SPANDIDOS
PUBLICATIONS

Mutational burden of resectable pancreatic cancer, as determined by whole transcriptome and whole exome sequencing, predicts a poor prognosis

ELISA GRASSI[1], SANDRA DURANTE[2], ANNALISA ASTOLFI[2], GIUSEPPE TARANTINO[2], VALENTINA INDIO[2], EVA FREIER[1], SILVIA VECCHIARELLI[1], CLAUDIO RICCI[3], RICCARDO CASADEI[3], FRANCESCA FORMICA[1], DARIA FILIPPINI[1], FRANCESCA COMITO[1], CARLA SERRA[4], DONATELLA SANTINI[5], ANTONIETTA D' ERRICO[5], FRANCESCO MINNI[3], GUIDO BIASCO[1,2] and MARIACRISTINA DI MARCO[1]

[1]Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Sant' Orsola-Malpighi Hospital; [2]Interdepartmental Center of Cancer Research University of Bologna, Sant' Orsola-Malpighi Hospital; Departments of [3]Medical and Surgical Sciences, [4]Internal Medicine and [5]Pathology, University of Bologna, Sant' Orsola-Malpighi Hospital, I-40138 Bologna, Italy

Abstract. Despite the genomic characterization of pancreatic cancer (PC), marked advances in the development of prognosis classification and novel therapeutic strategies have yet to come. The present study aimed to better understand the genomic alterations associated with the invasive phenotype of PC, in order to improve patient selection for treatment options. A total of 30 PC samples were analysed by either whole transcriptome (9 samples) or exome sequencing (21 samples) on an Illumina platform (75X2 or 100X2 bp), and the results were matched with normal DNA to identify somatic events. Single nucleotide variants and insertions and deletions were annotated using public databases, and the pathogenicity of the identified variants was defined according to prior knowledge and mutation-prediction tools. A total of 43 recurrently altered genes were identified, which were involved in numerous pathways, including chromatin remodelling and DNA damage repair. In addition, an analysis limited to a subgroup of early stage patients (50% of samples) demonstrated that poor prognosis was significantly associated with a higher number of known PC mutations (P=0.047). Samples from patients with a better overall survival (>25 months) harboured an average of 24 events, whereas samples from patients with an overall survival of <25 months presented an average of 40 mutations. These findings indicated that a complex genetic profile in the early stage of disease may be associated with increased aggressiveness, thus suggesting an urgent requirement for an innovative approach to classify this disease.

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is the fourth leading cause of cancer-associated mortality and has a 5-year survival rate of 5%. It has been predicted to become the second leading cause of cancer-associated mortality within a decade (1). Surgery remains the most effective treatment for PDAC; however, only 20% of patients are suitable for radical resection at the time of diagnosis, and a large number of cases that undergo curative surgery develop recurrence at a median of 10-20 months following resection (2). In addition, treatment of patients with recurrent or metastatic PDAC is limited to palliative chemotherapy. Therapeutic advances, including the use of FOLFIRINOX and gemcitabine plus nab-paclitaxel regimens, have resulted in a modest improvement in outcome (3,4).

It is well known that cancer is a complex disease caused by the accumulation of somatic mutations acquired by the genome of cells. With the advent of next-generation sequencing technologies, it is possible to detect a large number of recurring mutations and many highly mutated genes with high resolution. In particular, whole transcriptome massively parallel sequencing (RNA-seq) and whole exome sequencing (WES) are powerful tools used to detect mutations underlying carcinogenesis and to detect somatic events, such as nucleotide substitution mutations and gene translocations, with high resolution by sequencing the expressed gene (cell transcriptomes) and the protein coding regions of the genome (cell exome) (5). The genomic characterization of PDAC via the molecular analysis of all somatic alterations has generated much information, highlighting a complex mutational landscape. With the exception of well-known KRAS proto-oncogene, GTPase (KRAS), tumour protein p53 (TP53), cyclin-dependent kinase inhibitor 2A (CDKN2A) and SMAD family member 4 (SMAD4) alterations, which have a

Correspondence to: Dr Eva Freier, Department of Experimental, Diagnostic and Specialty Medicine, University of Bologna, Sant' Orsola-Malpighi Hospital, 9 Massarenti Street, I-40138 Bologna, Italy
E-mail: eva.freier@gmail.com

frequency of 71, 49, 22 and 20%, respectively, as reported in the literature, a large number of genomic rearrangements with mutational frequency <2% were detected (6).

The majority of single gene mutations in pancreatic cancer can be grouped in common cellular pathways. Jones *et al* identified 69 mutated gene sets in most of the 24 samples analysed, of which 31 could be grouped into 12 core signalling pathways. These pathways included KRAS signalling, transforming growth factor (TGF)-β pathway, DNA damage control, apoptosis and regulation of G1/S cell cycle transition, and other pathways, such as Hedgehog signalling, the homophilic cell adhesion pathway, integrin signalling, TGF-β signalling, Wnt/Notch signalling and regulation of the invasion pathway (7).

However, at present, no prognostic or therapeutic molecular markers are available for the better selection of patients for surgery and therapeutic strategies.

In order to better understand PDAC biology, the present study analysed 30 PDAC samples using either whole transcriptome (9 samples) or exome sequencing (21 samples). A total of 43 recurrently altered genes were identified, which are involved in numerous pathways, including chromatin remodelling and DNA damage repair. Subsequently, an analysis of mutational events in known PDAC mutated genes was conducted on early stage samples (50% of specimens), and a high number of mutations was revealed to be significantly associated with a poor prognosis. These findings indicated that a low level of gene mutations may predict a favourable prognosis following surgery in patients with early stage PDAC.

Materials and methods

Sample collection and sequencing analysis. The present study was conducted according to the principles expressed in the Declaration of Helsinki, and written informed consent was obtained from all participants. The study was approved by the Independent Ethics Committee of Sant' Orsola-Malpighi Hospital (Bologna, Italy). A total of 30 PDAC samples were obtained by ultrasound-guided biopsy or surgery for either RNA or DNA extraction between January 2011 and December 2015. The tissues were collected in cryogenic tubes and were stored at -20˚C in RNAlater solution and DNAlater solution (both from Ambion; Thermo Fisher Scientific, Inc., Waltham, MA, USA), respectively. Nucleic acid extraction was performed using the AllPrep RNA/DNA kit for tumour biopsies and the QiaAmp DNA Mini kit (both from Qiagen, Inc., Valencia, CA, USA) for peripheral blood samples of the same patients with PDAC; the results were matched with this normal DNA to identify somatic events.

Out of the 30 samples, 9 were analysed by RNA-seq and 21 were analysed by WES, performed at 75X2 or 100X2 bp on a HiScanSQ Illumina platform (Illumina, Inc., San Diego, CA, USA). Analysis was performed on a local CentOS5 Linux server by applying a customised bioinformatics pipeline. Following the conversion of data from binary base call (BCL) to FASTQ format (Illumina Bcl2Fastq function), short reads were processed to clean up sequencing adapters, and to filter or trim the reads for sequence quality (minimum Phred quality, 10; minimum length of trimmed sequence, 30), both of these steps were performed using AdapterRemoval (8). The

cleaned reads were mapped to the human reference genome hg38 using Burrows-Whealer Aligner in paired-end mode for WES data and with the Tophat/Bowtie pipeline for RNA-seq data (9). In order to remove the optical and polymerase chain reaction duplicates, Samtools was adopted, and GATK was used to optimize the alignment around insertions and deletions (InDels) (10,11). Single nucleotide variants (SNVs) and InDels were detected using GATK (HaplotypeCaller mode), Mutect and SNVmix2, applying manually curated quality filters (12,13).

The whole set of detected variants was refined adopting different filtering thresholds for each sample. These criteria were based on KRAS variant detection, which depends on the abundance of tumour cells within a sample. Evaluation of tumour cells was based on the presence and relative enrichment of the KRAS mutation; only samples with >10% tumour vs. normal allele were included in the study. Minimum depth of coverage and minimum ratio (proportion between the depth of coverage of the mutated allele and the total depth of coverage) were established in the range of 15-20 and 0.05-0.1, respectively. Therefore, the confidence in selecting genuine variants was increased.

The obtained variants were annotated with 1000 Genomes allele frequencies, dbSNP 149 rsIDs, Exac and EVS, using Annovar and Oncotator, in order to discard polymorphisms, and to identify novel or almost rare variants (population frequency <0.01) (14,15). Finally, to determine the pathogenic significance of the emerging variants, previous knowledge (Catalog of Somatic Mutations in Cancer) and bioinformatics mutation-prediction tools (PolyPhen2, Proven and SIFT) were used (16-18). The resulting list of variants was analysed to distinguish between somatic and germline mutations, according to the presence of alternative events compared with in the normal counterpart (peripheral blood samples) using the Samtools mpileup function. The following criteria were used: Total depth >5 and depth of alternative base=0 (depth=1 or 2 were admitted if the total depth was ≥15 or 30, respectively) for inclusion of the variant in the 'somatic' set, otherwise the events were marked as 'germline'; however, if total coverage of the normal counterpart was considered insufficiently informative (total depth <5) the variant was classified as 'undetermined'. The SNV and InDels were automatically annotated with regards to the common PDAC pathways using the Reactome database (https://reactome.org/PathwayBrowser/#TOOL=AT).

With regards to WES data, copy number alterations were identified using two software programs, Control-FREEC and ADTEX (19,20). These tools compute and normalise the copy number and β-allele frequency profiles of segments, thus providing copy number alterations. In the present study, WES data were analysed with the matched normal sample used as a control. The catalogue of human genomic structural variation (Database of Genomic Variants; http://dgv.tcag.ca/dgv/app/home) was used to remove physiological alterations.

Mutational load. The present study restricted the analysis to the early stage subgroup of patients (15 samples) and evaluated the mutational load. The threshold used for overall survival (OS) was derived from the current scientific literature, in order to stratify the patients into long survival (>25 months)

Table I. Patient characteristics.

Variable	Value
Age (years)	
Median	74
Range	33-83
Sex (%)	
Male	43
Female	57
Site of specimen (%)	
Pancreatic tumour	93
Hepatic metastasis	7
Site of tumour (%)	
Head	63
Body/tail	37
Stage (%)[a]	
I+II	60
III-LA	13
IV	27
Surgery (%)	
Yes	60
No	40

[a]Classified according to the 8th edition of the American Joint Committee on Cancer staging system (34).

and short survival (<25 months) groups; three early stage cases with an OS value very close to 25 months were excluded, and only the extreme cases were analysed (2).

A list of 365 genes previously reported in the literature was considered somatically mutated in PDAC according to the Catalogue of Somatic Mutations in Cancer (http://cancer.sanger.ac.uk/cosmic/browse/tissue?wgs=off&sn=pancreas&ss=NS&hn=carcinoma&sh=ductal_carcinoma&in=t&src=tissue&all_data=n). Matching this set with the genes carrying SNVs or InDels in the samples of the present study, the mutational load of each patient was estimated. Using the software suite SPSS version 23 (IBM Corporation, Armonk, NY, USA), one-way analysis of variance was used to highlight the statistical difference in terms of OS between the groups with high mutational load and the low mutational load, using 25 months as the OS cut-off (Fig. 4). Finally the median OS for the two groups of subjects was calculated according to a survival analysis and Kaplan-Meier estimation.

Results

Sequencing analysis data. A total of 30 PDAC samples were analysed by either whole transcriptome or exome sequencing. The patient characteristics are summarised in Table I. After matching with normal DNA to identify somatic events, an average of 71 coding non-synonymous novel disease-related SNVs (ranging from 8 of the case with the lowest number of mutation events to 304 of the one with the highest number) and three InDels (ranging from 0 of the case with the lowest number

of mutation events to 6 of the one with the highest number) were identified. In addition, intra- or inter-chromosomal rearrangements were detected in the set of samples analysed by RNA-seq. However, no recurrent fusion transcripts were detected.
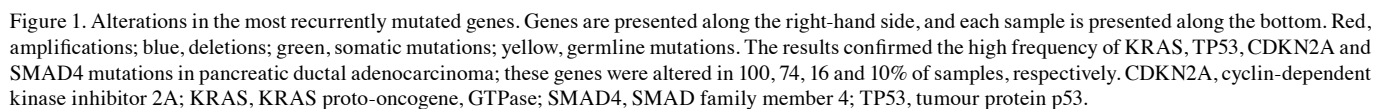
A total of 43 recurrently mutated genes were identified, and the high frequency of KRAS, TP53, CDKN2A and SMAD4 mutations was confirmed in PDAC; these genes were altered in 100, 74, 16 and 10% of samples, respectively. As expected, KRAS exhibited the highest prevalence of somatic mutations; the mutations affected the known hotspot at codon 12 (G12D, G12V, G12R and G12C in 16, 8, 4 and 2 patients, respectively). The prevalence of the KRAS mutation is however biased by the experimental design, since samples were included in the present study only if a mutation in any KRAS gene was detected, estimating the percentage of tumour cells in the sample via quantification of the KRAS mutation in the DNA. Somatic mutations in other genes were much less recurrent, as summarised in Fig. 1.

Pathway analysis demonstrated that the genes with a high frequency of aberrations were clustered into the following specific pathways: KRAS signalling, TGF-β signalling, chromatin remodelling, Wnt signalling, DNA damage repair, cell cycle and RNA processing (Fig. 2).

Mutational load data. In order to better understand the somatic driving events in PDAC carcinogenesis, analysis was restricted to the early stage subgroup of patients (50% of the samples), including only patients with primary, operable and untreated PDAC, which underwent pancreatectomy with curative intent. Since the RNA-seq data may lead to biased results in the mutational load evaluation, the present study normalised the number of mutations for the coding sequences covered at least 10X using Samtools; therefore, it was considered that RNA-seq data covered the transcriptome dimension. In addition, of the 9 cases analysed through RNA-seq, only 5 cases were retained for analysis of the mutational load, and these were almost equally distributed in the two subgroups. Cohort characteristics are presented in Table II. The median OS of the two groups, as estimated using a Kaplan-Meier analysis, were 46 and 13 months for patients classified into the long and short survival groups, respectively (Fig. 3).

The mutational load of each patient was evaluated and was matched to a list of 365 genes previously described as mutated in PDAC. Among the genes carrying SNVs or InDels in the present study samples, an average of 33 events were detected. Based on 25 months as the cut-off value, patients were divided into two subgroups; the subgroup with the better OS (>25 months) harboured an average of 24 events, whereas the subgroup with the worse OS (<25 months) presented an average of 40 mutations (P=0.047) (Fig. 4)

The most recurrent genes altered by point mutations, deletions or amplifications in the subgroup were KRAS (100%), TP53 (60%), mucin 4, cell surface associated (MUC4), olfactory receptor family 10 subfamily Z member 1 (33%), CDKN2A, GNAS complex locus, SMAD4, TGF-β receptor 2 (TGFBR2), filaggrin (27%), SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2, dynein axonemal heavy chain 14, titin, adhesion G protein-coupled receptor B1, Ral GTPase activating

Figure 1. Alterations in the most recurrently mutated genes. Genes are presented along the right-hand side, and each sample is presented along the bottom. Red, amplifications; blue, deletions; green, somatic mutations; yellow, germline mutations. The results confirmed the high frequency of KRAS, TP53, CDKN2A and SMAD4 mutations in pancreatic ductal adenocarcinoma; these genes were altered in 100, 74, 16 and 10% of samples, respectively. CDKN2A, cyclin-dependent kinase inhibitor 2A; KRAS, KRAS proto-oncogene, GTPase; SMAD4, SMAD family member 4; TP53, tumour protein p53.

protein non-catalytic β subunit (20%), ring finger protein 43 (RNF43), polybromo 1, obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF, proline-rich coiled-coil 2C, coiled-coil domain-containing 168, crystalline βγ domain-containing 2, E4F transcription factor 1 and POM121-like protein 12 (POM12L12) (13%), as shown in Fig. 5. However, no specific genes or pathways were correlated with OS.

Discussion

Based on chromosome structural variation and gene expression profiles, recent studies have classified PDAC into various subtypes with prognostic and biological relevance. In particular, Bailey *et al* (22) correlated four PDAC molecular subtypes with specific histopathological
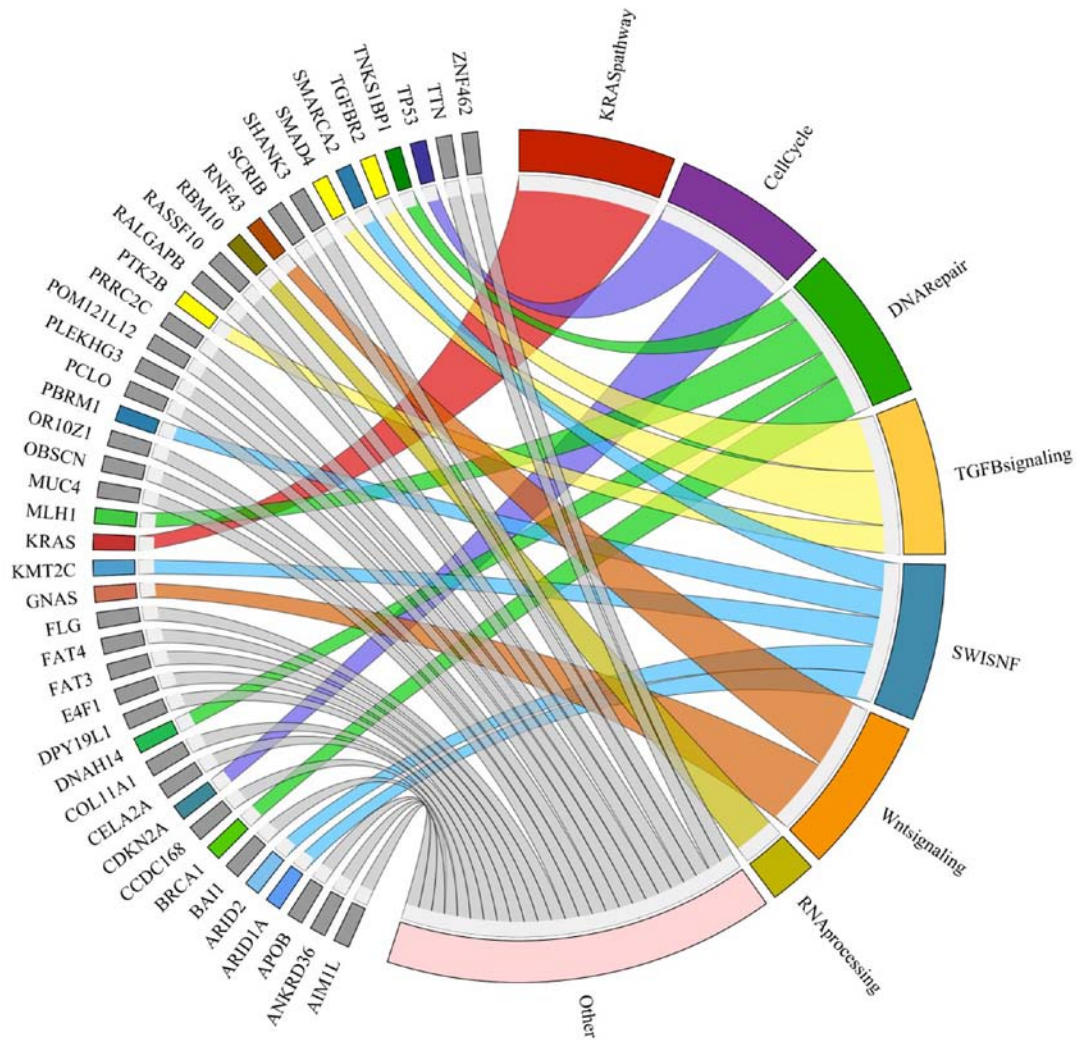
Figure 2. Genes with a high frequency of aberrations, as presented on the left-hand side of the image, were clustered into specific pathways, as presented on the right-hand side. Connections are colour-coded. The most recurrently detected pathways included KRAS signalling, TGF-β signalling, chromatin remodelling, Wnt signalling, DNA damage repair, cell cycle and RNA processing. KRAS, KRAS proto-oncogene, GTPase; TGF-β, transforming growth factor-β.
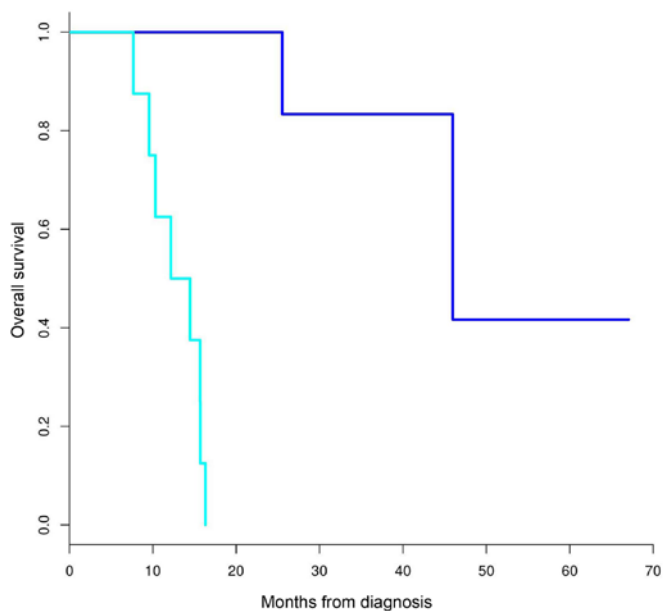

Figure 3. Median overall survival of early stage patients characterised into the long and short survival groups, as determined by Kaplan-Meier analysis. Dark blue, long survival group; light blue, short survival group.
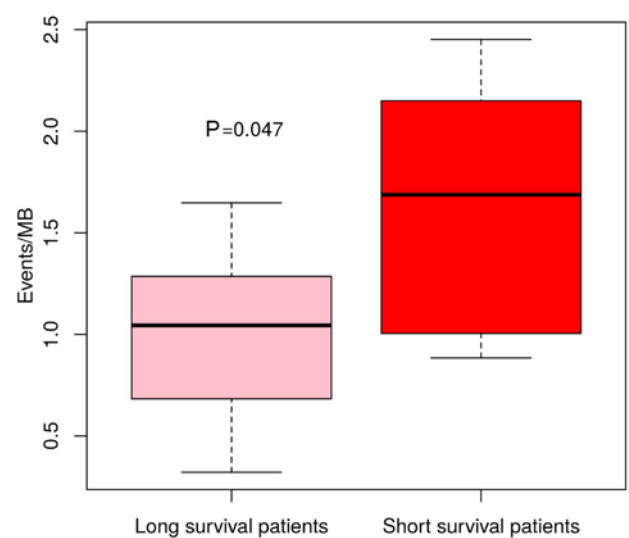

Figure 4. Based on 25 months as the cut-off value, patients were divided into two subgroups, with long (>25 months) and short overall survival (<25 months), respectively. The histogram represents the events/MB in the two subgroups. Samples from the long survival group harboured an average of 24 events, whereas samples from the short overall group possessed an average of 40 mutations (P=0.047). MB, mutational burden.
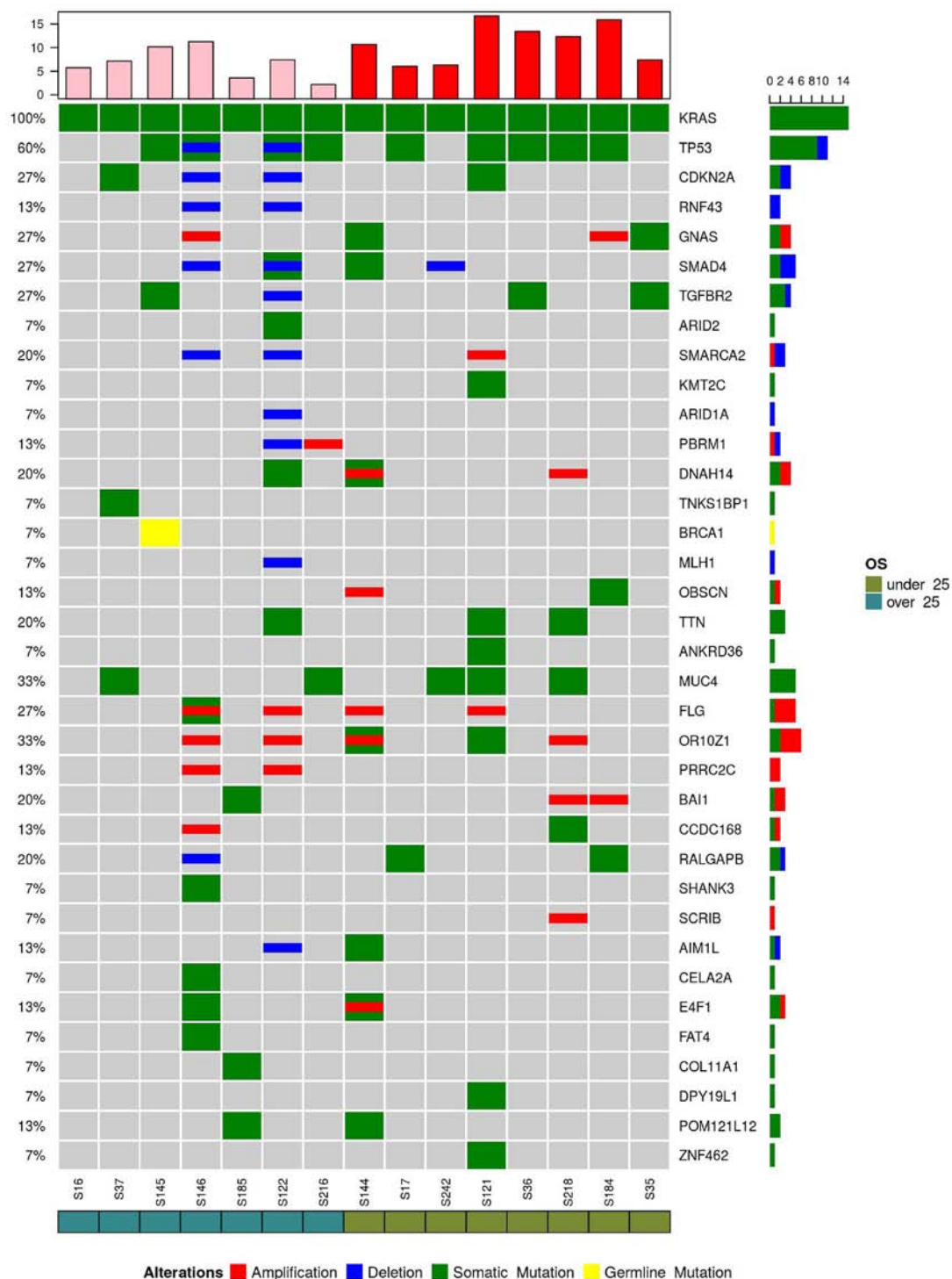
Figure 5. Recurrent genes altered by point mutations (green), deletions (blue) or amplifications (red) in the early stage samples are presented. A high frequency of mutations were detected in the following genes: KRAS (100%), TP53 (60%), MUC4, OR10Z1 (33%), CDKN2A, GNAS, SMAD4, TGFBR2, FLG (27%), SMARCA2, DNAH14, TTN, BAI1, RALGAPB (20%), RNF43, PBRM1, OBSCN, PRRC2C, CCDC168, AIM1L, E4F1 and POM121L12 (13%). AIM1L, crystalline βγ domain-containing 2; BAI1, adhesion G protein-coupled receptor B1; CCDC168, coiled-coil domain-containing 168; CDKN2A, cyclin-dependent kinase inhibitor 2A; DNAH14, dynein axonemal heavy chain 14; E4F1, E4F transcription factor 1; FLG, filaggrin; GNAS, GNAS complex locus; KRAS, KRAS proto-oncogene, GTPase; MUC4, mucin 4, cell surface associated; OBSCN, obscurin, cytoskeletal calmodulin and titin-interacting RhoGEF; OR10Z1, olfactory receptor family 10 subfamily Z member 1; PBRM1, polybromo 1; POM121L12, POM121-like protein 12; PRRC2C, proline-rich coiled-coil 2C; RALGAPB, Ral GTPase activating protein non-catalytic β subunit; RNF43, ring finger protein 43; SMAD4, SMAD family member 4; SMARCA2, SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 2; TGFBR2, transforming growth factor-β receptor 2; TP53, tumour protein p53; TTN, titin.

characteristics, whereas Waddell et al (23) identified a subtype of PDAC that was unstable with a high prevalence of inactivation of DNA maintenance genes, and was associated with therapeutic responsiveness for platinum-based chemotherapy and poly (ADP-ribose) polymerase inhibitors (21-23).

However, despite improvements regarding the biological understanding of PDAC, there is currently no available

molecular classification able to translate the histopathological classification into clinical practice. Molecular heterogeneity, due to high PDAC genetic instability, hinders the identification of prognostic or predictive biomarkers (24,25).

The present data confirmed this heterogeneous landscape, as characterised by a high number of mutations occurring with low frequency, with the exception of well-known mutations in KRAS, TP53, CDKN2A and SMAD4.

As expected, KRAS mutations were the most prevalent events detected in the present study. Numerous hotspot mutations were identified (G12D, G12V, G12R and G12C) in KRAS, which is a key oncogene during the onset of pancreatic cancer (26). A recent study demonstrated how numerous subclonal KRAS mutations coexist within tumour cells, suggesting a convergent evolution of various clones of advanced cancer with independent KRAS mutations (6). In the present study, the majority of single gene mutations detected in the samples can be grouped into common cellular pathways, including KRAS signalling, TGF-β signalling, chromatin remodelling, Wnt signalling, DNA damage repair, cell cycle and RNA processing; these findings are concordant with those of previous studies (7,27).

The 5-year OS of patients with surgically resected PDAC is 20-25%, and the majority of patients develop recurrence at a median of 10-20 months following resection (2).

By evaluating mutational load in the early stage subgroup the present study demonstrated that samples harbouring more genetic events were associated with worse OS, thus stressing the urgent requirement for more molecular features to support the clinical-pathological classification of PDAC. It may be hypothesized that tumours with high molecular events have a more aggressive phenotype; therefore, detection of these patients during the early diagnostic phase may identify patients that benefit from neoadjuvant treatment prior to surgery. However, no specific genes or pathways were revealed to be correlated with OS in this subgroup of patients. SMAD4 gene inactivation, which has previously been reported to be associated with recurrence and a poorer prognosis in patients following resection, was detected in 27% of the samples, but was not correlated with OS correlation (28,29).

In the present study, a high prevalence of MUC4 mutation was detected in samples from patients following resection (33%), thus confirming its importance in PDAC carcinogenesis. MUC4 is a type I membrane-bound mucin that promotes proliferation, motility, invasiveness, epithelial-mesenchymal transition, chemoresistance and tumour growth, the overexpression of which is considered an early event in pancreatic carcinogenesis, and is detected in the preneoplastic stage. Urey et al (30) investigated MUC4 expression by immunohistochemistry and indicated that low MUC4 expression is associated with a survival benefit in patients with resectable pancreatic cancer receiving adjuvant gemcitabine (30,31). However, no differences in MUC4 mutations were detected between samples in the long and short survival groups in the present study.

RNF43 deletion was detected in only two cases with intraductal papillary mucinous neoplasm (IPMN)-associated PDAC in the long survival group. Inactivating mutations of the RNF43 gene, which encodes a protein with intrinsic U3 ubiquitin ligase activity, have been reported in IPMN of the

Table II. Early-stage patient characteristics.

Variable	Value
Age (years)	
Median	67.6
Range	33-83
Sex (%)	
Male	53
Female	47
Outcome	
Death (%)	67
Survival (%)	33
Median OS (months)	23.5
Range OS (months)	7.66-67.09
T stage (%)[a]	
T1 (1a, 1b, 1c)	0
T2	7
T3	93
N stage (%)[a]	
N0	33
N1	20
N2	47
Grading (%)	
G1	20
G2	60
G3	20
Margins (%)	
Clear	47
Involved	53
Tumour location (%)	
Head	67
Body/tail	33
Type of surgery (%)	
Total pancreatectomy	47
Pancreaticoduodenectomy	27

[a]Classified according to the 8th edition of the American Joint Committee on Cancer staging system (34). OS, overall survival.

pancreas. According to a previous study (32), IPMN has a better outcome than conventional PDAC, and the present detection of RNF43 deletions in the long survival group only supports this hypothesis. Recent evidence has suggested that mutational inactivation of RNF43 confers Wnt dependency, thus indicating that the presence of RNF43 mutations may be used as a predictive biomarker for the use of Wnt inhibitors (33).

In conclusion, the present study reaffirmed the presence of a heterogeneous genetic landscape in PDAC, and provided evidence to suggest that a more complex genome is correlated with a more aggressive phenotype and poor prognosis. Further work is required to identify molecular biomarkers that better characterise pancreatic cancer, particularly in the early stage.

Acknowledgements

Funding

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Authors' contributions

EG, SD, SV, EF, GB and MDM designed and supervised the research. CR, RC, FM, EG, CS, DS, FF, DF, FC and ADE provided the samples. SV, SD and EG collected the specimens. SD, VI, GT and AA performed sequencing, bioinformatics and statistical analysis. EG, EF and MDM wrote the manuscript. All authors have read and approved the final manuscript.

Ethics approval and consent to participate

The present study was conducted according to the principles expressed in the Declaration of Helsinki and written informed consent was obtained from all participants. The study was approved by Independent Ethics Committee of Sant' Orsola-Malpighi Hospital (Bologna, Italy).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Vogelzang NJ, Benowitz SI, Adams S, Aghajanian C, Chang SM, Dreyer ZE, Janne PA, Ko AH, Masters GA, Odenike O, *et al*: Clinical cancer advances 2011: Annual Report on Progress Against Cancer from the American Society of Clinical Oncology. J Clin Oncol 30: 88-109, 2012.
2. Katz MH, Wang H, Fleming JB, Sun CC, Hwang RF, Wolff RA, Varadhachary G, Abbruzzese JL, Crane CH, Krishnan S, *et al*: Long-term survival after multidisciplinary management of resected pancreatic adenocarcinoma. Ann Surg Oncol 16: 836-847, 2009.
3. Conroy T, Desseigne F, Ychou M, Bouché O, Guimbaud R, Bécouarn Y, Adenis A, Raoul JL, Gourgou-Bourgade S, de la Fouchardière C, *et al*; Groupe Tumeurs Digestives of Unicancer; PRODIGE Intergroup: FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med 364: 1817-1825, 2011.
4. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, Seay T, Tjulandin SA, Ma WW, Saleh MN, *et al*: Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med 369: 1691-1703, 2013.
5. Meyerson M, Gabriel S and Getz G: Advances in understanding cancer genomes through second-generation sequencing. Nat Rev Genet 11: 685-696, 2010.
6. Cancer Genome Atlas Research Network. Electronic address: andrew_aguirre@dfci.harvard.eduCancer Genome Atlas Research Network: Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell 32: 185-203. e13, 2017.
7. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, Mankoo P, Carter H, Kamiyama H, Jimeno A, *et al*: Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 321: 1801-1806, 2008.
8. Schubert M, Lindgreen S and Orlando L: AdapterRemoval v2: Rapid adapter trimming, identification, and read merging. BMC Res Notes 9: 88, 2016.
9. Lang B, Trapnell C, Pop M and Salzberg L: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10: R25, 2009.
10. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, *et al*: 1000 Genome roject Data Processing Subgroup. The SequenceAlignment/Map format and SAMtools. Bioinformatics 25: 2078-2079, 2009.
11. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, *et al*: The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20: 1297-1303, 2010.
12. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES and Getz G: Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 31: 213-219, 2013.
13. Goya R, Sun MGF, Morin RD, Leung G, Ha G, Wiegand KC, Senz J, Crisan A, Marra MA, Hirst M, *et al*: SNVMix: Predicting single nucleotide variants from next-generation sequencing of tumors. Bioinformatics 26: 730-736, 2010.
14. Wang K, Li M and Hakonarson H: ANNOVAR: Functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 38: e164, 2010.
15. Ramos AH, Lichtenstein L, Gupta M, Lawrence MS, Pugh TJ, Saksena G, Meyerson M and Getz G: Oncotator: Cancer variant annotation tool. Hum Mutat 36: E2423-E2429, 2015.
16. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, *et al*: Catalog of Somatic Mutations in Cancer and bioinformatics mutation-prediction tools PolyPhen2. Nat Methods 7: 248-249, 2010.
17. Choi Y and Chan AP: PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31: 2745-2747, 2015.
18. Ng PC and Henikoff S: SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res 31: 3812-3814, 2003.
19. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, Janoueix-Lerosey I, Delattre O and Barillot E: Control-FREEC: A tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics 28: 423-425, 2012.
20. Amarasinghe KC, Li J, Hunter SM, Ryland GL, Cowin PA, Campbell IG and Halgamuge SK: Inferring copy number and genotype in tumour exome data. BMC Genomics 15: 732, 2014.
21. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, Cooc J, Weinkle J, Kim GE, Jakkula L, *et al*: Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med 17: 500-503, 2011.
22. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, Miller DK, Christ AN, Bruxner TJ, Quinn MC, *et al*; Australian Pancreatic Cancer Genome Initiative: Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 531: 47-52, 2016.
23. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, *et al*; Australian Pancreatic Cancer Genome Initiative: Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 518: 495-501, 2015.
24. Di Marco M, Astolfi A, Grassi E, Vecchiarelli S, Macchini M, Indio V, Casadei R, Ricci C, D' Ambra M, Taffurelli G, *et al*: Characterization of pancreatic ductal adenocarcinoma using whole transcriptome sequencing and copy number analysis by single-nucleotide polymorphism array. Mol Med Rep 12: 7479-7484, 2015.
25. Samuel N and Hudson TJ: The molecular and cellular heterogeneity of pancreatic ductal adenocarcinoma. Nat Rev Gastroenterol Hepatol 9: 77-87, 2011.
26. Collins MA and Pasca di Magliano M: Kras as a key oncogene and therapeutic target in pancreatic cancer. Front Physiol 4: 407, 2014.

27. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J, *et al*; Australian Pancreatic Cancer Genome Initiative: Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 491: 399-405, 2012.
28. Blackford A, Serrano OK, Wolfgang CL, Parmigiani G, Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Eshleman JR, *et al*: SMAD4 gene mutations are associated with poor prognosis in pancreatic cancer. Clin Cancer Res 15: 4674-4679, 2009.
29. Shin SH, Kim HJ, Hwang DW, Lee JH, Song KB, Jun E, Shim IK, Hong SM, Kim HJ, Park KM, *et al*: The DPC4/SMAD4 genetic status determines recurrence patterns and treatment outcomes in resected pancreatic ductal adenocarcinoma: A prospective cohort study. Oncotarget 8: 17945-17959, 2017.
30. Urey C, Andersson B, Ansari D, Sasor A, Said-Hilmersson K, Nilsson J and Andersson R: Low MUC4 expression is associated with survival benefit in patients with resectable pancreatic cancer receiving adjuvant gemcitabine. Scand J Gastroenterol 52: 595-600, 2017.
31. Jonckheere N, Lahdaoui F and Van Seuningen I: Targeting MUC4 in pancreatic cancer: miRNAs. Oncoscience 2: 799-800, 2015.
32. Koh YX, Chok AY, Zheng HL, Tan CS and Goh BK: Systematic review and meta-analysis comparing the surgical outcomes of invasive intraductal papillary mucinous neoplasms and conventional pancreatic ductal adenocarcinoma. Ann Surg Oncol 21: 2782-2800, 2014.
33. Jiang X, Hao HX, Growney JD, Woolfenden S, Bottiglio C, Ng N, Lu B, Hsieh MH, Bagdasarian L, Meyer R, *et al*: Inactivating mutations of RNF43 confer Wnt dependency in pancreatic ductal adenocarcinoma. Proc Natl Acad Sci USA 110: 12649-12654, 2013.
34. Brierley JD, Gospodarowicz MK and Wittekind C (eds): UICC TNM Classification of Malignant Tumours. 8th edition. Wiley-Blackwell, Oxford, 2016.