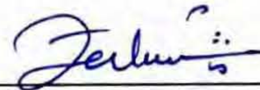
CERTIFICATE

This thesis, submitted by **Ms. Iqra Bibi** to the Department of Biochemistry, Faculty of Biological Sciences, Quaid-i-Azam University, Islamabad, Pakistan, is accepted in its present form as satisfying the thesis requirement for the Degree of Master of Philosophy in Biochemistry/Molecular Biology.

Supervisor:

Dr. Aneesa Sultan
Associate Professor

External Examinar:



Dr. Zertashia Akram
Associate Professor
Department of Biosciences
COMSATS University, Islamabad.

Chairperson:

Dr. Iram Murtaza
Professor

Dated: November 1, 2022

Investigating the Role of ZEB1 and YAP1 in Pancreatic Ductal Adenocarcinoma through Whole Exome Sequencing

DRSML

Submitted By

Iqra Bibi

Department of Biochemistry,

Faculty of Biological Sciences

Quaid-i-Azam University Islamabad, Pakistan

2020-2022

Investigating the Role of ZEB1 and YAP1 in Pancreatic Ductal Adenocarcinoma through Whole Exome Sequencing

A report submitted in fulfilment of requirements for

The degree of Master of Philosophy in

Biochemistry

By

Iqra Bibi

Department of Biochemistry,

Faculty of Biological Sciences

Quaid-i-Azam University Islamabad, Pakistan

2020-2022

DRSML QAU

Declaration

I hereby declare that any part of this dissertation is not plagiarized, and all work was carried out in accordance with regulation set by the Quaid-i-Azam University. No part of this thesis has been previously presented for any other degree. If found anything contrary, I shall be held responsible.

Iqra bibi

DRSML QAU

Certificate

This thesis entitled as "**Investigating the Role of ZEB1 and YAP1 in Pancreatic Ductal Adenocarcinoma through Whole Exome Sequencing**" submitted by **Iqra Bibi** is accepted in its present form by the Department of Biochemistry, Quaid-i-Azam University Islamabad, Pakistan; as a requirement for the degree of Master of Philosophy in Biochemistry/Molecular Biology.

Supervisor ______________________________

Dr. Aneesa Sultan

Associate Professor

External Examiner ______________________________

Assistant Professor

Chairperson ______________________________

Dr. Irum Murtaza

Professor

Dated ______________________________

DRSML QAU

Dedicated To,

My Supportive Parents and Sisters

DRSML QAU

Acknowledgements

First and foremost, praises and thanks to **ALLAH**, the Almighty, for His countless blessings throughout the research journey to complete the research successfully. Countless salutations upon our **Holy Prophet Hazrat Muhammad (PBUH)**, for showing the right path of life to the mankind.

I am incredibly grateful to the Chairman, Department of Biochemistry **Dr. Iram Murtaza** for providing us with a healthy research environment. It is a genuine pleasure to express my deep and sincere gratitude to my mentor and guide, **Dr. Aneesa Sultan,** for providing me the opportunity, invaluable guidance, and motivation throughout this research. It was a great honor to work and study under her supervision. I am extremely grateful for all the support and guidance she has offered me.

I am very thankful to the volunteers of the study as their contribution made the tasks possible. I extend my deepest gratitude to my PhD seniors **Maryam Naeem** and **Usman Tareen** senior for their sincere guidance and support throughout the research journey until completion. I am very grateful to my dear friends **Arooba Azeem, Aroma Tahir, Amna Sajid** and other fellows for their help and support throughout the research journey. I am highly thankful to staff of Department of Biochemistry for their support and guidance.

Last but not the least, I would like to pay my deepest gratitude to my parents **Aman Ullah** and **Roshaan Bibi** for their unconditional love, moral support, efforts, prayers and the trust they have in me. I am thankful to my sisters for their love, moral and emotional support without which I would not have been able to come this far. I am thankful to every person who helped me in this whole tenure.

DRSML QAU

Table of Contents

List of Figures ……………………………………………………………..……iv

List of Tables ………………………………………………………………….....v

List of Abbreviations ……………...vi

Abstract………………………………………………………………….........vii

Chapter 01. Introduction……………...1

1.1 Cancer ....…………………………………..………………………………..1

1.2 Pancreatic Cancer ....……………………………………………………….…1

1.3 Types of pancreatic cancer ....……………………………………….……….. …2

1.3.1 Pancreatic Ductal Adenocarcinoma (PDAC) …………………………………..3

1.3.1 Pancreatic Acinar Cell Carcinoma (PACC)....……………………………………3

1.3.3 Squamous Cell Carcinoma ....……………………………………………......4

1.4 Molecular subtypes of pancreatic cancer ....…………………………………….…5

1.4.1 Squamous Subtype ....……………………………………………………….5

1.4.2 Pancreatic Progenitor Subtype ....………………………………….……….. …6

1.4.3 Immunogenic Subtype ....………………………………………………..….6

1.4.4 ADEX subtype ....…………………………………..…………………….6

1.5 Symptoms of pancreatic cancer ....………………………………………….....6

1.6 Risk Factors of Pancreatic cancer ....…………………………………… ...…7

1.7 Diagnosis of Pancreatic cancer ....……………………………………...……..7

1.7.1Blood tests ....……………………………………..………………………....7

1.7.2 Computed Tomography (CT) scan ....……………………………….………...8

1.7.3 Magnetic Resonance Imaging (MRI) ....………………………………………8

1.7.4 Cholangiopancreatography ....…………………………………….……….. …9

1.8 Treatment and Therapeutic Strategies …………………...............................9

DRSML QAU

1.8.1 Surgery .....……………………………………………………………...…9

1.8.2 Chemotherapy .....…………………………………………………...…10

1.8.3 Radiotherapy .....……………………………………………….……...…10

1.9 Zinc finger E-box binding homeobox 1 (ZEB1) .....……………..………………11

1.9.1 Structure of ZEB1 .....……………………………………………..……..…11

1.9.2 ZEB1 Role in Cancer .....………………………………………..……………..11

1.9.3 ZEB1 Role in Pancreatic ductal carcinoma (PDAC) ………………..……..12

1.10 YES-Associated Protein 1 (YAP1) .....…………………………………………13

1.10.1 Structure of YAP1 .....…………………………………………………...…13

1.10.2 Role of YAP1 in tissue development and cancer ……………………………14

1.10.3 Role of YAP1 in PDAC .....……………………………...…………..…16

1.11 Driver Genes Mutations in PDAC ……………………………………………17

1.12 Aims and Objectives .....……………………...…………………………...…19

Chapter 02. Materials and methods…………………………………………………20

2.1 Ethical Approval .....…………………………………………...……………..…20

2.2 Tissue sample collection .....……………………………………...…………....20

2.3 DNA Extraction from tissue .....…………………………………………....…20

2.4 DNA Quantification .....……………………………………..………………...23

2.5 Agarose Gel Electrophoresis .....……………………………………..………..24

2.6 Whole exome Sequencing .....……………………………………..…………..25

2.7 Bioinformatics analysis .....………………………………………………….…..…26

2.7.1 Raw data import .....……………………………………………………....…26

2.7.2 Quality checking .....……………………………………………………...…26

2.7.3 Adapters' removal .....……………………………………..…………..…26

2.7.4 Mapping to the reference genome .....…………………..……………………26

2.7.5 Duplicates marking .....……………………………………………………...…26

DRSML QAU

2.7.6 Variant calling by MuTect2 .....…………………………………..…..…27

2.7.7 Annotation of variants .....……………………………………..…………...27

2.8 SIFT, PROVEAN and Polyphen2 .....……………………………………..….27

2.9 Mutation Tester and Mutation Assessor .....……………………………………..28

2.10 Damage Pred Count .....…………………………………..……………….28

2.11 CADD Raw and CADD Phred .....……………………………………..….28

Chapter 03. Results……………………………………………………..……...29

3.1 Demographic Details .....………………………………………………...….29

3.1.1 Stages of PDAC Patients .....…………………………………………....… 29

3.2 Quantitative Analysis .....………………………………………..……………….30

3.3 Qualitative Analysis .....…………………………………………………..…..….31

3.4 WES Results analysis .....……………………………………………...……..….32

3.4.1 Quality check analysis of sample Raw reads .....………………………………32

3.4.2 Adapter content Analysis .....………………………………………………..….33

3.4.3 Variants' detection of PDAC .....…………………………………………..…...….33

3.5 Analysis of Annotated variants .....…………………………………………..…..….34

3.5.1 Nucleotide variation in ZEB1 .....…………………………………………...….34

3.5.2 Nucleotide variation in YAP1 .....……………………………………...…..….34

3.5.3 Nucleotide Variation in driver genes .....………………………………………35

3.5.4 Amino acid substitution analysis .....……………………………………....….36

3.5.5 Damaging or tolerant nature of variants .....………………………..……….36

3.5.6 Scoring the damaging potential of variation .....……………………….………36

3.5.7 Deleteriousness of variants .....………………………………………..….36

Chapter 04. Discussion………......…………………………………………..…... 37

Chapter 05. Conclusion……………………………………………………………..39

Chapter 06. References…………………………………………………...…...40

DRSML QAU

List of Figures

1.1 Schematic representation of trends, risk factors and burden of pancreatic cancer across the globe ……………………………………………………………..….2
1.2 Anatomy of Human pancreas …………………………………………………2
1.3 Morphologic changes during PDAC ………………………………………..…4
1.4 Schematic representation of ZEB1 protein ……………………….……….11
1.5 Schematic representation of YAP1 protein ……………………………..…14
1.6 Role of YAP1 in Hippo Signaling Pathway …………………………………15
1.7 Schematic illustration of ZEB1 and YAP1 interaction ……...………………16
2.1 Schematic Illustration of WES data analysis ……………………………………25
3.1 Cancer Stages Frequency in PDAC Patients ……………………………………30
3.2 Gel electrophoresis image …………………………………………..…….31
3.3 Basic statistics of Raw Reads using FASTQ …………………….……………32
3.4 Quality score graph of FastQ Tool ………………………………….……….32
3.5 Adapter Content of sequence ………………………………………………..33
3.6 Single nucleotide variation in Zeb1 gene ……………………………………34
3.7 Single nucleotide variation in YAP1 gene ……………………..…………35
3.8 SNV in Econ 5 of KRAS gene …………………………………………...… 35
3.9 SNV in Exon 2 of p53 ……………………………………………………..…35
3.10: SNV in Econ 6 of SMAD4 gene ……………………………………………36

DRSML QAU

List of Tables

2.1: List of Reagents for DNA extraction and their Chemical Composition ………..20

2.2: Chemical Composition of Reagent used in Agarose Gel Electrophoresis……....23

2.3: Reference genomes used for alignment Specie ……………………………….....25

3.1: Demographic details of study subjects ……………………………………….29

3.2: Quality and Concertation of extracted DNA ……………………………………30

3.3: Number of variants in study samples ………………………………………...33

DRSML QAU

List of Abbreviations

ADEX	Aberrantly Differentiated Endocrine Exocrine
BWA	Burrows-Wheeler Aligner
CA-19	Cancer Antigen 19
CADD	Combined Annotation Dependent Depletion
CBC	Complete Blood Count
CEA	Carcinoembryonic Antigen
CP	Chronic Pancreatitis
CT-Scan	Computed Tomography Scan
EDTA	Ethylenediamine Tetra Acetic Acid
EMT	Epithelial to Mesenchymal Transition
ERCP	Endoscopic Retrograde Cholangiopancreatography
EtBr	Ethidium Bromide
FNA	Fine Needle Aspiration
FU	5-Fluorouracil
GATK	Genome Analysis Toolkit
GEM	Gemcitabine
HbA1c	Glycated Hemoglobin
IPMN	Intraductal Papillary Mucinous Neoplasm
IRB	International Review Board
LFT	Liver Function Test
LncRNA	Long Non-Coding RNA
MCN	Mucinous Cystic Neoplasm

DRSML QAU

MRI	Magnetic Resonance Imaging
NGS	Next Generation Sequencing
PanINs	Pancreatic Intraepithelial Neoplasia
PanNETs	Pancreatic Neuroendocrine Tumor
PC	Pancreatic Cancer
PDAC	Pancreatic Ductal Adenocarcinoma
PET	Positron Emission Tomography
PolyPhen	Polymorphism Phenotyping
PROVEAN	Protein Variation Effect Analyzer
RFT	Renal Function Test
SCC	Squamous Cell Carcinoma
SDS	Sodium Dodecyl Sulphate
SIFT	Scale Invariant Feature Transform
SNP	Single Nucleotide Polymorphism
TCGA	The Cancer Genome Atlas
TID	TEAD Interaction Domain
TLR	Toll Like Receptor
WES	Whole Exome Sequencing
YAP-1	Yes-Associated Protein 1
ZEB1	Zinc finger E-box binding homeobox 1
ZEB1-AS1	ZEB1 Antisense 1

DRSML QAU

Abstract

Pancreatic cancer is one of the deadliest among cancer types, ranked as seventh leading cause of cancer mortality. Pancreatic Ductal Adenocarcinoma (PDAC), the most lethal and common pancreatic malignancy with poor survival rate of <5 year and fourth leading cause of cancer related deaths worldwide. It develops in pancreatic ducts and rapidly metastasized to nearby organs including lungs and liver. Different mutations in driver genes including KRAS, Tp53, SMAD4, and CDKN2A have reported to be actively involved in poor prognosis of PDAC. This study aims to identify the role of ZEB1 and YAP1 in PDAC cases with the objective of genetic characterization in freshly resected tumor tissue samples matched with control through Whole Exome Sequencing. Analysis of sequencing data using bioinformatics tools, have presented a novel substitutive nucleotide variation (31309911-12; GG>TA) in ZEB1-AS1, while no variation have been observed in exonic region of YAP1 genes. However, multiple single nucleotide variations (SNVs) and indels were present in intronic and intergeneric regions of both genes that may held responsible for their upregulation and active role in promotion of tumorigenesis of PDAC. Furthermore, a novel SNV have been found in exonic region of SMAD4 (51058160; G>A), KRAS (51058160; G>A) and in p53 (7674945; G>A) while no variation have been observed in exonic region of CDKN2A. In silico studies using multiple tools including, SIFT, Polyphen2, PROVEAN, Mutation tester, CADD Raw and Clin Pred have confirmed the novelty of identified variants. Transcriptome analysis and molecular docking can further help to validate the results and pave the way to devise diagnostic marker for early detection and promising therapeutic approaches.

Keywords: Pancreatic Ductal Adenocarcinoma (PDAC), Zinc finger E-box binding homeobox 1 (ZEB1), ZEB1 antisense 1 (ZEB1-AS1), Yes-Associated Protein 1 (YAP1**),** Whole Exome Sequencing (WES)

DRSML QAU

1. Introduction

1.1 Cancer

Cancer, a Non-communicable disease (NCDs) that is generally known as abnormal division of cells due to multiple underlying causes such as genetic mutations that accumulated overtime and mostly results into death. According to GLOBOCAN report, 2020 overall the incidence of cancer cases have increased to 19.3 million and the mortality rate has risen to 10.0 million across the globe in 2020. In addition to other severe impacts on human health and socio-economic status of people worldwide, the pandemic of Corona Virus (COVID-19) has also significantly affected the registration process in PBCR, in developing and under-developed countries that had caused delayed reporting and eventually affect corresponding rate of incidence years before to 2020 (Ferlay*et al*., 2021).

The increased mortality rate can be credited to poor prognostic rate of many cancers, the asymptomatic nature, unavailability of efficient diagnostic tools and therapeutic options. There is an immense need to provide awareness about the severity of disease and to promote the regular annual medical checkups among masses to get the disease diagnosed at earlier stages, ensuring the provision of appropriate treatment and eventually increase the life expectancy.

1.2 Pancreatic Cancer

Pancreatic cancer, one of the deadliest among cancer types and well known for its poor prognosis. There is a parallel association between disease incidence and mortality rates, and 1 to 5-year survival rate is reported that only accounts for 10% of pancreatic cancer patients in the United States (Viale, 2020). The overall incidence rate of pancreatic cancer has risen to 458,918 cases while approximately 432,242 deaths have been reported by Haung *et al*., (2021). Recent studies have reported the 0.5% to 1.0% per year increase in PDAC incidence and it is expected to turn out as the second-leading cause of cancer-associated deaths by 2030. (Park *et al*., 2021).

DRSML QAU

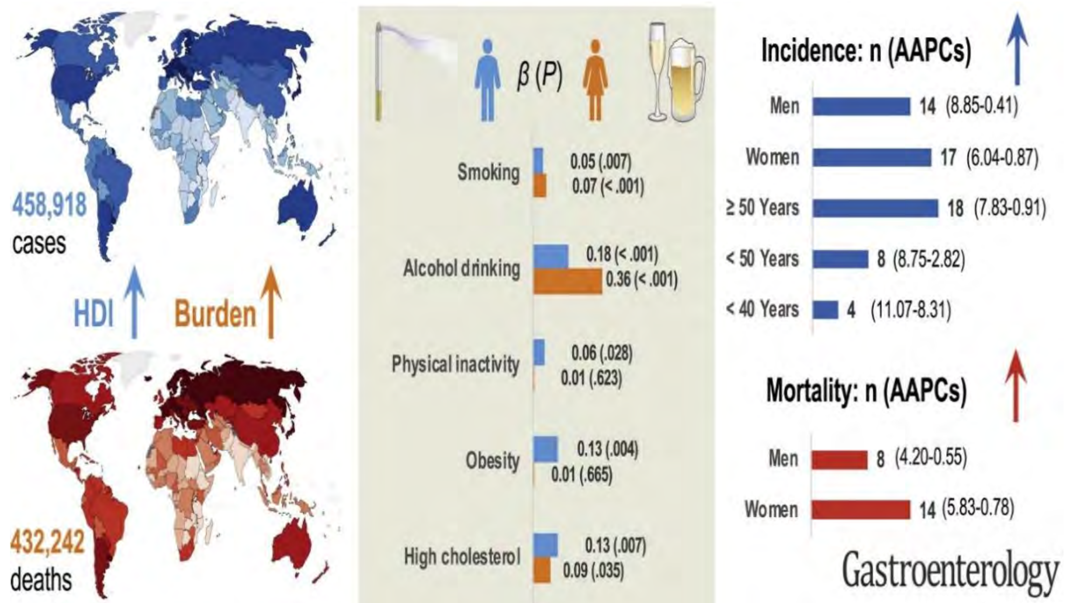


Figure 1.1: *Schematic representation of trends, risk factors and burden of pancreatic cancer across the globe* (Huang *et al*., 2021)

There are multiple factors that can be attributed for low survival rate including the most important is the delayed diagnosis. Pancreatic cancer is mostly asymptomatic in nature in majority patients until the disease progression to an advanced stage. The general symptoms include jaundice, abdominal pain, abnormality in liver functioning, early onset of diabetes, vomiting, dyspepsia, back pain, and unintended weight loss (Schmidt *et al*., 2016).

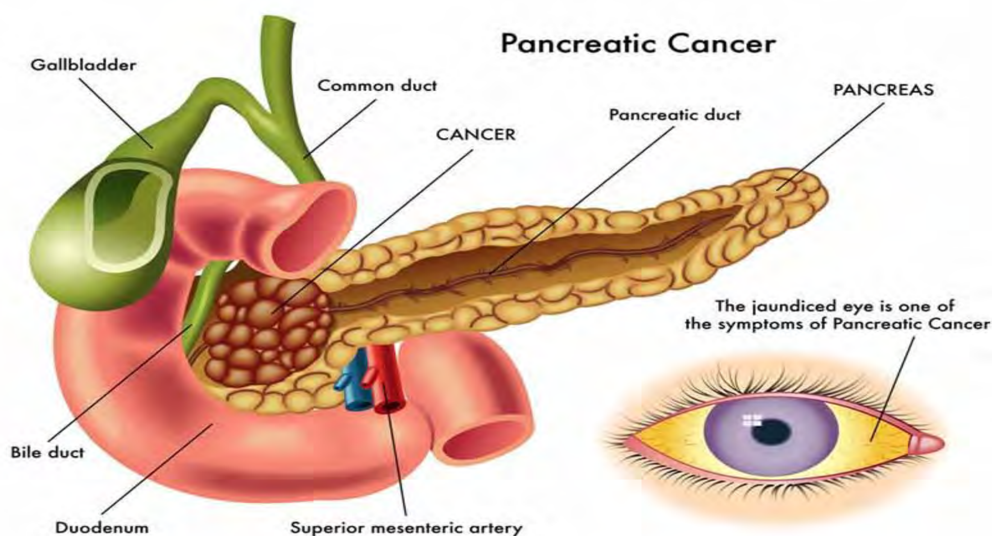


Figure 1.2: *Anatomy of Human pancreas (adapted from web)*

Mostly the surgical resection of tumor is being performed as a therapeutic option but even after that, most of the patients have experienced the reoccurrence and only 20% of completely resected patients experiences the 5-year survival (Viale, 2020). Furthermore, the pancreatic tumors biology mediates the early recurrence, metastasis, and resistance to chemo and radiotherapy (Kamisawa *et al*., 2016).

1.3 Types of Pancreatic Cancer

There are two major types of pancreatic cancer categorized based on origin. These include Exocrine (Non-endocrine) and Neuroendocrine pancreatic cancer. Exocrine as the name implies originates from exocrine cells, of pancreatic ducts and exocrine gland. Several digestive enzymes are secreted by exocrine gland that facilitates digestion of carbohydrates, proteins, and fats in duodenum. On the other hand, Neuroendocrine originates from pancreatic endocrine gland, producing insulin and glucagon hormone, critical for blood sugar maintenance. Exocrine pancreatic cancers are more abundant and accounts for >95 percent while neuroendocrine malignancies are comparatively uncommon, accounting for < 5% of all pancreatic cancers (Serafini *et al*., 2017; Liao *et al*., 2021)

DRSML QAU

1.3.1 Pancreatic ductal adenocarcinoma (PDAC)

Pancreatic ductal adenocarcinoma (PDAC) that develops in lining of the pancreatic ducts. It is a most common type and accounts for about 90 percent of all pancreatic cancer. It develops from precursor lesions, of which pancreatic intraepithelial neoplasia are the most prevalent (PanINs). The buildup of genetic mutation especially in driver genes including KRAS, CDKN2A, Tp53 and SMAD4 is linked to the progression of these benign ductal lesions to in situ carcinomas and invasive malignancies (Soreide, 2018). PDAC is highly destructive and possess poor prognostic rate that can be credited to its increased metastatic potential. Currently the surgical removal is the only option but in advance cases, it is mostly avoided. Being heterogeneous in nature, PDAC is further divided into many subtypes based on differences of functional behavior in different preclinical models, rate of survival and response to chemotherapy.

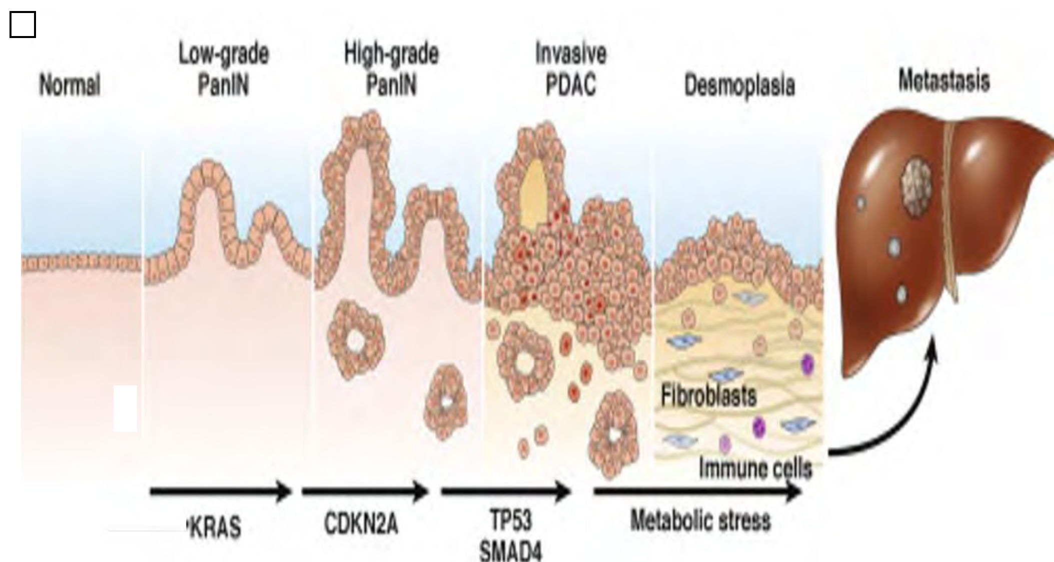


Figure 1.3: *Morphologic changes during PDAC* (Soreide, 2018)

DRSML QAU

Quasi-mesenchymal (QM) or basal-like/squamous is considered as most malignant molecular subtype of PDAC (Collisson *et al*., 2011). During epithelial–mesenchymal transition (EMT) process, these cancer cells adapted the mesenchymal phenotype. Different transcription factors, including SNAIL1 and SNAIL2/SLUG, causes the lower expression of epithelial markers such as E-cadherin and over-expression of mesenchymal markers such as vimentin that eventually stimulates the process of EMT. Wilson *et al*, (2020) demonstrated the facilitative role EMT in generation of cancer stem cells (CSCs) in many epithelial tissues. The stemness of cancer cells further increase the growth of tumor, metastasis as well as resistance to therapy and tumor relapse post-therapy (Rodriguez *et al*., 2019).

1.3.2 Pancreatic Acinar cell carcinoma (PACC)

It is another rare type of exocrine cancer accounting for less than 1% pancreatic malignancies. It develops from enzyme secreting acinar cells of pancreas. Molecular analysis of PACC cases revealed the germline BRCA2 and BRCA1 mutations and somatic BRCA1/2 alterations in approximately 7% and 16% of PACCs, respectively. It is similar in symptoms to adenocarcinoma such as nausea, severe abdominal pain, and unintended weight loss, except jaundice that rarely occurs in acinar cell carcinoma. In addition, skin rashes and joint pain is reported in some patients that is credited to increased level of enzymes (Kryklyva *et al*.,2019).

1.3.3 Squamous Cell Carcinoma

Of all exocrine pancreatic neoplasms, the incidence of primary squamous cell carcinoma of the pancreas ranges from 0.5 to 2 percent. Only case reports are used in the literature to describe SCC cases (Adachi *et al*., 2011). In patients who received curative resection, the median survival for pancreatic SCC was observed to be seven months, with a range of six to 16 months (Brown *et al*., 2005).

Other rare type of exocrine pancreatic cancer includes

- **Adeno-squamous Carcinoma**, which is a more aggressive tumor with a worse prognosis and possess characteristics of both ductal adenocarcinoma and squamous cell carcinoma.
- **Colloid carcinomas**, which develop from intraductal papillary mucinous neoplasm (IPMN), a benign cyst and comparatively better prognosis than other types of pancreatic cancer. Pancreatic colloid tumor contains floating malignant cells in a gelatinous substance called mucin, hence it is not frequently metastasized, increasing the chance of cure.

1.4 Molecular subtypes of pancreatic cancer

Pancreatic cancer has been further classified into four molecular subtypes based on involvement of different transcription factors (TFs) and mediators in pancreatic cancer progression. Transcriptome analysis has revealed 10 out of 26 differentially expressed genes that provides the basis for this classification (Collison *et al*., 2019). These subtypes includes:

1.4.1 Squamous Subtype

Squamous subtype comprises different genes involved in inflammation, metabolic reprogramming, TGF-β, EGF, and MYC signaling pathways (Collison *et al*., 2019). In addition to pancreatic cancer, other cancer types also possess higher expression of these genes as reported in TCGA database. Furthermore, mutated p53 along with upregulated p63 have reported to promote EMT and eventually the tumorigenicity of squamous subtype of pancreatic cancer.

DRSML QAU

1.4.2 Pancreatic Progenitor Subtype

Baily *et al*., (2016) have reported various TFs including FOXA2/A3, PDX1, HNF4G/4A and HNF1B/1A in progenitor subtype of pancreatic cancer. These TFs, especially PDX1 plays an active role in fate determination of endodermal cells of pancreas and regulate the development of ductal, endocrine, and exocrine of pancreas. Apomucins such as MUC1/5AC facilitates the clustering of PDAC linked IPMNs, increasing the severity of progenitor subtype.

1.4.3 Immunogenic Subtype

It possesses similar characteristics to progenitor subtype and actively involve TFs of immune regulation including B-cell pathway, CD4+ and CD8+ cells, Toll-like receptor pathway and antigen presenting cells. Induced upregulation of PD1 and CTLA4 can activate the tumor suppressor pathway and could be targeted as a therapeutic option to improve survival rate of pancreatic cancer (Collison *et al*., 2019).

1.4.4 ADEX subtype

This subtype involves signaling pathways critical for development and differentiation at advance stage of pancreatic cancer. Increased expression of NR5A2, MIST1, INS, MAFA, and NEUDO1 and their downstream effectors are actively involved in differentiation of acinar and endocrine cells of pancreas (Collison *et al*., 2019).

DRSML QAU

1.5 Symptoms of pancreatic cancer

Pancreatic cancer is well known for its asymptomatic nature that has severely reduced the rate of early diagnosis and eventually improves the death rate worldwide. Abdominal pain is one the frequent symptoms reported in 78-82% patients of pancreatic cancer (Li *et al*., 2013) and mostly confused to with gastrointestinal diseases but comparatively more severe in advance stages of cancer. Some other common symptoms includes, jaundice, lethargy, abnormal functioning of liver, new onset of diabetes, back pain, dyspepsia, nausea, and unintentional weight loss (Schmidt *et al*., 2016).

1.6 Risk Factors of Pancreatic cancer

- **Genetic factors:**

Although the familial basis of pancreatic cancer only accounts for about 10% but it can be an efficient risk factor for development of the disease (Hruban et al., 2010). Research studies have demonstrated that genetic syndromes, including atypical mole melanoma, Peutz-Jeghers, and Lynch syndrome harboring germline mutations in *CDKN2A*, *STK11* and, DNA repair genes, respectively accounts for familial pancreatic cancer (Kastrinos *et al*., 2009). Furthermore the germline mutations in *PRSS1* and *SPINK1*, known to cause hereditary pancreatitis. In addition, the *CFTR* gene, *PALB2*, ATM and *BRCA2*, have been also reported to induce familial pancreatic cancer (Jones *et al*., 2009 ; Robert *et al*., 2012).

- **Obesity**:

The National Institutes of Health (NIH) cohort study reported the higher incidence of pancreatic cancer in overweight and obese patients relative to patients with normal weight as per Body Mass Index (BMIs) value (Stolzenberg *et al*., 2013). Rebours *et al* (2015) has credited the fatty infiltration of pancreas for pancreatic intraepithelial neoplasia formation that further proceeds to occurrence of pancreatic ductal adenocarcinoma.

DRSML QAU

- **Diabetes and smoking:**

Diabetes have also been linked with pancreatic cancer development as Pereira *et al*, (2020) reported 5-4-fold increase in risk of pancreatic cancer in recently diagnosed diabetic patients relative to prolonged diabetic patients possessing only 1·5-fold increase. Yaun *et al*., (2017) have proposed the Smoking as potential factor to double the risk of pancreatic cancer and shows its involvement in 25% of PDAC cases.

1.7 Diagnosis of Pancreatic cancer

Following are some commonly used approaches to diagnose pancreatic cancer:

1.7.1 Blood tests

Different blood tests can be helpful in pancreatic cancer diagnose including:

- **Liver function tests,** as Jaundice or yellowing of the skin and eyes is an initial symptom of pancreatic cancer. Physicians usually recommend blood test to

check levels of bilirubin that is a chemical made by the liver to rule out the association of jaundice with abnormality in liver and to ensure the accurate diagnosis to eventually increase the rate of survival among cancer patients.

- **Tumor markers,** certain factors that gets increased in blood of cancer patients. For pancreatic cancer it includes **CA 19-9 and Carcinoembryonic antigen (CEA)**. Both cannot be use as final diagnostic approach as there increased level is not reported in all pancreatic cancer patients and might be credited for some other diseases such as biliary obstructions. However, their level can be helpful in confirmed PDAC cases and can be helpful to assess the effect of ongoing treatment or retrieval of cancer after complete resection (Pelzer *et al*., 2013; Mizrahi *et al*., 2020)
- **Other blood tests** including CBC or chemistry panel are also performed to evaluate general health status of patient such as renal and bone marrow functionality to identify their capability for major operation.

1.7.2 Computed Tomography (CT) scan

The CT scan is an initial and reliable diagnostic test of pancreatic cancer as it provides the sharp and detailed cross-sectional images of internal organs. It doesn't only show the clear image of pancreas but can also present the status of metastasis to other nearby organs and involvement of lymph nodes. It can eventually guide the physician to opt surgical resection or other therapeutic approach. Currently the **multiphase CT scan** or a **pancreatic protocol CT scan** is getting attention for pancreatic cancer diagnostic. It involves the administration of an intravenous (IV) contrast injection followed by multiple sets of scans at specific time intervals (Valls *et al*., 2002).

1.7.3 Magnetic Resonance Imaging (MRI)

Magnetic Resonance Imaging is used to make refined images of different body parts. It utilizes strong magnetic and radio waves hence can provide detailed picture of internal organs. **MR cholangiopancreatography (MRCP)** and MR **angiography (MRA)** are the recommended MRI types used to examine the use pancreatic ducts and the surrounding blood vessels, respectively.

DRSML QAU

1.7.4 Cholangiopancreatography

It is an emerging imaging technique that is being utilized to examine the ducts including pancreatic and bile ducts. It helps to determine that either pancreatic tumor has caused blockage, constriction or dilation of ducts. Based on different protocol, Cholangiopancreatography is of following two types:

- **Endoscopic retrograde cholangiopancreatography (ERCP):**

ERCP involves insertion of an endoscope down the throat, passing through the esophagus and stomach, reaching to duodenum. The hepatopancreatic duct commonly known as Ampulla of Vater, is checked for any constriction or blockage in ducts through X-rays. A small brush can also be attached with endoscopic tube to obtain cells as a biopsy sample or for stenting to keep the pancreatic or bile duct open if pressed by any tumor mass (Purnak *et al*., 2021).

- **Magnetic resonance cholangiopancreatography (MRCP):**

MRCP is like MRI scan and is a non-invasive approach. It does not employ any phase contrast dye and being non-invasive, it can only be used to just determine the physical appearance of ducts and could not help in stenting or getting biopsy samples (Elsherif *et al*., 2020).

- **Percutaneous transhepatic cholangiography (PTC):**

It is not commonly performed unless ERCP could not be practiced for some personal reason in some patients. PTC involves the insertion of a thin needle through abdominal skin and into a bile duct of liver followed by injection of a contrast dye and taking x-rays while its passing through the pancreatic and bile ducts. Like ERCP technique, it can also provide fluid or tissue samples or stenting. However, being more invasive in nature it may be more painful as compared to ERCP (Hubers *et al*., 2018).

1.8 Treatment and Therapeutic Strategies

1.8.1 Surgery

Currently the surgical resection is only best option that possess offers curative potential. Multidisciplinary team of medical specialists determines the resect ability

DRSML QAU

status after evaluation with high-quality cross-sectional imaging through CT scan or ERCP.

- **Pancreaticoduodenectomy** that is called ''Whipple procedure'', is mostly used to remove tumors spanning over head of pancreas. It also involves the removal of certain other regions including duodenum, proximal part of jejunum, gall bladder, a common bile duct and small part of the stomach as well.
- **Distal pancreatectomy**, often combined with splenectomy is used to remove tumors localized in body and tail of the pancreas. Research studies have revealed some beneficial aspects of automated approaches such as laparoscopy as they are safer than traditional approach of open surgery. It has been found to cause decreased blood loss, improved survival, and lower post-operative mortality (Kendrick and Cusati, 2010; Croome *et al*., 2014).

1.8.2 Chemotherapy

Systemic chemotherapy that usually aims to relive symptoms associated with pancreatic cancer and to improve life expectancy. Currently the Folfirinox, Gemcitabine, and the combination regimen have been reported to increase the overall survival up to year from months. Adjuvant therapy for about 6 months is prescribed after surgery to ease the disease symptoms and to prevent the reoccurrence of disease (Neoptolemos *et al*., 2001). The provision of chemotherapy after surgery is very critical and must be provided for précised duration at proper intervals as analysis of retrospective studies have reported improved survival in patients who started the adjuvant therapy between 28 and 59 days relative to those who started before or after of this duration (Ma *et al*., 2019).

1.8.3 Radiotherapy

Radiations are well known for killing of rapidly growing cancerous cells through inducing DNA damage, but it the cellular death of normal cells is a considerable risk associated with radiation therapy. Sultana *et al* (2007), reported the efficacy of External beam radiotherapy (EBRT) in combination with 5-fluorouracil (5-FU) or gemcitabine. Though this combination might enhance the risk of adverse effects on

DRSML QAU

normal cells but in locally advanced cases of PDAC has shown considerable improvement using chemoradiotherapy over the administration of EBRT alone.

1.9 Zinc finger E-box binding homeobox 1 (ZEB1)

ZEB1 (zinc finger E-box binding homeobox 1), an important transcription factor of **zfh** family, which is also known as **TCF8 or Delta EF1** (Caramel *et al*., 2018). It is a protein encoded by ZEB1 gene, found on human chromosome 10p11.22 that was initially discovered the brain of the rat embryo, with mutations or deletions leading to severe deformities in development of limbs, spinal cord, and brain.

1.9.1 Structure of ZEB1

As a transcription factor, the ZEB family binds to the 5′-CANNTG-3′ region of the E-promoter DNA to regulate transcription of several gens. ZEB1 protein comprises thousand amino acids in humans and multiple domains including two zinc-finger clusters at its N- and C-terminal ends, a homeodomain, a Smad interaction domain, and a C-terminal binding protein (CtBP) interaction domain in the middle region. CtBP participates in the control of ZEB1 activity through interacting with other regulatory factors as it cannot directly bind to DNA (Zhang *et al*., 2015).

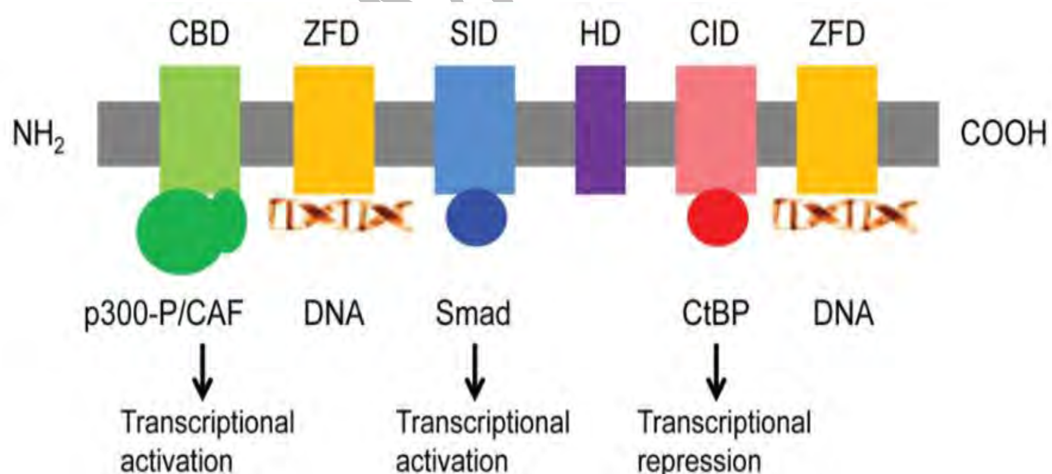


Figure 1.4: *Schematic representation of ZEB1 protein* (Zhang *et al*., 2015).

1.9.2 ZEB1 Role in Cancer

It plays a critical role in Epithelial-mesenchymal transition (EMT), a key step of embryogenesis and metastasis as well. EMT involves change in shape of epithelial cells, modification of molecules of adhesion that facilitates to obtain migratory and invasive potential (Lambert *et al*., 2017). ZEB1 actively participate in cytoskeletal

remodeling of basement membrane and improve invasive potential of tumor cells into surrounding tissues by downregulating factors of cell polarity and upregulating the matrix metalloproteases (MMPs) including MMP-1, MMP-9, and MMP-14 (Brabletz S and Brabletz T, 2010).

Induced expression of ZEB1 in different cancer types such as colon, lung, and breast cancer of human have resulted in increased invasive and migratory capabilities of tumor cells *in vitro* and metastatic potential *in vivo*. Chaffer *et al*., (2013) have presented the role of ZEB1 in conversion of non-cancer stem cells (CSC) to CSC in response to TGFβ production. CSC possess self-renewal and regenerating ability that increase the severity of breast cancer. ZEB1 blocks the expression of miR-200 and other stemness-repressing miRNAs such as miR-183 and miR-203, to regulate the stemness of cancer cells that eventually increases tumorigenicity (Wellner *et al*., 2009).

1.9.3 ZEB1 Role in Pancreatic ductal carcinoma (PDAC)

It involves in initiation of pancreatic tumorigenesis as its overexpression have been found in noninvasive lesion of pancreatic cells that is also termed as pancreatic intraepithelial lesions (PanIN) (Zheng *et al*., 2015). Krebs *et al*., (2017) has demonstrated the reduced number and grading of PanINs and acinar ductal metaplasia in response to Zeb1 depletion in mice, indicating the significance of ZEB1 for regional invasion, colonization abilities, and metastasis.

DRSML QAU

ZEB1 has been found to collaborate with ELK3 in PDAC. ELK3 that is a member of the ternary complex factor (TCF), and forms a ternary complex with serum response factor (SRF) to regulate expression of genes (Li *et al*., 2000). Kong *et al* (2016) have revealed the higher ELK3 levels in breast cancer while Yang et al (2015) have presented the negative effect of ELK3 knockdown on growth of tumor and progression of squamous cell carcinoma. Analysis based on using JASPAR database revealed that promoter of ELK3 possesses binding motifs and five potential binding sites for ZEB1. Furthermore, Zeb1 has been found to increase mRNA and protein levels of ELK3 by binding as confirmed by results of QRT-PCR and western blot. The binding of zeb1 to ELK3 promoter between 641 and 631 bp is credited to be responsible for zeb1 induced expression of ELK3 as confirmed by ChIP-qPCR results (Zhao *et al*., 2021).

Zhao *et al* (2021) have further proposed the effect of ELK3 on TGFβ as its depletion has been found to inhibit TGFβ-induced healing, migration, and invasive potential of cell wound, suggesting the role of ELK3 as a tumor oncogene causing progression of pancreatic cancer.

Long non-coding RNA (lncRNA) have drawn the attention of researcher as an efficient regulator of an EMT process through interaction with various micro-RNAs. For example, the ZEB1 antisense 1 (ZEB1-AS1), the lncRNA originated from ZEB1 promoter region and critical for its expression regulation. ZEB1-AS1 sequesters the miR-101 which can directly target and blocks the expression of Zeb1 expression in colorectal cancer (CRC), thus maintains the higher levels of ZEB1 to eventually promotes EMT and increased progression towards advance stage of cancer. This finding was further supported by induced overexpression miR-101 combined with of ZEB1-AS1 inhibition has significantly decreased the proliferative and migratory potential of cancel cells in CRC (Xiong *et al*., 2018).

DRSML QAU

1.10 YES-Associated Protein 1 (YAP1)

Yes-Associated protein 1 (YAP1) that is also known as YAP or YAP65. It is a protein acting as a coregulator of transcription factors that are involved proliferation of cells, downregulating the apoptotic genes, maintenance of normal growth of organs, homeostasis of tissues and cancer as well.

1.10.1 Structure of YAP1

YAP1, a transcriptional co-activator lacking the DNA binding domain and found to interact with different other transcription factors including TEADs, RUNX2, SMADs and multiple proteins involved in remodeling of chromatin.

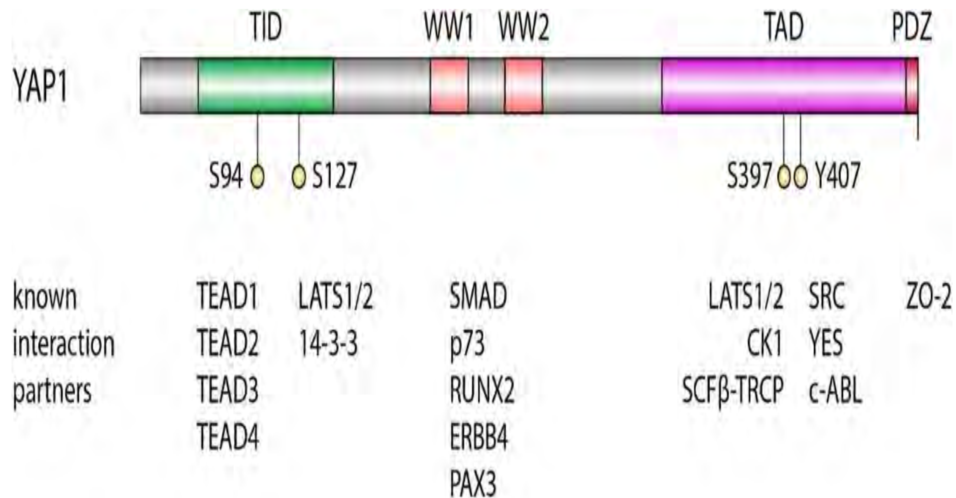


Figure 1.5: *Schematic representation of YAP1 protein* (Zhao *et al*., 2010).

Yap1 contains TEAD interaction domain (TID) in the N-terminal region where Serine 94 interact with E240 and Y406 of TEAD1 through hydrogen bonding and eventually facilitating the direct interaction of YAP1 and TEAD1. At Serine 127, there is a 14-3-3 interaction domain, targeted by kinase LATS1/2 and play an important role in cytoplasmic retention of YAP1 through interacting with 14-3-3 proteins upon phosphorylation by LATS1/2 (Zhao *et al*., 2010).

WW domains, a key distinctive feature among isoforms of YAP in terms of number. YAP1 contains two WW domains, that identifies a PPxY motif in respective protein targets during interaction with different proteins and transcription factors including RUNX2, p63/p73, SMADs, ERBB4 and PAX3 (Manderfield *et al*., 2014).

YAP1 contains a transactivation domain (TAD) in C-terminal region. Research studies have demonstrated that c-ABL, SRC, or YES factor phosphorylates the YAP1 at Y407 and can affect the functionality and interaction of YAP1 with certain proteins including β-catenin or p73 (Levy *et al*., 2008; Rosenbluh *et al*., 2012). At the very far end of c-terminus, a PDZ binding motif is present that facilitates the interaction YAP1 with ZO-2 and found to be critical for YAP1 activity and its localization in nucleus (Oka and Sudol 2009; Oka *et al*., 2010).

1.10.2 Role of YAP1 in tissue development and cancer

Hippo pathway inhibits the YAP1 accumulation in nuclei through its phosphorylation by key kinases including LATs1/2 and MST1/2 during normal development of tissues. However, upon tissue injury, YAP1 gets activated and promote self-renewal

DRSML QAU

of stem cell and triggers new cell generation to repair damaged tissue. It further interacts with different signaling pathways such as Wnt, BMP, EGF, TGF beta and HGF to promote differentiation and proliferation of cells (Wang *et al*., 2017).

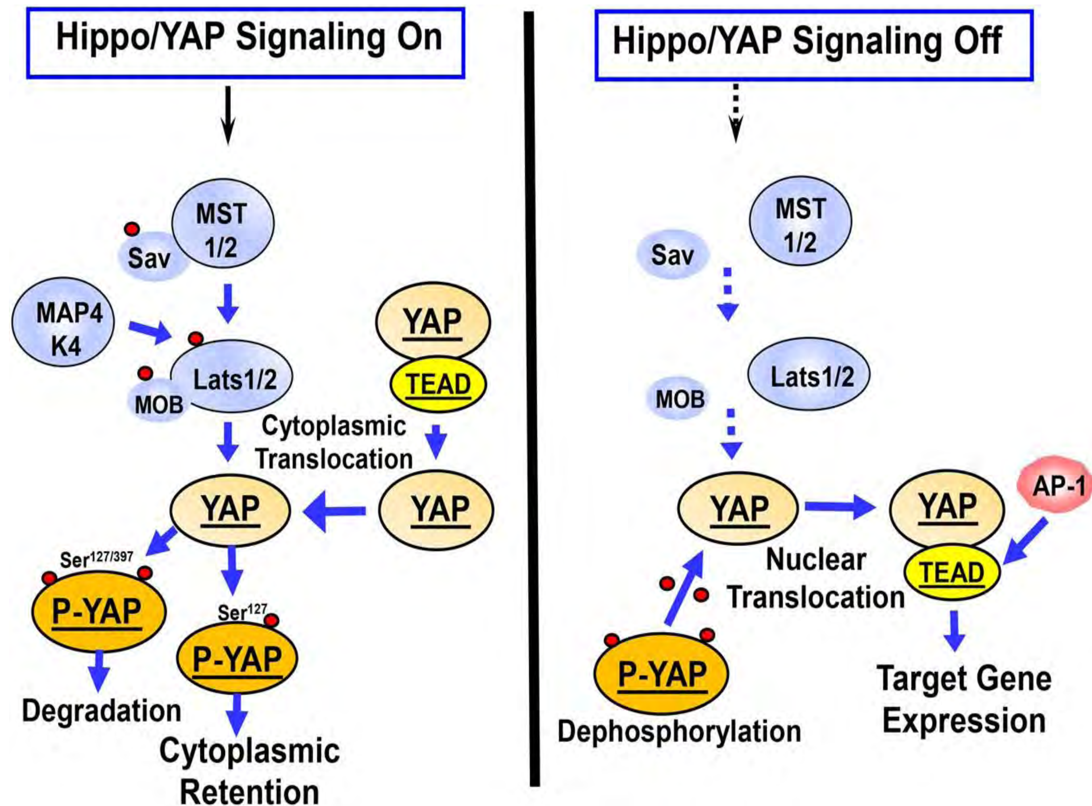


Figure 1.6: *Role of YAP1 in Hippo signaling pathway* (Rozengurt *et al*., 2018)

DRSM

When the Hippo signaling is activated in response to different signals including cell density, polarity, or mechanical stimuli, the Mst1/2 kinases phosphorylate and activate Lats1/2 in association with its regulatory protein MOB1/2. MAP4Ks function as alternative kinases that phosphorylate Lats1/2 in addition to Mst1/2. YAP is phosphorylated at Ser-127 and Ser-397 that promotes its cytoplasmic retention and degradation, respectively. YAP is dephosphorylated and moved into the nucleus upon inactivation of Hippo pathway, where it binds to and activates the TEAD transcription factors and promotes the activation of many genes involved in regulation of cell proliferation and differentiation (Rozengurt *et al*., 2018).

Initial studies revealed the role of YAP1 as tumor suppressor as it facilitates the pro-apoptotic function of P53 genes such as P63 and P73 through binding via its WW domain (Strano *et al*., 2005). Furthermore, its lower expression in different tumor types have also been corelated with decreased patient survival (Cottini *et al*., 2014).

However, recent findings have demonstrated the function of YAP1 as pro-oncogene. It interacts with transcription factors of TEAD family and eventually elicit the pro-proliferative and pro-survival transcriptional program. Mutational analysis and nuclear staining of various cancer samples have shown loss of function mutations in upstream factors of Hippo pathway such as LATs1/2 and MST1/2 that eventually causes higher nuclear expression of YAP1 which cause abnormality in cell proliferation and differentiation leading to increase tumorigenesis (Deel *et al*., 2015).

1.10.3 Role of YAP1 in PDAC

YAP1 has been revealed as an independent prognostic marker associated with relapse and poor survival in pancreatic cancer. Inhibition of interaction between YAP1 and TEAD have been found to significantly affect the development and preservation of a fibrotic microenvironment of tumor. Therefore, it is regarded as a biomarker produced from pancreatic cancer tissue that is therapeutically and physiologically meaningful.

Pancreatic ductal adenocarcinoma (PDAC) is divided into numerous molecular subgroups according to transcriptomic profiling, each of which has unique histological and clinical characteristics (Moffitt *et al*., 2015). The squamous subtype is augmented in YAP1 activation and linked to a poor prognosis. Analysis of expression profile using TCGA datasets, revealed higher levels of genes those under regulation of YAP1 in squamous subtype of PDAC. These findings were further supported by tissue microarray analysis and decreased tumorigenicity upon YAP1 depletion presented the significance of YAP1 activation in lower survival rate of patients suffering from squamous subtype of PDAC (Tu *et al*., 2019).

ZEB1 act as a co-activator for certain YAP target genes such as Connective tissue growth factor (CTGFR) that is an extracellular matrix (ECM) protein and mediates the oncogenic properties of YAP and increasing the metastasis, therapeutic resistance, and poor survival among cancer patients (Wang *et al*., 2009).

DRSML QAU

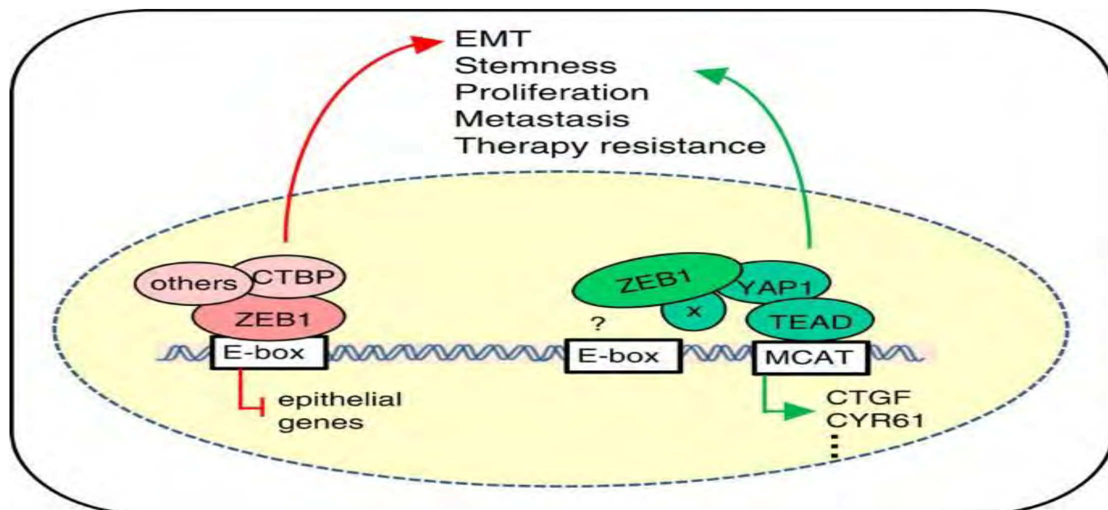


Figure 1.7: *Schematic representation ZEB1 and YAP1 interaction* (Lehman *et al*., 2016)

DRSML QAU

Lehman *et al*., 2016 reported the correlation between YAP1 and ZEB1 gene in breast cancer cell lines as in presence of both ZEB1 and YAP1 in wild type (WT) form has resulted the normal expression of CTGFR gene. On other hand, reduced expression of CTGFR was observed in case of Doxycyclin induced depletion of ZEB1 along with WT YAP1 and knockdown of YAP1 and WT ZEB1, indicating the interdependence of both genes for regulation of certain regulatory genes.

1.11 Driver Genes Mutations in PDAC

In recent decades, the genetic analysis using different approaches of sequencing such as Sanger sequencing and Next generation sequencing (NGS) have revealed several mutations in variety of tumor suppresser genes and oncogenes (Vogelstein *et al*., 2013). Wood and Hruban, (2012) have reported total four major driver gene mutation in PDAC, that comprises one oncogene mutation in KRAS and three mutations in tumor suppressor genes including Tp53, CDKN2A and SMAD4.

- ***KRAS***

KRAS, is the most frequently mutated oncogene to be reported in > 90% of PDAC cases. Generally it encodes a small GTPase protein, involved in variety of regulatory pathways of metabolism, proliferation and differentiation (Karnoub and Weinberg, 2008). It responds to different growth factors and are activated in GTP bound form (induced by GEFs) while inactivated in GDP bound state (induced by GAP). Stephen *et al*., (2014) have reported that mutations in 12, 13 and 61 codons of *KRAS* has

caused insensitivity to GAP, thus maintaining the constitutive activation of KRAS in GTP bound form. Mutated KRAS binds to several downstream effectors that promotes over proliferation of cancer cells and facilitates the disease progression.

- ***TP53***

TP53 gene is an essential tumor suppressor gene that actively responds to cellular stress such as DNA damage by regulating cell cycle, initiating apoptosis and senescence, to stimulates DNA repair leading to inhibition of tumor growth (Aubrey *et al*., 2016). It is the second most highly mutated gene in PDAC and missense mutations are more common, inducing gain-of-function (GOF) that increases the tumorigenesis (Aubrey *et al*., 2016; Yamamoto and Iwakuma, 2018).

- ***CDKN2A***

CDKN2A is another highly mutated tumor suppressor gene in PDAC, that normally encodes several important regulators of cell-cycle including p16INK4a, p15INK4 and p14ARF (Sharpless and Sherr, 2015). Mostly it acquires loss of function resulting in disruption of cell cycle leading to abnormal cell proliferation and differentiation that improves tumorigenesis and decrease the survival rate. Furthermore, DNA methyltransferase have been reported to induce hypermethylation of the CDKN2A promoter that mediates the silencing of tumor suppressor genes, thus ultimately increasing the metastasis of different cancer types including PDAC (Cui *et al*., 2015; Tang *et al*., 2015).

DRSML QAU

- **SMAD4**

Somatic mutation in Smad4 are more specific and mostly reported in >50% of invasive pancreatic malignancies. It is a signal transducers of Smad family and an important mediator of TGF-β signaling pathways that involves in several physiological processes, including cell growth, differentiation, and proliferation. Research studies have revealed the role of SMAD4 in angiogenesis and suppression of immune system hence presenting its significance in tumor development (Ahmed *et al*., 2017).

1.12 Aim and Objectives

This study aims to identify the role of ZEB1 and YAP1 in PDAC with the following objectives:

1. Genetic characterization of PDAC through Whole Exome Sequencing in Pakistani population.
2. Analysis of disease-causing sequence variants using bioinformatics tools.

DRSML QAU

2. Materials and Methods

2.1 Ethical Approval

Ethical approval was obtained to conduct the current research project from Institutional Review Board (IRB) of Quaid-i-Azam University Islamabad, Pakistan. To get patients' consent, an approval form was designed as per the criteria set out by the research and the ethical committee. The consent form was duly filled out patient's attendant and maintained as a record. Study subjects comprises the PDAC patients, and their samples were collected from PIMS Islamabad, Quaid-i-Azam International (QIH) and Holy family hospital, Rawalpindi.

2.2 Tissue sample collection

The subjects of study comprised of fresh tissues of tumor and normal tissues (as control) soon after the surgical resection of PDAC patient. Samples were collected in cryo-tubes and stored in liquid nitrogen to prevent degradation of DNA and proteins. Later, sample tubes were carefully stored at -80 in laboratory.

2.3 DNA Extraction from tissue

- **Preparation of Solutions for Genomic DNA Extraction**

Following solutions described in given table were prepared prior to proceed DNA extraction from tissue samples.

Table 2.1: *List of Reagents for DNA extraction and their Chemical Composition*

Sr.NO	Solutions	Chemical Composition
1.	**70% Ethanol**	➢ Absolute Ethanol (70mL) ➢ Distilled water (30mL)

DRSML QAU

2.	**Lysis buffer**	➢ 100mM EDTA (2.92g, pH 8.0) ➢ 10mM Tris HCl (0.12g, pH 8.0) ➢ 50mM NaCl (0.29g) ➢ 0.5 % SDS (0.5g) ➢ Distilled water (100ml)
3.	**Proteinase K**	➢ Proteinase K (0.1g) ➢ 10mM Tris HCl (50µl) ➢ 20mM CaCl2 (2000µl) ➢ Glycerol (5ml) ➢ Nuclease free water (5ml)
4.	**Sodium Acetate (pH 5.0)**	➢ 3M Sodium acetate (12.31g) ➢ Distilled water (50ml)
5.	**Isopropanol**	➢ 2- Propanol
6.	**Tris EDTA buffer (pH 8.0)**	➢ 1mM EDTA (0.029g) ➢ 10mM Tri's hydroxyl (methylamino) methane (0.12g) ➢ Distilled water (50ml)

- **Procedure:**

DNA was extracted manually using standard Phenol chloroform (Organic) method (Köchl *et al*., 2005; Sambrook and Russell, 2006). Following are the steps involved in DNA extraction process:

Day 1

1. Addition of lysis buffer

i. Pancreatic tissue samples of about 20 mg were mixed with 500µl of lysis solution for homogenization and followed by incubation at room temperature for 20-30 minutes.

ii. To ensure efficient homogenization, samples were gently mixed.

iii. After that, centrifugation of samples were performed at 13000 rpm for 3 minutes to get separation of phases.

iv. Upon centrifugation, we got the pallet containing DNA and the supernatant. Pellet was further treated with lysis buffer multiple times to reduce risk of contamination as far as possible while the supernatant was discarded.

v. Next, we added lysis solution (400 µl) to pellet along with 20% SDS (Sodium Dodecyl sulphate) (13µl), and proteinase K (25 µl), followed by incubation of samples overnight at 37 ◦ C.

Day 2

DNA Clean Up (Phenol: Chloroform: Isoamyl alcohol (PCI) Treatment)

vi. The incubated samples were further treated with 500 µl of PC1 solution including, phenol, chloroform and isoamyl alcohol.

vii. Centrifugation was proceeded at 13000 rpm for 10 minutes to ensure gentle and through mixing.

viii. After that, DNA was further purified and separated by transferring aqueous phase to other Eppendorf tube.

ix. Chloroform (500 µl) and isoamyl alcohol (C: I, 24:1) was again added to the aqueous layer and centrifuged again at 13000 rpm for 10 minutes.

2. DNA Precipitation

x. Next, we shifted the aqueous layer into another tube, and further treated with sodium acetate (55 µl) and chilled isopropanol (500 µl).

xi. Samples were incubated at -20 ◦ C for 45 minutes.

xii. After that, samples were again centrifuged for 10 minutes at 13000 rpm.

DRSML QAU

3. DNA Washing

xiii. We got the supernatant that was discarded, and 70% ethanol (500 µl) was added to pellet, followed by centrifugation for 5 minutes at 7500 rpm to ensure maximum removal of all contaminations. Again, supernatant was discarded, and pellet was air dried.

4. DNA Storage

xiv. TE Buffer (Tris EDTA) was added to DNA pellet and kept the pellet resuspended in buffer and stored at 4 ◦ C.

2.4 DNA Quantification

Quantitative analysis of DNA was performed using Thermo scientific Multi Skan Go Instrument. Concentration of DNA was measured in ng/ul unit while ratio of 260/280 ratio represented the quality of DNA sample (Waye *et al*., 1989).

2.5 Agarose Gel Electrophoresis:

➢ **Preparation of Solutions for Agarose Gel Electrophoresis**

Following tables comprises of the solutions prepared before to proceed agarose gel electrophoresis using extracted DNA samples.

PSML QAU

Table 2.2: *List of Reagents for Gel Electrophoresis and their Chemical Composition*

Sr.NO	Chemical	Composition
1.	**1%Agarose gel**	➢ 1g agarose gel, 1X TBE buffer (100 ml)
2.	**10X TBE**	➢ Tris Base (0.89 M, 54g) ➢ Boric Acid (0.025M, 27g) ➢ EDTA (3.65g, pH 8.3) ➢ Distilled water (500ml)

3.	**1X TBE**	➢ 10X TBE (10ml) ➢ Distilled water (90ml)
4.	**Ethidium Bromide**	➢ EtBr dissolved (400mg) ➢ 2.5mM ethidium bromide ➢ Distilled water (20ml
5.	**Bromophenol blue**	➢ Bromophenol blue (0.25%) ➢ Sucrose (40%) ➢ Distilled water (100 ml

DRSML QAU

- **Procedure**

i. 1% agarose gel was prepared by dissolving 1 gram of agarose in 100 ml of 1X TAE buffer (Tris Acetic acid EDTA). Solution was heated in oven for a minute to get complete dissolution and clarity.

ii. Ethidium Bromide (7 µl) was added in solution of gel, followed by pouring into the gel casting tray with inserted combs at appropriate distance.

iii. Once gel got solidified, we carefully removed the combs and caster was transferred to gel tank already filled with 1X TAE buffer.

iv. Extracted DNA (6 µl) was mixed with 6X bromophenol blue dye (2 µl) and loaded carefully in wells.

v. The gel was run under 500 mA of current at 75 volts for 1 hour.

vi. UV Trans-Illuminator bio–Doc Analyzer was used for gel visualization.

2.6 Whole exome Sequencing

After DNA extraction and quantification, genomic DNA was subjected to whole exome sequencing (WES), a widely used method to explore variations that potentially causes different genetic disorders and other disease including cancer. In contrast to whole genome sequencing, WES is more economical as only exome sequences that codes for proteins are targeted. Furthermore, it offers several benefits in biological

research studies as it provides high-quality and precise data set that ensure convenient and faster analysis of data sets.

2.7 Bioinformatics analysis

Bioinformatical analysis was performed to evaluate the sequencing data using multiple tools including SIFT, PROVEAN, Clin Pred, etc.

Complete Workflow

Fastq format of sequence date was used for detection of variants and annotation by using GATK variant calling pipeline. Following flowchart represents the workflow of data analysis:

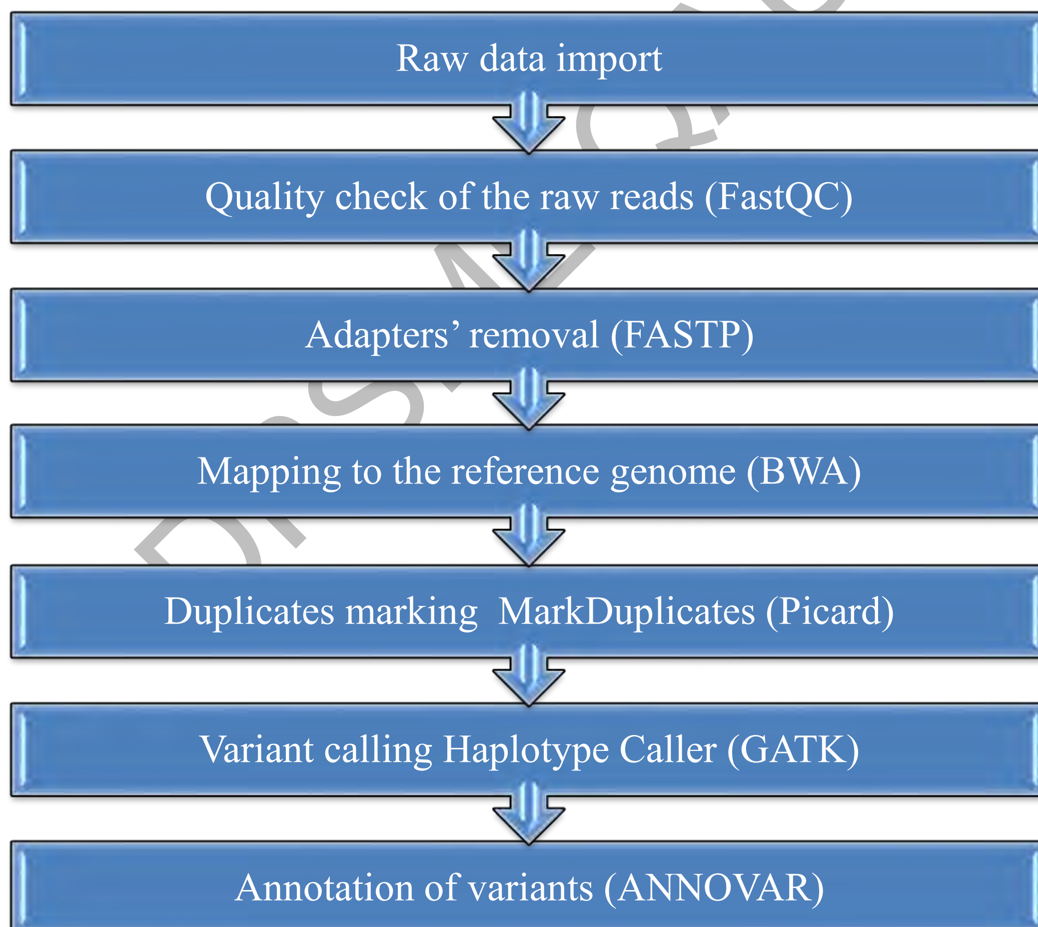


Figure 2.1: *Schematic Illustration of WES data analysis*

2.7.1 Raw data import

Linus system was utilized for further processing of raw data obtained in the form of FASTQ files.

2.7.2 Quality checking

FastQC tool was used to check quality control where fastq format was provided as input file. Html report was generated as output file that contains information about different quality control parameters (Andrew, 2010).

2.7.3 Adapters' removal

Certain Adapter sequences (over-represented or PCR primer sequences) were found in html file that was removed by using FASTP tool (Chen *et al*., 2018).

2.7.4 Mapping to the reference genome

Mapping to the indexed reference genome is an integral part of sequence analysis. BWA tool was used to perform mapping with fasta format of reference genome. As an input, Fastq files of the sequenced samples were provided, and an aligned BAM format was obtained as an output file (Li and Durbin, 2009).

The genome that was utilized as a reference is shown in the table below.

Species	Reference Genome	
Human	**Hg38**	Homo_sapiens.GRCh38.dna.primary_assembly. fa (https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.39/)

Table 2.3: *Reference genomes used for alignment Species*

SML QAU

2.7.5 Duplicates marking

The aligned BAM files were given as input to Mark Duplicates (Picard) tool to mark the duplicates sequences. Picard tool generates a new BAM file as an output that comprises the identified duplicates in the SAM flags field for each sequence read.

2.7.6 Variant calling by MuTect2

MuTect2 software was used for the variant calling, as it possess the highest rate of validation of about 90% in detection of mutation. After processing of key steps and alignments of sequence reads with a reference genome, data of tumor sample with matched normal control sample was provided to MuTect2 software. (Cibulskis *et al*., 2013)

Based on the normal samples, a Panel of Normals (PoN) was generated that is widely used resource for analysis of somatic variant. PoN varies as per type of variant under consideration however, it has some common features including:

- These are generated using normal samples that represents the healthy tissue which is assumed to lack any somatic variations.
- The major aim for generating PoN is to increase the reliability of outcomes of variation analysis by identifying recurring technical artifacts (Benjamin *et al*., 2019).

The variant calling file (VCF) was obtained as final output for each tumor sample.

DRSML QAU

2.7.7 Annotation of variants

Variant's annotation was performed using an ANNOVAR tool that provided with variants identifies in VCF file as an input. ANNOVAR is a reliable, easy to use and an efficient annotation tool that is widely used for identification of functional consequences of genetic variation found in data of high-throughput sequencing. It generated 3 output files

- A **vcf file** containing ANNOVAR annotations in the INFO field.
- A tab-delimited text file **(.txt)** to make manual Excel or other software inspection easier.
- An additional column-filled, **avinput file** that contains comments on the variations.

2.8 SIFT, PROVEAN and Polyphen2

- **Sorting Tolerant from Intolerant' (SIFT)** was used to assess the effect of amino acid substitution (AAS) on function of protein. It is an online available

bioinformatic based tool that assumes the conservation of protein sequence throughout evolution and hence expect the change in functionality of protein due to AAS (Kumar *et al*., 2009).

- ***Pro*tein *V*ariation *E*ffect *An*alyzer (PROVEAN)** is another efficient tool that was used for analysis of variants identified through WES (Choi *et al*., 2015). Unlike SIFT, it is also capable to determine the functional impact of amino acid insertions, and deletions in addition to AAS by using approach of alignment-based scores.(Choi *et al.*, 2012).
- **PolyPhen-2 (Polymorphism Phenotyping v2)** was also used to support the findings of SIFT and PROVEAN. It determines the effect of AAS on protein function and stability and widely used in WES analysis due to its capability of analyzing large sets of data (Adzhubei *et al*., 2013).

2.9 Mutation Tester and Mutation Assessor

Mutation Taster was used to determine the effect of changes in sequence of nucleotides through identification of functional impact of AAS, variation in intronic and intron-exon region that may influence. It was opted due to its highest accuracy of about 88% over other similar tools (Schwarz *et al*., 2014). Furthermore, Mutation assessor was used to predict either damaging or tolerant nature of missense variant (Gnad *et al*., 2013).

DRSML QAU

2.10 Damage Pred Count

Damage Pred Count/ Clin Pred was used to identify any nonsynonymous variation. Variation possessing values lower than 0.5 represents tolerant while values higher than 0.5 shows damaging nature of variation (Alirezaie *et al*., 2018).

2.11 CADD Raw and CADD Phred

Combined Annotation-Dependent Depletion (CADD) was used to determine the deleteriousness of variant. Higher values of Raw CADD scores represent that variant is deleterious based on comparison of tumor to normal control sequences. Furthermore, PHRED CADD was used to evaluate causal or fine-mapping interpretation of variant present in specific locus. It well predicts the pathogenic potential of Single Nucleotide Variations (SNVs) matched with reference genome Hg38 (Rentzsch, *et al*., 2019)

3. Results

3.1 Demographic Detail

Demographic details of study subjects are described in the table below:

Table 3.1: *Demographic details of study subjects*

S.no	Sex	Age (yr)	Wt (kg)	Histopatho -logical difference	Symptoms	Co. morbids	Hereditary diseases	Treatment
S1	M	56	54	Poor	Abdominal pain, Jaundice, and Vomiting	none	none	Whipple (folforinox)
S2	F	50	52	Poor	Weight loss, and Abdominal pain	none	none	Colostomy, and Whipple
S3	F	60	48	Poor	Weight loss, Abdominal pain, and Obstructive jaundice	Diabetes and Jaundice	Diabetes	Stenting, and Whipple

3.1.1 Stages of PDAC Patients

Following graph represents the stages of PDAC of study sample, 2 out of 3 patients were at stage II while one patient possesses stage III of PDAC.

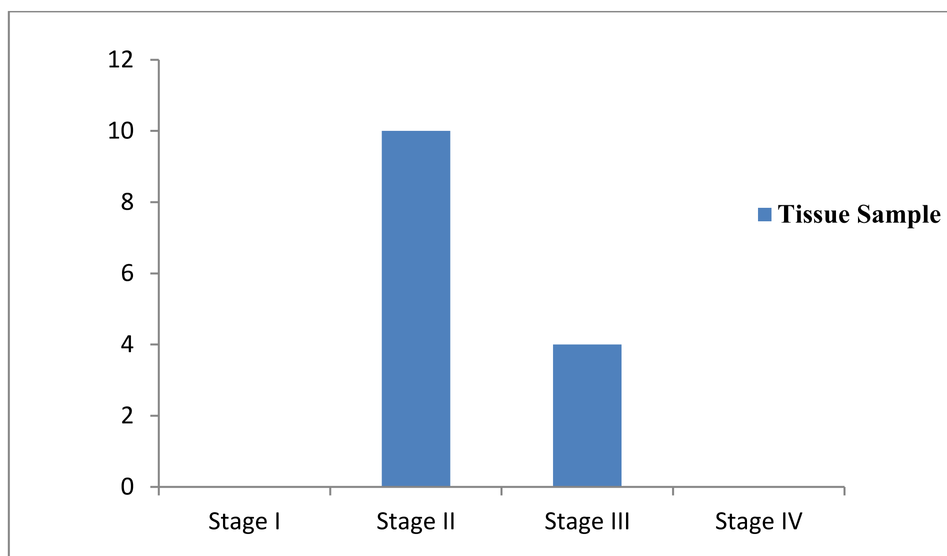


Figure 3.1: *Cancer Stages Frequency in PDAC Patients.*

3.2 Quantitative Analysis

Ratio of 260/280 represented the fair quality of extracted DNA and concentration is measured in ng/ul, described in given table:

Table 3.2: *Quality and Concertation of extracted DNA*

(Tumor sample=3 and control=3, n=12 as each sample was run in duplicate)

Sr. No	Nucleic Acid 260/280	Nucleic Acid Conc. in (ng/ul)
1	1.89	770
2	1.83	780
3	1.79	720
4	1.81	773
5	1.88	780
6	1.82	719

DR ML QAU

7	1.85	831
8	1.83	777
9	1.84	704
10	1.88	810
11	1.80	790
12	1.87	650

3.3 Qualitative Analysis

On 1% agarose gel, along with ladder and control DNA, 12 bands appeared because all 6 samples (3T, 3C) were run in duplicates. 1KB ladder was loaded in 1st wells with DNA samples in adjacent wells. High-quality bands can be seen on the gel because the extracted DNA from tissue samples was highly intact and had a size of more than 20kb. DNA bands are represented in following image of gel with comparison to 1KB Ladder.

SMLQAU

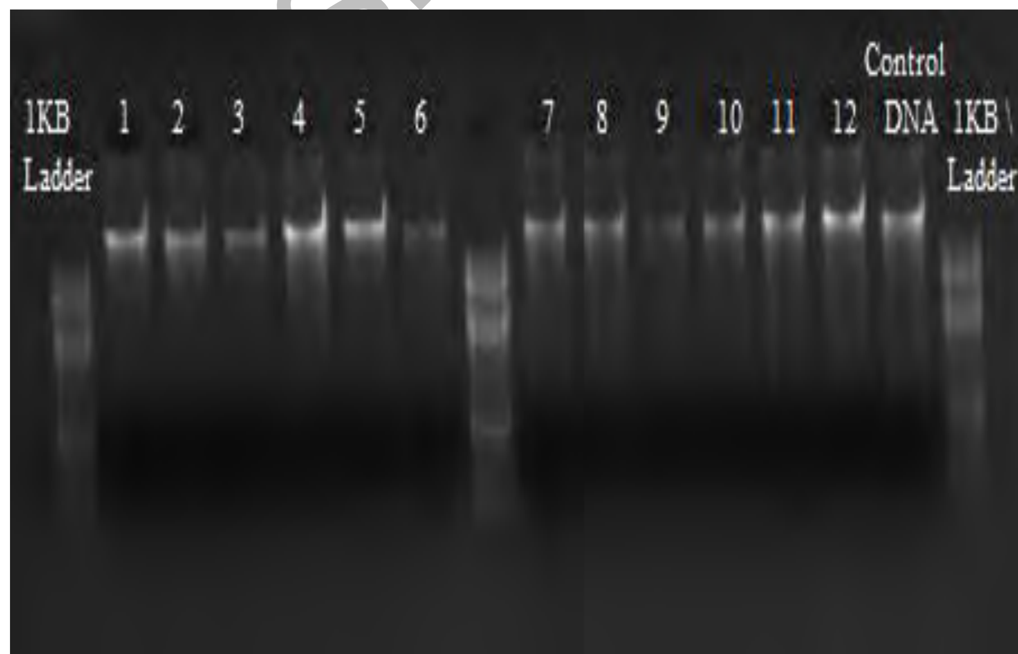


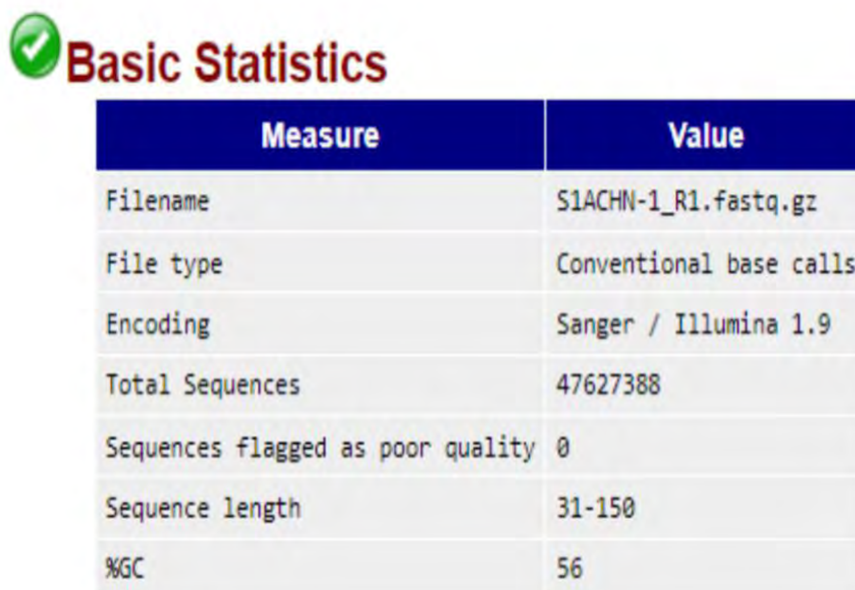
Figure 3.2: *Gel electrophoresis image*

3.4 WES Results analysis

Analysis of WES results was performed using variety of bioinformatics tools.

3.4.1 Quality check analysis of sample Raw reads

The sample of FastQC results' report is shown below:



Basic Statistics

Measure	Value
Filename	S1ACHN-1_R1.fastq.gz
File type	Conventional base calls
Encoding	Sanger / Illumina 1.9
Total Sequences	47627388
Sequences flagged as poor quality	0
Sequence length	31-150
%GC	56

Figure 3.3: *Basic statistics of Raw Reads using FASTQ*

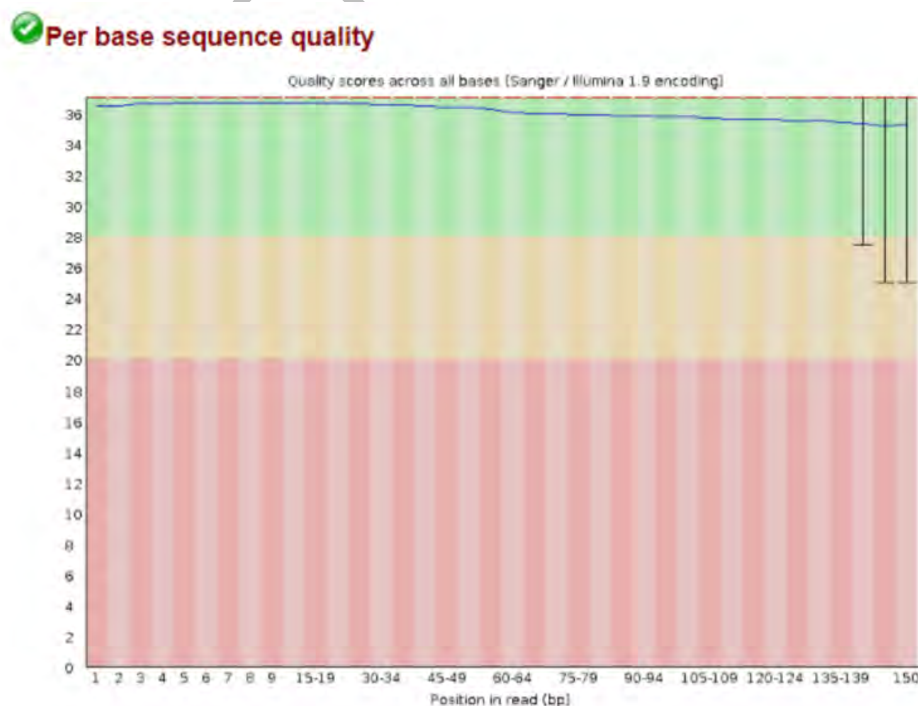


Figure 3.4: *Quality score graph of FastQ Tool*

3.4.2 Adapter content Analysis

The straight line along at x-axis represents the absence of adapter content.

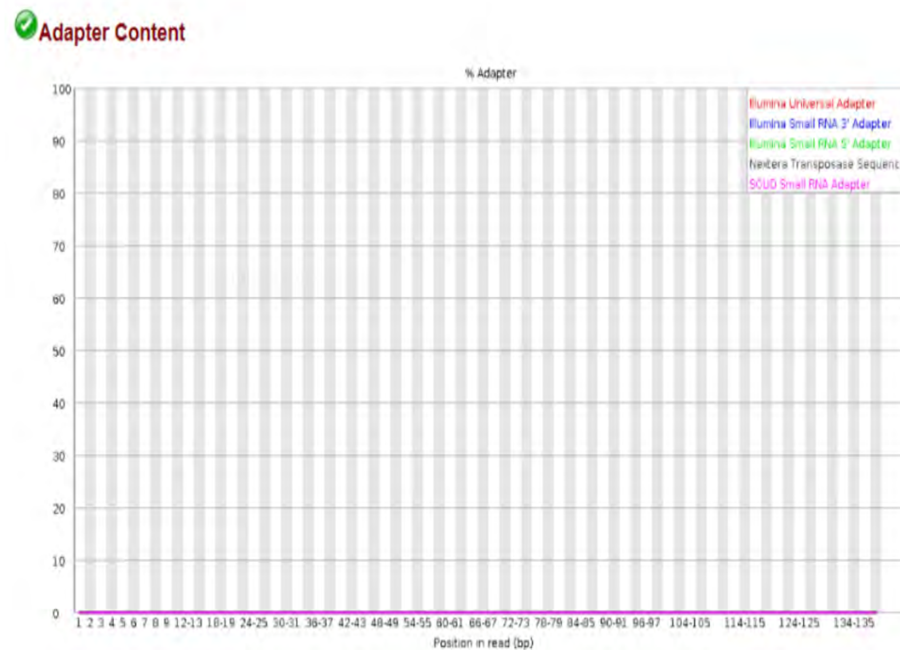


Figure 3.5: *Adapter Content of sequence*

DRSML

3.4.3 Variants' detection of PDAC

Mutect2 (GATK) identified variants in all samples. The number of variants identified in each sample are shown in the following table:

Table 3.3: *Number of variants in study samples*

Species	Pancreatic Samples IDs	Number of Variants
Homo sepians	S7APC	1590779
	S8APC	356593
	S1APC	633838

3.5 Analysis of Annotated variants

The variant calling file generated through ANNOVAR tool for study samples (n=3) matched with controlled sample (n=3) have described multiple variations in exonic region of highly mutated genes such as KRAS and SMAD4 in PDAC. Analysis of annotated variants presented a novel substitutive variation (31309911-12; GG>TA) in nucleotide sequence of ZEB1-AS1, coding for Long non-coding RNA. On other hand, analysis of annotated variants have not presented any variation in exonic region of YAP1 however, both ZEB1 and YAP1 genes have shown multiple variations in intronic, intergenic, upstream, downstream, 5'UTR and 3'UTR region.

3.5.1 Nucleotide variation in ZEB1

Following graph represents the frequency of SNVs in intronic, intergenic, upstream, Downstream, 3'UTR and 5'UTR of ZEB1. These SNVs were frequently observed in intronic and intergenic regions.

QAU

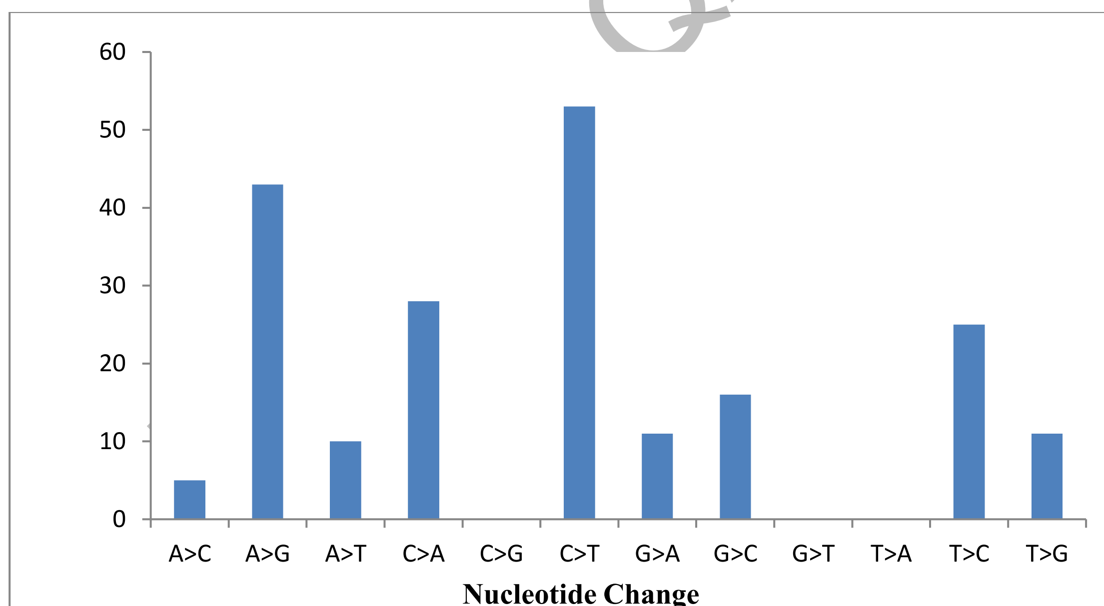


Figure 3.6: *Single nucleotide variation in ZEB1 gene*

3.5.2 Nucleotide variation in YAP1

Following graph represents the frequency of SNVs in intronic, intergenic, upstream, Downstream, 3'UTR and 5'UTR of YAP1. These SNVs were frequently observed in intronic and intergenic regions.

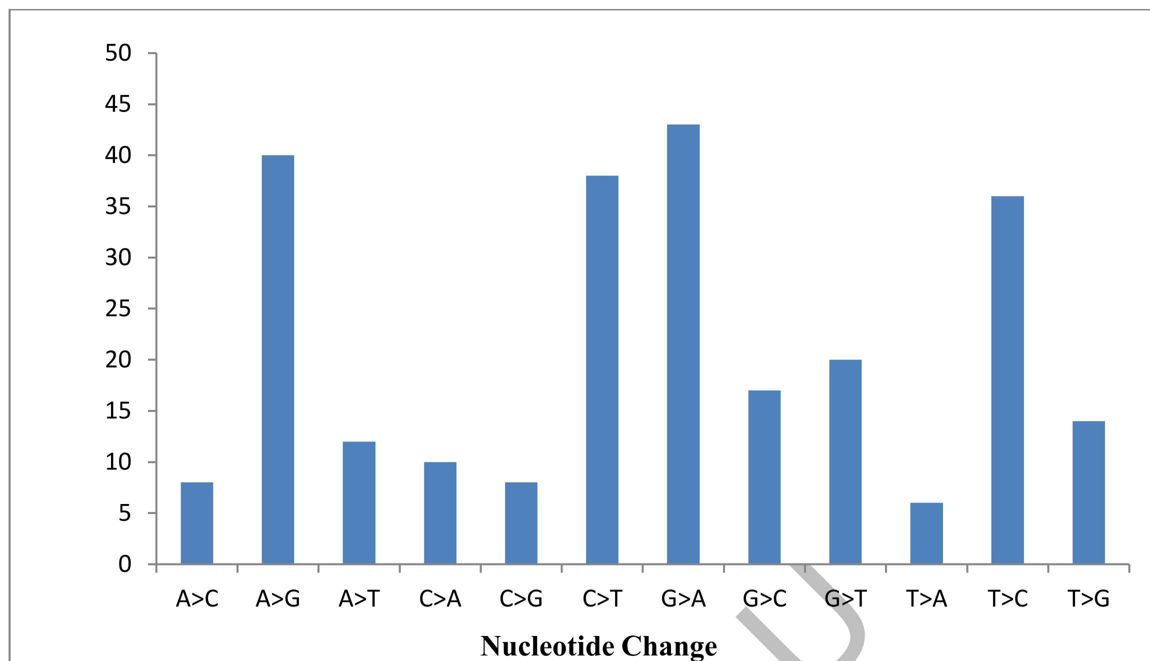


Figure 3.7: *Single nucleotide variation in YAP1 gene.*

3.5.3 Nucleotide Variation in driver genes

No variation have been observed in exonic region of CDKN2A while 1 SNVs have been found in KRAS in 5th Exon (51058160; G>A) **(Figure: 3.8)** and in p53 at 2nd Exon (7674945; G>A) **(Figure: 3.9)**. Furthermore, a novel SNV have been found in exonic region of SMAD4 in 6th Exon (51058160; G>A) **(Figure: 3.10).**

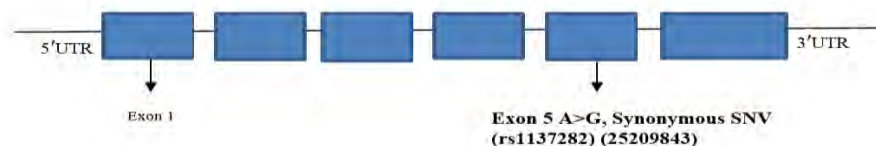


Figure 3.8: *SNV in Econ 5 of KRAS gene*

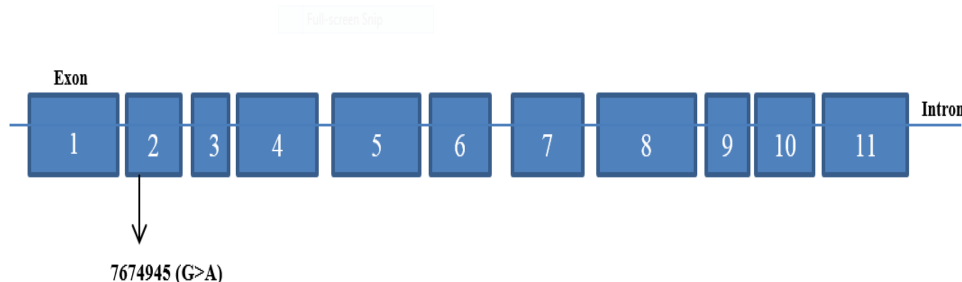


Figure: 3.9: *SNV in Exon 2 of p53* gene

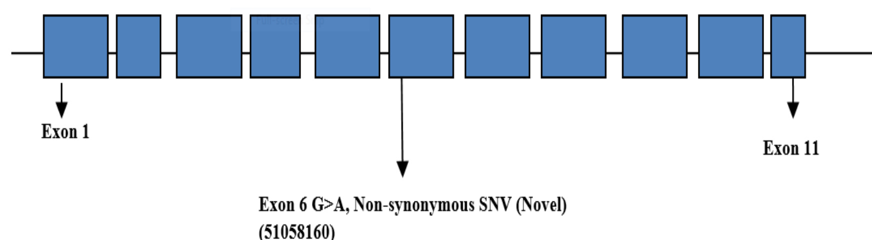


Figure 3.10: *SNV in Econ 6 of SMAD4 gene*

3.5.4 Amino acid substitution analysis

SIFT, PROVEAN and Polyphen2 results have not provided information about amino acid substitutions as these tools perform the variant analysis using different databases. The ‘’no result’ ’confirms that no data regarding the identified variant is available in literature and databases, thus the ZEB1-AS1 variant is novel.

3.5.5 Damaging or tolerant nature of variants

No information about variant have been obtained by using Mutation tester and mutation Assessor as these tools also interprets the variations based on data available in databases, hence indicating the novelty of ZEB1-AS1 variant.

3.5.6 Scoring the damaging potential of variation

Clin pred could not provide any information about the damaging or tolerant nature of identified variant depicting the lack of information in databases, thus further confirming the novelty of identified variant.

3.5.7 Deleteriousness of variants

CADD raw and CADD Phred have not provided any score to discovered variant. As it uses algorithm by comparing the variation to well reported deleterious variation for assigning scores as per degree of deleteriousness of variation. This further confirms the novel nature of discovered variant.

DRSML GAU

4. Discussion

Pancreatic cancer, one of deadliest type of cancer affecting millions of individuals across the globe with minimal survival rate that accounts for 1-5 years. This is credited to asymptomatic nature of pancreatic cancer or sometime its symptoms are confused with gastric issues causing the delayed diagnosis and eventually lowers the chances of cure. Situation is worst in low-income countries like Pakistan, as annual screening for major infections and disorders are not in regular practice. There is high need to provide awareness about lethality of pancreatic cancer among masses to ensure the timely diagnosis and overall survival rate.

PDAC is the most common malignancy (>90%) among pancreatic cancer types. Genetic aberration in both tumor suppressor genes including SMAD4, Tp53, CDKN2A and oncogenes such as KRAS, have been reported to significantly increase the disease progression and poor prognosis. Research studies have reported the overexpression of ZEB1 and YAP1 in PDAC cases (Lambert *et al*., 2017; Tu *et al*., 2019). ZEB1 is an important inducer of EMT process as it facilitates cytoskeletal remodeling by downregulating factors of cell polarity and upregulating various MMPs, eventually increasing the invasive potential of tumor cell (Brabletz S and Brabletz T, 2010; Tu *et al*., 2019). Yap1 is a key player of Hippo signaling pathway, acting as transcriptional co-activator of multiple genes involved in cell proliferation, organ development, tissue homeostasis and chromatin remodeling in response to cellular stress. YAP1 induced inhibition of apoptotic genes results in abnormal cell growth leading towards tumor development (Wang *et al*., 2017; Rozengurt *et al*., 2018).

DRSML QAU

This study was aimed to perform genetic analysis of PDAC patients using WES, a widely used method for genetic and molecular characterization of various hereditary disorders and cancer. Annotation of variants, identified in study sample was performed using an ANNOVAR tool that tabulated the data in form of VCF file. As concerning the driver genes, analysis of annotated variants have presented no variation in exonic region of CDKN2A while 1 SNVs have been found in KRAS in 5th Exon (51058160; G>A) and in p53 at 2nd Exon (7674945; G>A). Furthermore, a novel SNV have been found in exonic region of SMAD4 in 6th Exon (51058160; G>A).

We have found a novel substitutive variation (31309911-12; GG>TA) in nucleotide sequence of ZEB1-AS1 through analysis of annotated variants. Bioinformatics analysis using variety of tools such as SIFT, PROVEAN, Polyphen2**,** Mutation Tester and assessor have not provided any information about effect of this variant on sequence of amino acids, as these tools uses different databases for variant detection. The ‘’no result’ ’confirms that no data regarding the identified variant is available in literature and databases, thus confirming the novelty of identified variant.

Furthermore, Clin Pred, CADD Raw and CADD Phred tools have not provided any scores to represent the damaging potential and deleteriousness of (31309911-12; GG>TA) variant, further validating the novelty of variation. Xiong et al., (2018) reported the regulatory role of ZEB1-AS1 on the expression of ZEB1 gene. It is the lncRNA originated from ZEB1 promoter region and hinders the miR-101 induced inhibition of Zeb1 expression in CRC, thus mediating its upregulation to eventually promotes EMT and increased progression towards advance stage of cancer. Wallner *et al*., (2009) have also reported the ZEB1 induced blockage of certain stemness-repressing miRNAs including miR-200, miR-183 and miR-203, to promote stemness in cancer cells and eventually increasing the tumorigenicity.

On other hand, analysis of annotated variants have not presented any variation in exonic region of YAP1. However, both ZEB1 and YAP1 genes have shown multiple variations in intronic, intergenic, upstream, downstream, 5’UTR and 3’UTR region that might be responsible for their higher expression of both genes in PDAC. These findings demands further validation through *in vitro* and *in vivo* studies using animal models and cell lines, transcriptome analysis, molecular docking, and in silico evaluation. These would collectively help to further investigate the root cause of disease and diagnostic marker that would eventually help in early diagnosis and devising potential therapeutics to improves the survival rate among PDAC patients.

DRSML QAU

5. Conclusion

At the end, we conclude that a novel substitutive variation (31309911-12; GG>TA) in nucleotide sequence of ZEB1-AS1, might be causing upregulation of ZEB1 to facilitate the process of EMT and eventually the progression of disease. In addition, no variation have been found in exonic region of YAP1 gene hence its higher expression could not be credited to nucleotide change in exome region. However, multiple variations have been observed in intronic and intergenic regions of both genes that have link with its upregulation in PDAC and poor survival rate. Further studies are still needed to validate the research findings and to explore therapeutic target to eventually devise cure for this deadliest cancer type.

DRSML QAU

References:

Adachi, K. (2011). Primary squamous cell carcinoma of the pancreas: a case report. *JOP. Journal of the Pancreas*, *12*(2), 181-184.

Adzhubei, I., Jordan, D. M., &Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current protocols in human genetics*, *76*(1), 7-20.

Ahmed, S., Bradshaw, A. D., Gera, S., Dewan, M. Z., & Xu, R. (2017). The TGF-β/Smad4 signaling pathway in pancreatic carcinogenesis and its clinical significance. *Journal of clinical medicine*, *6*(1), 5.

Alirezaie, N., Kernohan, K. D., Hartley, T., Majewski, J., & Hocking, T. D. (2018). ClinPred: prediction tool to identify disease-relevant nonsynonymous single-nucleotide variants. *The American Journal of Human Genetics*, *103*(4), 474-483.

Al-Shehri, A., Silverman, S., & King, K. M. (2008). Squamous cell carcinoma of the pancreas. *Current oncology*, *15*(6), 293-297.

Andrews, S., Krueger, F., Segonds-Pichon, A., Biggins, L., Krueger, C., &Wingett, S. (2010). FastQC. *A quality control tool for high throughput sequence data*, *370*.

Aubrey, B. J., Strasser, A., & Kelly, G. L. (2016). Tumor-suppressor functions of the TP53 pathway. *Cold Spring Harbor perspectives in medicine*, *6*(5), a026062.

Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., ... &Grimmond, S. M. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. Nature, 531(7592), 47-52.

Benjamin, D., Sato, T., Cibulskis, K., Getz, G., Stewart, C., & Lichtenstein, L. (2019). Calling somatic SNVs and indels with Mutect2. *BioRxiv*, 861054.

Bixler, H. A., Castro, M. J., & Stewart III, J. (2011). Cytologic differentiation of squamous elements in the pancreas. *Diagnostic cytopathology*, *39*(7), 536-540.

Brown, H. A., Dotto, J., Robert, M., & Salem, R. R. (2005). Squamous cell carcinoma of the pancreas. *Journal of clinical gastroenterology*, *39*(10), 915-919.

DRSML QAU

Brabletz, S., & Brabletz, T. (2010). The ZEB/miR-200 feedback loop—a motor of cellular plasticity in development and cancer? *EMBO reports*, *11*(9), 670-677.

Caramel, J., Ligier, M., & Puisieux, A. (2018). Pleiotropic roles for ZEB1 in cancer. *Cancer research*, *78*(1), 30-35.

Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics*, *34*(17), i884-i890.

Chaffer, C. L., Marjanovic, N. D., Lee, T., Bell, G., Kleer, C. G., Reinhardt, F., ... & Weinberg, R. A. (2013). Poised chromatin at the ZEB1 promoter enables breast cancer cell plasticity and enhances tumorigenicity. *cell*, *154*(1), 61-74.

Choi, Y., Sims, G. E., Murphy, S., Miller, J. R., & Chan, A. P. (2012). Predicting the functional effect of amino acid substitutions and indels.

Choi, Y., & Chan, A. P. (2015). PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745-2747

Cibulskis, K., Lawrence, M. S., Carter, S. L., Sivachenko, A., Jaffe, D., Sougnez, C., ... & Getz, G. (2013). Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nature biotechnology*, *31*(3), 213-219.

Collisson, E. A., Bailey, P., Chang, D. K., &Biankin, A. V. (2019). Molecular subtypes of pancreatic cancer. *Nature reviews Gastroenterology & hepatology*, *16*(4), 207-220.

Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., ... & Gray, J. W. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nature medicine*, *17*(4), 500-503.

Cottini, F., Hideshima, T., Xu, C., Sattler, M., Dori, M., Agnelli, L., ... &Tonon, G. (2014). Rescue of Hippo coactivator YAP1 triggers DNA damage–induced apoptosis in hematological cancers. *Nature medicine*, *20*(6), 599-606.

Croome, K. P., Farnell, M. B., Que, F. G., Reid-Lombardo, K., Truty, M. J., Nagorney, D. M., & Kendrick, M. L. (2014). Total laparoscopic pancreaticoduodenectomy for pancreatic ductal adenocarcinoma: oncologic advantages over open approaches?. *Annals of surgery*, *260*(4), 633-640.

DRSML QAU

Cui, C., Gan, Y., Gu, L., Wilson, J., Liu, Z., Zhang, B., & Deng, D. (2015). P16-specific DNA methylation by engineered zinc finger methyltransferase inactivates gene transcription and promotes cancer metastasis. *Genome biology*, *16*(1), 1-12.

Deel, M. D., Li, J. J., Crose, L. E., &Linardic, C. M. (2015). A review: molecular aberrations within hippo signaling in bone and soft-tissue sarcomas. *Frontiers in oncology*, *5*, 190.

Elsherif, S. B., Virarkar, M., Javadi, S., Ibarra-Rovira, J. J., Tamm, E. P., & Bhosale, P. R. (2020). Pancreatitis and PDAC: association and differentiation. *Abdominal Radiology*, *45*(5), 1324-1337.

Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., & Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International journal of cancer*, *149*(4), 778-789.

Gnad, F., Baucom, A., Mukhyala, K., Manning, G., & Zhang, Z. (2013). Assessment of computational methods for predicting the effects of missense mutations in human cancers. *BMC genomics*, *14*(3), 1-13.

Huang, J., Lok, V., Ngai, C. H., Zhang, L., Yuan, J., Lao, X. Q., ... & Wong, M. C. (2021). Worldwide burden of, risk factors for, and trends in pancreatic cancer. *Gastroenterology*, *160*(3), 744-754.

Hubers, J., Patel, R., Dalvie, P., &Megna, B. (2018). Percutaneous Transhepatic CholangioscopyWith Electrohydraulic Lithotripsy in a Patient With Choledocholithiasis in a Metal Biliary Stent. *The American Journal of Gastroenterology*, *113*, S832-S833.

Hruban, R. H., Canto, M. I., Goggins, M., Schulick, R., & Klein, A. P. (2010). Update on familial pancreatic cancer. *Advances in surgery*, *44*(1), 293-311.

Jones, S., Hruban, R. H., Kamiyama, M., Borges, M., Zhang, X., Parsons, D. W., ... & Klein, A. P. (2009). Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science*, *324*(5924), 217-217.

Kamisawa, T., Wood, L. D., Itoi, T., &Takaori, K. (2016). Pancreatic cancer. *The Lancet*, *388*(10039), 73-85.

DRSML QAU

Karnoub, A. E., & Weinberg, R. A. (2008). Ras oncogenes: split personalities. *Nature reviews Molecular cell biology*, *9*(7), 517-531.

Kastrinos, F., Mukherjee, B., Tayob, N., Wang, F., Sparr, J., Raymond, V. M., ... &Syngal, S. (2009). Risk of pancreatic cancer in families with Lynch syndrome. *Jama*, *302*(16), 1790-1795.

Kendrick, M. L., &Cusati, D. (2010). Total laparoscopic pancreaticoduodenectomy: feasibility and outcome in an early experience. *Archives of surgery*, *145*(1), 19-23.

Köchl, S., Niederstätter, H., & Parson, W. (2005). DNA extraction and quantitation of forensic samples using the phenol-chloroform method and real-time PCR. In Forensic DNA typing protocols (pp. 13-29). Humana Press.

Kong, S. Y., Kim, K. S., Kim, J., Kim, M. K., Lee, K. H., Lee, J. Y., ... & Park, K. S. (2016). The ELK3-GATA3 axis orchestrates invasion and metastasis of breast cancer cells in vitro and in vivo. *Oncotarget*, *7*(40), 65137.

Krebs, A. M., Mitschke, J., Losada, M. L., Schmalhofer, O., Boerries, M., Busch, H., ... & Brabletz, T. (2017). The EMT-activator Zeb1 is a key factor for cell plasticity and promotes metastasis in pancreatic cancer. *Nature cell biology*, *19*(5), 518-529.

Kryklyva, V., Haj Mohammad, N., Morsink, F. H., Ligtenberg, M. J., Offerhaus, G. J. A., Nagtegaal, I. D., ... &Brosens, L. A. (2019). Pancreatic acinar cell carcinoma is associated with BRCA2 germline mutations: a case report and literature review. *Cancer Biology & Therapy*, *20*(7), 949-955.

Kumar, P., Henikoff, S., & Ng, P. C. (2009). Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nature protocols*, *4*(7), 1073-1081.

Lambert, A. W., Pattabiraman, D. R., & Weinberg, R. A. (2017). Emerging biological principles of metastasis. *Cell*, *168*(4), 670-691.

Lehmann, W., Mossmann, D., Kleemann, J., Mock, K., Meisinger, C., Brummer, T., ... & Brabletz, T. (2016). ZEB1 turns into a transcriptional activator by interacting with YAP1 in aggressive cancer types. *Nature communications*, *7*(1), 1-15.

DRSML QAU

Levy, D., Adamovich, Y., Reuven, N., &Shaul, Y. (2008). Yap1 phosphorylation by c-Abl is a critical step in selective activation of proapoptotic genes in response to DNA damage. *Molecular cell*, *29*(3), 350-361.

Li, R., Pei, H., & Watson, D. K. (2000). Regulation of Ets function by protein–protein interactions. *Oncogene*, *19*(55), 6514-6523.

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *bioinformatics*, *25*(14), 1754-1760.

Li, J., Li, Y., Cao, G., Guo, K., Zhang, L., & Ma, Q. (2013). Early manifestations of pancreatic cancer: The effect of cancer–nerve interaction. *Medical hypotheses*, *81*(2), 180-182.

Liao, W., Clift, A. K., Patone, M., Coupland, C., González-Izquierdo, A., Pereira, S. P., & Hippisley-Cox, J. (2021). Identifying symptoms associated with diagnosis of pancreatic exocrine and neuroendocrine neoplasms: a nested case-control study of the UK primary care population. *British Journal of General Practice*, *71*(712), e836-e845.

Ma, S. J., Oladeru, O. T., Miccio, J. A., Iovoli, A. J., Hermann, G. M., & Singh, A. K. (2019). Association of timing of adjuvant therapy with survival in patients with resected stage I to II pancreatic cancer. *JAMA network open*, *2*(8), e199126-e199126.

Manderfield, L. J., Engleka, K. A., Aghajanian, H., Gupta, M., Yang, S., Li, L., ... & Epstein, J. A. (2014). Pax3 and hippo signaling coordinate melanocyte gene expression in neural crest. *Cell reports*, *9*(5), 1885-1895.

Mizrahi, J. D., Surana, R., Valle, J. W., & Shroff, R. T. (2020). Pancreatic cancer. *The Lancet*, *395*(10242), 2008-2020.

Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S., Hoadley, K. A., ... & Yeh, J. J. (2015). Virtual microdissection identifies distinct tumor-and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nature genetics*, *47*(10), 1168-1178.

Neoptolemos, J. P., Dunn, J. A., Stocken, D. D., Almond, J., Link, K., Beger, H., ... & members of the European Study Group for Pancreatic Cancer. (2001).

DRSML CAU

Adjuvant chemoradiotherapy and chemotherapy in resectable pancreatic cancer: a randomised controlled trial. *The Lancet*, *358*(9293), 1576-1585.

Oka, T., &Sudol, M. (2009). Nuclear localization and pro-apoptotic signaling of YAP2 require intact PDZ-binding motif. *Genes to Cells*, *14*(5), 607-615.

Oka, T., Remue, E., Meerschaert, K., Vanloo, B., Boucherie, C., Gfeller, D., ... &Sudol, M. (2010). Functional complexes between YAP2 and ZO-2 are PDZ domain-dependent, and regulate YAP2 nuclear localization and signalling. *Biochemical Journal*, *432*(3), 461-478.

Pereira, S. P., Oldfield, L., Ney, A., Hart, P. A., Keane, M. G., Pandol, S. J., ... & Costello, E. (2020). Early detection of pancreatic cancer. *The lancet Gastroenterology & hepatology*, *5*(7), 698-710.

Pelzer, U., Hilbig, A., Sinn, M., Stieler, J., Bahra, M., Dörken, B., &Riess, H. (2013). Value of carbohydrate antigen 19-9 in predicting response and therapy control in patients with metastatic pancreatic cancer undergoing first-line therapy. *Frontiers in oncology*, *3*, 155.

Purnak, T., El Hajj, I. I., Sherman, S., Fogel, E. L., McHenry, L., Lehman, G., ... & Easler, J. J. (2021). Combined versus separate sessions of endoscopic ultrasound and endoscopic retrograde cholangiopancreatography for the diagnosis and management of pancreatic ductal adenocarcinoma with biliary obstruction. *Digestive Diseases and Sciences*, *66*(8), 2786-2794.

Rebours, V., Gaujoux, S., d'Assignies, G., Sauvanet, A., Ruszniewski, P., Lévy, P., ... &Couvelard, A. (2015). Obesity and Fatty Pancreatic Infiltration Are Risk Factors for Pancreatic Precancerous Lesions (PanIN) Fatty Pancreatic Infiltration and PanIN Lesions. *Clinical Cancer Research*, *21*(15), 3522-3528.

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic acids research*, *47*(D1), D886-D894.

Rodriguez-Aznar, E., Wiesmüller, L., Sainz Jr, B., & Hermann, P. C. (2019). EMT and stemness—Key players in pancreatic cancer stem cells. *Cancers*, *11*(8), 1136.

DRSML QAU

Rosenbluh, J., Nijhawan, D., Cox, A. G., Li, X., Neal, J. T., Schafer, E. J., ... & Hahn, W. C. (2012). β-Catenin-driven cancers require a YAP1 transcriptional complex for survival and tumorigenesis. *Cell*, *151*(7), 1457-1473.

Rozengurt, E., Sinnett-Smith, J., &Eibl, G. (2018). Yes-associated protein (YAP) in pancreatic cancer: at the epicenter of a targetable signaling network associated with patient survival. *Signal transduction and targeted therapy*, *3*(1), 1-10.

Sambrook, J., & Russell, D. W. (2006). Purification of nucleic acids by extraction with phenol: chloroform. Cold Spring Harbor Protocols, 2006(1), pdb-prot4455.

Serafini, S., Da Dalt, G., Pozza, G., Blandamura, S., Valmasoni, M., Merigliano, S., & Sperti, C. (2017). Collision of ductal adenocarcinoma and neuroendocrine tumor of the pancreas: a case report and review of the literature. *World Journal of Surgical Oncology*, *15*(1), 1-5.

Schmidt-Hansen, M., Berendse, S., & Hamilton, W. (2016). Symptoms of pancreatic cancer in primary care: a systematic review. *Pancreas*, *45*(6), 814-818.

Schwarz, J. M., Cooper, D. N., Schuelke, M., &Seelow, D. (2014). MutationTaster2: mutation prediction for the deep-sequencing age. *Nature methods*, *11*(4), 361-362.

Sharpless, N. E., &Sherr, C. J. (2015). Forging a signature of in vivo senescence. *Nature Reviews Cancer*, *15*(7), 397-408.

Stephen, A. G., Esposito, D., Bagni, R. K., & McCormick, F. (2014). Dragging ras back in the ring. *Cancer cell*, *25*(3), 272-281.

Strano, S., Monti, O., Pediconi, N., Baccarini, A., Fontemaggi, G., Lapi, E., ... &Blandino, G. (2005). The transcriptional coactivator Yes-associated protein drives p73 gene-target specificity in response to DNA Damage. *Molecular cell*, *18*(4), 447-459.

Soreide, K. (2018). Sweet predictions speak volumes for early detection of pancreatic cancer. *Gastroenterology*, *155*(2), 265-268.

Stolzenberg-Solomon, R. Z., Schairer, C., Moore, S., Hollenbeck, A., & Silverman, D. T. (2013). Lifetime adiposity and risk of pancreatic cancer in the NIH-

DRSML QAU

AARP Diet and Health Study cohort. *The American journal of clinical nutrition*, *98*(4), 1057-1065.

Sultana, A., Tudur Smith, C., Cunningham, D., Starling, N., Tait, D., Neoptolemos, J. P., &Ghaneh, P. (2007). Systematic review, including meta-analyses, on the management of locally advanced pancreatic cancer using radiation/combined modality therapy. *British journal of cancer*, *96*(8), 1183-1190.

Tang, B., Li, Y., Qi, G., Yuan, S., Wang, Z., Yu, S., ... & He, S. (2015). Clinicopathological significance of CDKN2A promoter hypermethylation frequency with pancreatic cancer. *Scientific Reports*, *5*(1), 1-12.

Tu, B., Yao, J., Ferri-Borgogno, S., Zhao, J., Chen, S., Wang, Q., ... & Ying, H. (2019). YAP1 oncogene is a context-specific driver for pancreatic ductal adenocarcinoma. *JCI insight*, *4*(21).

Valls, C., Andía, E., Sanchez, A., Fabregat, J., Pozuelo, O., Quintero, J. C., ... &Jorba, R. (2002). Dual-phase helical CT of pancreatic adenocarcinoma. *Am J Roentgenol*, *178*(4), 821-826.

Viale, P. H. (2020). The American Cancer Society's facts & figures: 2020 edition. *Journal of the Advanced Practitioner in Oncology*, *11*(2), 135.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz Jr, L. A., &Kinzler, K. W. (2013). Cancer genome landscapes. *science*, *339*(6127), 1546-1558.

Wang, M. Y., Chen, P. S., Prakash, E., Hsu, H. C., Huang, H. Y., Lin, M. T., ... & Kuo, M. L. (2009). Connective tissue growth factor confers drug resistance in breast cancer through concomitant up-regulation of Bcl-xL and cIAP1. *Cancer research*, *69*(8), 3482-3491.

Waye, J. S., Presley, L. A., Budowle, B., Shutler, G. G., &Fourney, R. M. (1989). A simple and sensitive method for quantifying human genomic DNA in forensic specimen extracts. Biotechniques, 7(8), 852.

Wellner, U., Schubert, J., Burk, U. C., Schmalhofer, O., Zhu, F., Sonntag, A., ... & Brabletz, T. (2009). The EMT-activator ZEB1 promotes tumorigenicity by repressing stemness-inhibiting microRNAs. *Nature cell biology*, *11*(12), 1487-1495.

DRSML QAU

Wilson, M. M., Weinberg, R. A., Lees, J. A., &Guen, V. J. (2020). Emerging mechanisms by which EMT programs control stemness. *Trends in cancer*, *6*(9), 775-780.

Wood, L. D., &Hruban, R. H. (2012). Pathology and molecular genetics of pancreatic neoplasms. *Cancer journal (Sudbury, Mass.)*, *18*(6), 492.

Xiong, W. C., Han, N., Wu, N., Zhao, K. L., Han, C., Wang, H. X., ... & He, P. (2018). Interplay between long noncoding RNA ZEB1-AS1 and miR-101/ZEB1 axis regulates proliferation and migration of colorectal cancer cells. *American journal of translational research*, *10*(2), 605.

Yang, H., Schramek, D., Adam, R. C., Keyes, B. E., Wang, P., Zheng, D., & Fuchs, E. (2015). ETS family transcriptional regulators drive chromatin dynamics and malignancy in squamous cell carcinomas. *Elife*, *4*, e10870.

Yamamoto, S., & Iwakuma, T. (2018). Regulators of oncogenic mutant TP53 gain of function. *Cancers*, *11*(1), 4.

Yuan, C., Morales-Oyarvide, V., Babic, A., Clish, C. B., Kraft, P., Bao, Y., ... &Wolpin, B. M. (2017). Cigarette smoking and pancreatic cancer survival. *Journal of Clinical Oncology*, *35*(16), 1822.

Zhao, B., Li, L., Tumaneng, K., Wang, C. Y., & Guan, K. L. (2010). A coordinated phosphorylation by Lats and CK1 regulates YAP stability through SCFβ-TRCP. *Genes & development*, *24*(1), 72-85.

Zhao, Q., Ren, Y., Xie, H., Yu, L., Lu, J., Jiang, W., ... & Li, B. (2021). ELK3 Mediated by ZEB1 Facilitates the Growth and Metastasis of Pancreatic Carcinoma by Activating the Wnt/β-Catenin Pathway. *Frontiers in Cell and Developmental Biology*, 2150.

Zheng, X., Carstens, J. L., Kim, J., Scheible, M., Kaye, J., Sugimoto, H., ... & Kalluri, R. (2015). Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature*, *527*(7579), 525-530.

DRSML QAU