, 0.02%, 0.01%. Four independent PCR amplifications were done for each serial diluted point and for six wild-type DNA samples to determine the read error rate for that specific genomic position. PCR amplifications from 2ng, library construction and deep sequencing were done following the same protocol as for the cfDNA.

Measurement of the CA19-9 plasma level

Measurements of CA19-9 were performed on plasma EDTA samples from the pilot study. Analyses were done using an immunoradiometric assay by Beckmann Coulter (Marseille, France). Samples have been randomized through the batches of analyses. We used the clinically accepted cut-off of 37 kU/l for CA19-9 positivity [53].

Bioinformatics and statistical analyses

We used Needlestack, a variant caller algorithm suitable for the detection of low-abundance mutations [43] (<https://github.com/IARCbioinfo/needlestack>). The approach is based on the inclusion of sequencing data of

a sufficient number of samples to robustly estimate the sequencing error rates at each position considered and for each possible base change. Reads were mapped to the human whole genome and BAM files were generated by the Ion Torrent PGM server using default parameters. Reads with a mapping quality below 20 were excluded from subsequent analysis. At each position and for each candidate variant, sequencing errors are modeled using a robust negative binomial regression [54] to avoid bias of the over-dispersion parameter due to the potential presence of genetic variants. We use a linear link and a zero intercept, and detected variants as being outliers from this error model. We calculated for each sample a p-value for being a variant (outlier from the regression) that we further transformed into q-values to account for multiple testing. q-values are reported in Phred scale $QVAL = -10 \log_{10}(q\text{-value})$, and we used a threshold of $QVAL > 30$ to call variants. For each variant, we also calculated the relative variant strand bias defined by:

$$RVSb = \frac{\max(AO_p DP_m, AO_m DP_p)}{AO_p DP_m + AO_m DP_p}$$

where DP and AO denote respectively the total number of reads and the number of reads matching the candidate variant, with the subscripts p and m referring to the forward and reverse strands respectively.

ACKNOWLEDGMENTS

We thank C. Voegele for precious help with operational tasks, H. Renard for the case-control study data management, and AS. Navionis for technical assistance.

CONFLICTS OF INTEREST

None.

GRANT SUPPORT

IARC; the recruitment of subjects was supported by the MH CZ - DRO (MMCI, 00209805) and TMD was supported by la Ligue Nationale (Française) Contre le Cancer.

REFERENCES

1. Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R. EURO CARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. Eur J Cancer. 2009; 45: 931-91.
2. Thomson CS, Forman D. Cancer survival in England and the influence of early diagnosis: what can we learn from recent EURO CARE results? Br J Cancer. 2009; 101 Suppl 2: S102-9.

3. Bliss LA, Witkowski ER, Yang CJ, Tseng JF. Outcomes in operative management of pancreatic cancer. *J Surg Oncol.* 2014; 110: 592-98.
4. Klapman J, Malafa MP. Early detection of pancreatic cancer: why, who, and how to screen. *Cancer Control.* 2008; 15: 280-7.
5. Goonetilke KS, Siriwardena AK. Systematic review of carbohydrate antigen (CA 19-9) as a biochemical marker in the diagnosis of pancreatic cancer. *Eur. J. Surg. Oncol.* 2007; 33: 266-270.
6. Huang Z, Liu F. Diagnostic value of serum carbohydrate antigen 19-9 in pancreatic cancer: a meta-analysis. *Tumour Biol.* 2014; 35: 7459-65.
7. Ballehaninna UK, Chamberlain RS. The clinical utility of serum CA 19-9 in the diagnosis, prognosis and management of pancreatic adenocarcinoma: An evidence based appraisal. *J Gastrointest Oncol.* 2012; 3: 105-19.
8. Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer.* 2011; 11: 426-37.
9. Partensky C. Toward a better understanding of pancreatic ductal adenocarcinoma: glimmers of hope? *Pancreas.* 2013; 42: 729-39.
10. Caldas C, Kern SE. K-ras mutation and pancreatic adenocarcinoma. *Int J Pancreatol.* 1995; 18: 1-6.
11. Kanda M, Matthaei H, Wu J, Hong SM, Yu J, Borges M, Hruban RH, Maitra A, Kinzler K, Vogelstein B, Goggins M. Presence of somatic mutations in most early-stage pancreatic intraepithelial neoplasia. *Gastroenterology.* 2012; 142: 730-33.e9.
12. Maitra A, Fukushima N, Takaori K, Hruban RH. Precursors to invasive pancreatic cancer. *Adv Anat Pathol.* 2005; 12: 81-91.
13. Morris JPt, Wang SC, Hebrok M. KRAS, Hedgehog, Wnt and the twisted developmental biology of pancreatic ductal adenocarcinoma. *Nat Rev Cancer.* 2010; 10: 683-95.
14. Castells A, Puig P, Mora J, Boadas J, Boix L, Urgell E, Sole M, Capella G, Lluís F, Fernández-Cruz L, Navarro S, Farré A. K-ras mutations in DNA extracted from the plasma of patients with pancreatic carcinoma: diagnostic utility and prognostic significance. *J Clin Oncol.* 1999; 17: 578-84.
15. Dianxu F, Shengdao Z, Tianquan H, Yu J, Ruoqing L, Zurong Y, Xuezhi W. A prospective study of detection of pancreatic carcinoma by combined plasma K-ras mutations and serum CA19-9 analysis. *Pancreas.* 2002; 25: 336-41.
16. Jiao L, Zhu J, Hassan MM, Evans DB, Abbruzzese JL, Li D. K-ras mutation and p16 and preproenkephalin promoter hypermethylation in plasma DNA of pancreatic cancer patients: in relation to cigarette smoking. *Pancreas.* 2007; 34: 55-62.
17. Maire F, Micard S, Hammel P, Voitot H, Levy P, Cugnenc PH, Ruszniewski P, Puig PL. Differential diagnosis between chronic pancreatitis and pancreatic cancer: value of the detection of KRAS2 mutations in circulating DNA. *Br J Cancer.* 2002; 87: 551-4.
18. Mulcahy HE, Lyautey J, Lederrey C, qi Chen X, Anker P, Alstead EM, Ballinger A, Farthing MJ, Stroun M. A prospective study of K-ras mutations in the plasma of pancreatic cancer patients. *Clin Cancer Res.* 1998; 4: 271-5.
19. Yamada T, Nakamori S, Ohzato H, Oshima S, Aoki T, Higaki N, Sugimoto K, Akagi K, Fujiwara Y, Nishisho I, Sakon M, Gotoh M, Monden M. Detection of K-ras gene mutations in plasma DNA of patients with pancreatic adenocarcinoma: correlation with clinicopathological features. *Clin Cancer Res.* 1998; 4: 1527-32.
20. Dabritz J, Preston R, Hanfler J, Oettle H. Follow-up study of K-ras mutations in the plasma of patients with pancreatic cancer: correlation with clinical features and carbohydrate antigen 19-9. *Pancreas.* 2009; 38: 534-41.
21. Mora J, Urgell E, Farré A, Comas L, Montserrat E, Gonzalez-Sastre F. Agreement between K-ras sequence variations detected in plasma and tissue DNA in pancreatic and colorectal cancer. *Clin Chem.* 2006; 52: 1448-9.
22. Sausen M, Phallen J, Adleff V, Jones S, Leary RJ, Barrett MT, Anagnostou V, Parpart-Li S, Murphy D, Kay Li Q, Hruban CA, Scharpf R, White JR, et al. Clinical implications of genomic alterations in the tumour and circulation of pancreatic cancer patients. *Nat Commun.* 2015; 6: 7686.
23. Takai E, Totoki Y, Nakamura H, Morizane C, Nara S, Hama N, Suzuki M, Furukawa E, Kato M, Hayashi H, Kohno T, Ueno H, Shimada K, et al. Clinical utility of circulating tumor DNA for molecular assessment in pancreatic cancer. *Sci Rep.* 2015; 5: 18425.
24. Tjensvoll K, Lapin M, Buhl T, Oltedal S, Steen-Ottosen Berry K, Gilje B, Soreide JA, Javle M, Nordgard O, Smaaland R. Clinical relevance of circulating KRAS mutated DNA in plasma from patients with advanced pancreatic cancer. *Mol Oncol.* 2016; 10: 635-43.
25. Uemura T, Hibi K, Kaneko T, Takeda S, Inoue S, Okochi O, Nagasaka T, Nakao A. Detection of K-ras mutations in the plasma DNA of pancreatic cancer patients. *J Gastroenterol.* 2004; 39: 56-60.
26. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, et al. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med.* 2014; 6 :224ra24.
27. Couraud S, Vaca-Paniagua F, Villar S, Oliver J, Schuster T, Blanche H, Girard N, Tredaniel J, Guilleminault L, Gervais R, Prim N, Vincent M, Margery J, et al. Noninvasive diagnosis of actionable mutations by deep sequencing of circulating free DNA in lung cancer from never-smokers: a proof-of-concept study from BioCAST/IFCT-1002. *Clin Cancer Res.* 2014; 20 : 4613-24.
28. Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forshew T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med.* 2013; 368 :1199-209.

29. Leary RJ, Kinde I, Diehl F, Schmidt K, Clouser C, Duncan C, Antipova A, Lee C, McKernan K, De La Vega FM, Kinzler KW, Vogelstein B, Diaz LA, et al. Development of personalized tumor biomarkers using massively parallel sequencing. *Sci Transl Med.* 2010; 2: 20ra14.
30. Narayan A, Carriero NJ, Gettinger SN, Kluytenaar J, Kozak KR, Yock TI, Muscato NE, Ugarelli P, Decker RH, Patel AA. Ultrasensitive measurement of hotspot mutations in tumor DNA in blood using error-suppressed multiplexed deep sequencing. *Cancer Res.* 2012; 72: 3492-8.
31. Newman AM, Bratman SV, To J, Wynne JF, Eclov NC, Modlin LA, Liu CL, Neal JW, Wakelee HA, Merritt RE, Shrager JB, Loo BW, Alizadeh AA, et al. An ultrasensitive method for quantitating circulating tumor DNA with broad patient coverage. *Nat Med.* 2014; 20: 548-54.
32. Jiang P, Chan CW, Chan KC, Cheng SH, Wong J, Wong VW, Wong GL, Chan SL, Mok TS, Chan HL, Lai PB, Chiu RW, Lo YM. Lengthening and shortening of plasma DNA in hepatocellular carcinoma patients. *Proc Natl Acad Sci U S A.* 2015; 112.
33. Forshev T, Murtaza M, Parkinson C, Gale D, Tsui DW, Kaper F, Dawson SJ, Piskorz AM, Jimenez-Linan M, Bentley D, Hadfield J, May AP, Caldas C, et al. Noninvasive identification and monitoring of cancer mutations by targeted deep sequencing of plasma DNA. *Sci Transl Med.* 2012; 4: 136ra68.
34. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller K, Wilson PJ, Patch AM, Wu J, Chang DK, Cowley MJ, Gardiner BB, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. *Nature.* 2012; 491: 399-405.
35. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, Quinn MC, Robertson AJ, Fadlullah MZ, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. *Nature.* 2015; 518: 495-501.
36. Hudson TJ, Anderson W, Artz A, Barker AD, Bell C, Bernabe RR, Bhan MK, Calvo F, Eerola I, Gerhard DS, Guttmacher A, Guyer M, Hemsley FM, et al. International network of cancer genome projects. *Nature.* 2010; 464: 993-8.
37. El Messaoudi S, Rolet F, Mouliere F, Thierry AR. Circulating cell free DNA: Preanalytical considerations. *Clin Chim Acta.* 2013; 424:222-30.
38. Thierry AR, Mouliere F, El Messaoudi S, Mollevi C, Lopez-Crapez E, Rolet F, Gillet B, Gongora C, Dechelotte P, Robert B, Del Rio M, Lamy PJ, Bibeau F, et al. Clinical validation of the detection of KRAS and BRAF mutations from circulating tumor DNA. *Nat Med.* 2014; 20: 430-5.
39. Yong E. Cancer biomarkers: Written in blood. *Nature.* 2014; 511: 524-6.
40. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA, Jr., Kinzler KW. Cancer genome landscapes. *Science.* 2013; 339: 1546-58.
41. Gormally E, Vineis P, Matullo G, Veglia F, Caboux E, Le Roux E, Peluso M, Garte S, Guarrera S, Munnia A, Airoidi L, Autrup H, Malaveille C, et al. TP53 and KRAS2 mutations in plasma DNA of healthy subjects and subsequent cancer occurrence: a prospective study. *Cancer Res.* 2006; 66: 6871-6.
42. Krimmel JD, Schmitt MW, Harrell MI, Agnew KJ, Kennedy SR, Emond MJ, Loeb LA, Swisher EM, Risques RA. Ultra-deep sequencing detects ovarian cancer cells in peritoneal fluid and reveals somatic TP53 mutations in noncancerous tissues. *Proc Natl Acad Sci U S A.* 2016.
43. Fernandez-Cuesta L, Perdomo S, Avogbe PH, Leblay N, Delhomme TM, Gaborieau V, Abedi-Ardekani B, Chanudet E, Olivier M, Zaridze D, Mukeria A, Vilensky M, Holcatova I, et al. Identification of Circulating Tumor DNA for the Early Detection of Small-cell Lung Cancer. *EBioMedicine.* 2016; 117-23.
44. Hafner C, Toll A, Gantner S, Mauerer A, Lurkin I, Acquadro F, Fernandez-Casado A, Zwarthoff EC, Dietmaier W, Baselga E, Parera E, Vicente A, Casanova A, et al. Keratinocytic epidermal nevi are associated with mosaic RAS mutations. *J Med Genet.* 2012; 49: 249-53.
45. Hafner C, Groesser L. Mosaic RASopathies. *Cell Cycle.* 2013; 12: 43-50.
46. He XY, Liu BY, Yao WY, Zhao XJ, Zheng Z, Li JF, Yu BQ, Yuan YZ. Serum DJ-1 as a diagnostic marker and prognostic factor for pancreatic cancer. *J Dig Dis.* 2011; 12: 131-7.
47. Melo SA, Luecke LB, Kahlert C, Fernandez AF, Gammon ST, Kaye J, LeBleu VS, Mittendorf EA, Weitz J, Rahbari N, Reissfelder C, Pilarsky C, Fraga MF, et al. Glypican-1 identifies cancer exosomes and detects early pancreatic cancer. *Nature.* 2015; 523: 177-82.
48. Yi JM, Guzzetta AA, Bailey VJ, Downing SR, Van Neste L, Chiappinelli KB, Keeley BP, Stark A, Herrera A, Wolfgang C, Pappou EP, Iacobuzio-Donahue CA, Goggins MG, et al. Novel methylation biomarker panel for the early detection of pancreatic cancer. *Clin Cancer Res.* 2013; 19: 6544-55.
49. Liu R, Chen X, Du Y, Yao W, Shen L, Wang C, Hu Z, Zhuang R, Ning G, Zhang C, Yuan Y, Li Z, Zen K, et al. Serum microRNA expression profile as a biomarker in the diagnosis and prognosis of pancreatic cancer. *Clin Chem.* 2012; 58: 610-8.
50. Brenner DR, Wozniak MB, Feyt C, Holcatova I, Janout V, Foretova L, Fabianova E, Shonova O, Martinek A, Ryska M, Adamcakova Z, Flaska E, Moskal A, et al. Physical activity and risk of pancreatic cancer in a central European multicenter case-control study. *Cancer Causes Control.* 2014; 25: 669-81.
51. Urayama KY, Holcatova I, Janout V, Foretova L, Fabianova E, Adamcakova Z, Ryska M, Martinek A, Shonova O, Brennan P, Scelo G. Body mass index and body size in early adulthood and risk of pancreatic cancer in a central European multicenter case-control study. *Int J Cancer.* 2011; 129: 2875-84.

52. Devonshire AS, Whale AS, Gutteridge A, Jones G, Cowen S, Foy CA, Huggett JF. Towards standardisation of cell-free DNA measurement in plasma: controls for extraction efficiency, fragment size bias and quantification. *Anal Bioanal Chem.* 2014; 406:6499-512.
53. Steinberg W. The clinical utility of the CA 19-9 tumor-associated antigen. *Am J Gastroenterol.* 1990; 85: 350-5.
54. Aeberhard WH, Cantoni E, Heritier S. Robust inference in the negative binomial regression model with an application to falls data. *Biometrics.* 2014; 70: 920-31.