DISSERTATION

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Mohanachary Amaravadi

born in Warangal, India

Oral-examination: July 29th, 2015

Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

M.Sc. Mohanachary Amaravadi

born in: Warangal, India

Oral-examination: July 29th, 2015

This Page is

Intentionally Left Blank

Identification of Pathogenic Virus Sequences in Pancreatic Ductal Adenocarcinoma

Referees: Prof. Dr. Roland Eils

PD Dr. Ralf Bischoff

This Page is

Intentionally Left Blank

DECLARATION

With this, I declare that the thesis entitled “Identification of pathogenic virus sequences in pancreatic ductal adenocarcinoma” has been written only by the undersigned. I have not used other sources or materials than those expressly indicated in this thesis. Verbatim text passages taken from other works in the spirit of science are well identified wherever used, stating the sources. The work documented in this thesis was carried out in the division of Functional Genome Analysis at the Deutsches Krebsforschungszentrum (DKFZ) in Heidelberg during the period of October 2011 to May 2015. I also declare that I have neither applied for an examination at any other institution nor have I used the thesis, in this or any other form, in any other institution for examination purposes. Further, this work is not being submitted as a dissertation for evaluation to any other faculty at the Ruprecht-Karls-Universität Heidelberg.

Heidelberg, 30.07.2015

Place and Date

Amaravadi, Mohanachary

Digitally signed by Amaravadi, Mohanachary
DN: cn=Amaravadi, Mohanachary, o=German Cancer Research Center, ou=Functional Genome Analysis, email=m.amaravadi@dkfz.de, c=DE
Date: 2015.07.30 10:58:12 +02'00'

Signature

ACKNOWLEDGEMENTS

It was my joy to spend time with wonderful colleagues at the German Cancer Research Center (DKFZ), Heidelberg. With immense encouragement from my family, friends, colleagues and mentors, I grew and matured as a responsible individual at DKFZ both personally and professionally. I would like to express my warm regards to all those who supported me to complete this study.

I would like to express my deepest gratitude to my boss and advisor, Dr. Jörg Hoheisel, for his excellent guidance, caring, patience, and providing me with a fantastic atmosphere for doing independent research. I would like to thank Prof. Roland Eils and Prof. Paul Schnitzler for guiding my thesis research for the past several years and helping me to develop as a better scientist. I would also like to thank Dr. Karin Müller-Decker, who was willing, to collaborate for in vivo *mouse experiments at a very short notice and during a dire hour of need. In addition, I would like to thank Dr. Ralf Bischoff, who was instrumental in extending the funding for my PhD and also for examining this thesis. Special thanks go to Dr. Agnes Hotz-Wagenblatt, who let me experience the next generation sequence analysis of tumor samples. Her assistance as my bioinformatics partner to this project was immensely resourceful. My dearest colleague and co-supervisor, Dr. Andrea Bauer, has helped me to design the project with insightful suggestions; my warmest regards to you Andrea.*

I would like to thank Dr. Sandeep Kumar Botla, Dr. Panduranga Sivarama Krishna Rachakonda, Pouria Jandaghi, Mehdi Manoochehri, Dr. Smiths Sengkwawoh Lueong, Laureen Sander, and Shakhawan Mustafa, who as good friends were always willing to help a great length even sharing my scientific responsibilities. I would also like to show profound appreciation for the staff members, Marie-Christine Leroy-Schell, and Anke Mahler, for their crucial role in meticulous processing of all the purchases made for this project.

I would also like to thank, my parents for always supporting and encouraging me, and my dearest younger sister for annoying, and so training me to deal with scientific "problems" during my PhD. I could not thank more my best pals, Yogesh Vegunta, Balakrishna Reddy Gadwala and Shalini Menon for being loyal friends at times of need by providing affable moral support.

Finally, I am very thankful to the German Academic Exchange Service (DAAD), for providing extremely generous financial support through their prestigious doctoral fellowship.

INSPIRATION

"Fill the brain with high thoughts, highest ideals, place them day and night before you, and out of that will come great work. Each work has to pass through these stages — ridicule, opposition, and then acceptance. Those who think ahead of their time are sure to be misunderstood."

— Swami Vivekananda

Volume II, Chapter 2, The Real Nature of Man (delivered in London)
Volume V, Epistles - First Series, XLVII, USA, 9th July 1895

"If medicine has been the great tutor of biology (William Harvey) then viruses have been the grand dons of oncology."

— Clodagh C O'Shea

William Scandling Developmental Chair
Molecular and Cell Biology Laboratory
The Salk Institute for Biological Studies
California, United States of America

DEDICATION

I would humbly like to dedicate this thesis to the family members of fallen souls, suffered because of an awful disease called cancer.

— Mohanachary Amaravadi

German Cancer Research Center (DKFZ)
Heidelberg, German

TABLE OF CONTENTS

SUMMARY

Pancreatic ductal adenocarcinoma (PDAC) is by far the most common type of pancreatic cancer. It constitutes about 90% of tumors of the exocrine pancreas. The aggressive nature of PDAC along with a lack of diagnostic markers contributes to high lethality of this disease, which is nearly identical to its incidence. Studies from malignancies such as hepatocellular carcinoma and cervical cancer, along with the fact that liver and pancreas are in a close proximity, provided a plausible basis for the hypothesis of virus association in PDAC tumor development. However, there have been no established reports about virus(es) associated with pancreatic cancer.

The present study identified a new cancer-associated virus in human PDAC samples, called Meleagrid herpesvirus-1 (MeHV-1), or also known commonly as herpesvirus of turkeys, by two different and independent approaches: experimental genomic subtraction and digital microRNAome subtraction between healthy and PDAC patients. In the first approach, a genome-wide experimental comparison of DNA from PDAC tissues to DNA from tissues of healthy individuals was performed by representational difference analysis (RDA). Using this technique, differences in sequence composition were selectively isolated and amplified with very high sensitivity. Virus sequences associated with the occurrence of pancreatic cancer were detected by this process. The second approach, performed in parallel, involved a sequence analysis of the complete microRNA (miRNA) content of PDAC tissue samples. The sequencing data was digitally compared to databases of human and viral sequences so as to identify viral miRNAs. Because of the limited number of molecules, this analysis form did not need any experimental selection and amplification in order to achieve a sufficiently enough sensitivity to find viral microRNAs. The common results of the two analyses strongly suggested that MeHV-1 plays a crucial role in PDAC tumor progression. One of the viral

microRNAs – hvt-miR-H14-3p – was studied in detail at the functional level by both *in vitro* and *in vivo* experiments in order to define the molecular mechanism of action with regard to its effect on pancreatic tumor carcinogenesis.

The key findings from this work include:

- Identification of MeHV-1 DNA sequences in the PCR difference products (DPs) resulting from RDA on genomic DNA from PDAC and healthy tissues.
- A tumor-specific MeHV-1 signature was also identified in the miRNA sequence analysis of tumor DNA, using an independent methodological approach.
- RT-qPCR analyses showed that hvt-miR-H14-3p from MeHV-1 was expressed at significantly higher levels in PDAC and chronic pancreatitis (CP) tissues – CP being a chronic inflammation of the pancreas and a well-known risk factor of PDAC – than in healthy tissues. This observation was further verified using independent digital PCR platforms.
- Metastatic and non-metastatic PDAC cell lines overexpressing hvt-miR-H14-3p showed a significant increase in migration and invasion compared to the respective controls, interestingly, without any significant change in proliferation.
- Hvt-miR-H14-3p was found targeting cellular p27, down-regulating its expression.
- The functional consequences of viral sequences identified *in vitro* could also be confirmed *in vivo* in NOD *scid* gamma mice.

In conclusion, this study is very significant in elucidating functional consequences of viral sequences in PDAC for the definition of relevant molecular effects responsible for carcinogenesis.

ZUSAMMENFASSUNG

Das duktale Adenokarzinom des Pankreas (PDAC) ist mit etwa 90% aller Fälle die häufigste Form von Krebs des exokrinen Pankreas. Aufgrund seiner sehr aggressiven Natur und dem Mangel an geeigneten Verfahren zur frühen Diagnose, ist die Todesrate bei PDAC nahezu identisch mit der Zahl der Erkrankungen. Studien an anderen Tumoren, wie etwa dem Leberzell- und Zervixkarzinom, und auch die Tatsache dass Leber und Pankreas dicht beieinander liegen, bildeten eine vage aber plausible Basis für die Hypothese, dass auch die Entwicklung des PDAC mit der Aktivität von Viren verknüpft sein könnte. Bisher gab es aber keine fundierten Berichte über eine direkte Assoziation eines Virus mit PDAC.

In der hier beschriebenen Studie wurde mittels zweier unabhängiger Verfahren in Tumoren menschlicher Patienten ein Virus identifiziert, das mit dem Auftreten von PDAC assoziiert ist. Es handelt sich um das Meleagrid Herpesvirus-1 (MeHV-1), ein Herpesvirus des Truthahns. Beim ersten Verfahren wurde ein genomweiter Vergleich der DNA aus Krebsgeweben mit der DNA aus gesunden Geweben mittels Repräsentativer Differenz Analyse (RDA) durchgeführt. RDA erlaubt es, durch eine experimentelle Substraktion mit hoher Sensitivität Unterschiede in Genomen zu finden und selektiv zu isolieren. Durch diesen Prozess konnten virale Sequenzen identifiziert werden, die mit dem Auftreten des Tumors korrelieren. Im zweiten Ansatz wurde durch Sequenzierung analysiert, welche microRNA Moleküle in Patientengeweben aber nicht in gesundem Pankreas auftauchen. Da nur ein sehr kleiner Teil des Genoms analysiert wurde, war keine experimentelle Selektion notwendig; stattdessen fand diese durch einen *in silico* Vergleich der Sequenzergebnisse miteinander und mit Datenbanken menschlicher und viraler Sequenzen statt. Die unabhängig gewonnenen Ergebnisse beider Ansätze deuteten darauf hin, dass MeHV-1 eine wichtige Rolle in der PDAC Karzinogenese spielt.

Eine der viralen microRNAs – hvt-miR-H14-3p – wurde anschließend detailliert auf funktioneller Ebene durch *in vitro* wie auch *in vivo* Studien untersucht, um den molekularen Wirkmechanismus zu identifizieren, durch den die Karzinogenese des Pankreastumors beeinflusst wird.

Die wichtigsten Ergebnisse der Arbeit umfassen:

- Die Identifizierung von MeHV-1 DNA Sequenzen in den Differenz-Produkten (DPs) des RDA Vergleichs der genomischen DNA aus PDAC und gesunden Geweben.
- Die Identifizierung einer tumorspezifischen MeHV-1 Signatur auch auf Ebene der microRNA durch ein zweites, unabhängiges Analyseverfahren.
- Analysen mittels RT-qPCR zeigten, dass hvt-miR-H14-3p in Geweben von Patienten mit PDAC oder Chronischer Pankreatitis (CP) – einer chronischen Entzündung des Pankreas, die ein bekannter Risikofaktor für die Entwicklung von PDAC ist – signifikant stärker exprimiert wurde als in Geweben von gesunden Personen. Dies wurde durch unabhängig durchgeführtes digitales PCR bestätigt.
- Metastatische und nicht-metastatische PDAC Zelllinien, in denen hvt-miR-H14-3p überexprimiert wurde, zeigten einen signifikanten Anstieg der Zell-Migration und Invasion, während interessanterweise kein Unterschied in der Proliferation vorlag.
- Hvt-miR-H14-3p reduziert die Expression des zellulären Proteins p27.
- Die Wirkung der viralen microRNA konnte *in vivo* in NOD *scid* gamma Mäusen bestätigt werden.

Zusammengefasst dokumentiert diese Arbeit einen Zusammenhang zwischen PDAC und dem Auftreten viraler Sequenzen und zeigt einen molekularen Mechanismus, über den die Karzinogenes beeinflusst werden kann.

1. INTRODUCTION

1.1. Cancer

Trillions of living cells make up the human body. In a healthy human being, cells grow and divide to produce new cells, and die in a very controlled and programmed manner. During the early development of the human body, normal cells divide relatively faster to allow growth of the individual. After adolescence and throughout adulthood, most cells divide only to replace deteriorated or dying cells or to repair tissue damage. Cancer originates when the orderly processes that control the growth and multiplication of normal cells become amiss. Cancer is a collection of related diseases characterized by the uncontrolled growth of cells and spread of the abnormal cells to different tissues of the body, frequently resulting in death. Cancer is a global problem. There are various causes of cancer: some of them are external factors such as tobacco use, infectious organisms, chemicals, radiation, and there are also internal factors like inherited mutations, hormone aberration and immune disruption that could potentially lead to cancer. These factors may act in tandem or independently to trigger or promote cancer development. These changes predominantly result in alterations or mutations in the genetic material of cells, thus affecting their behavior. The specific factor and the order and speed at which multiple factors accumulate, together with the genetic makeup of the individual influence the rate at which cancer develops and progresses.

During the past few years, the identity of many cancer pathways was derived from whole-genome sequencing of tumor types, and also from the exhaustive information of genomics, transcriptomics, proteomics, and metabolomics. Collaborations between various international cancer research centers generated data to provide particular insights into the nature of cancer cell and its development. Early stage detection of cancer can be improved using such data,

but a more refined molecular classification of tumor types is also possible. They also reveal aberrant signaling and alterations in the cancer cells which establish a basis for targeted therapy. Malignant cancers are also made up of various cell populations such as fibrous, inflammatory, vascular, and immunological cells. Any one or more of these cell types could be critical to tumor development and hence may offer an approach to prevention or therapy.

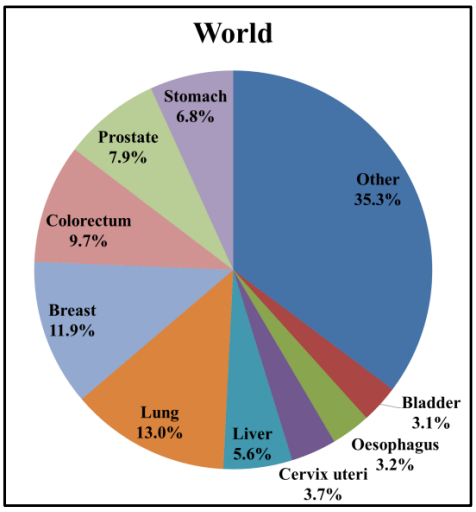


Figure 1: Estimated world cancer incidence proportions by major organ sites in both sexes (redrawn based on the original data from Stewart and Wild, World Cancer Report, 2014).

Malfunctions in a cell by genetic mutations often lead to tumor initiation, which when coupled with interactions between cancer cells and their surrounding environment, known as the tumor microenvironment, influence the development and progression of disease. Cancer is a major cause of morbidity and mortality, with annually more than 14 million new cases and 8 million cancer-related deaths, affecting populations in all countries. More than 60% of the world's cancer cases occur in Africa, Asia, and Central and South America, and these regions account for about 70% of cancer-related deaths (Stewart and Wild 2014). The diversity and distribution of cancer in the world indicates marked and extreme differences with respect to particular tumor types. The dedicated efforts of researchers working

throughout the biomedical research enterprise continue to expand our knowledge of cancer. Such data are vital to understanding disease causes and for the development of preventive measures.

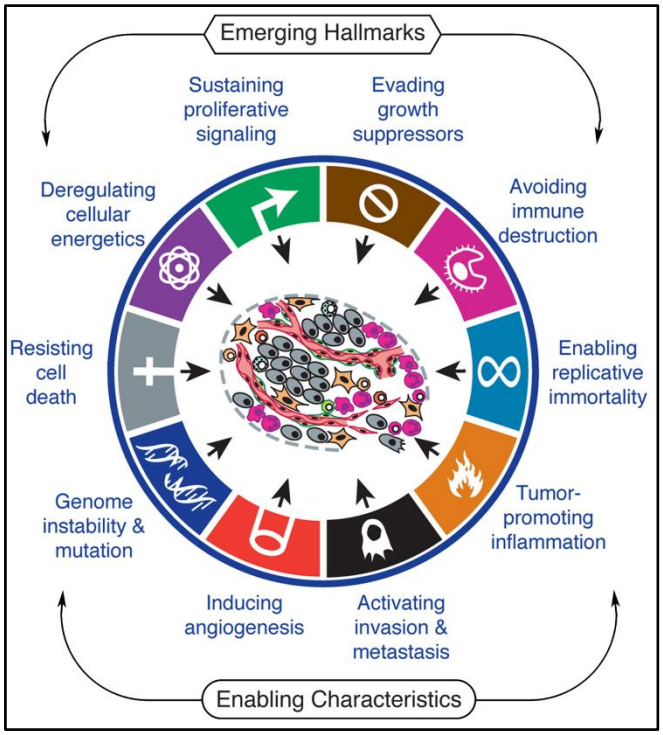


Figure 2: The emerging hallmarks and enabling characteristics of cancer (redrawn based on the original illustration by Hanahan and Weinberg, 2011).

The classic hallmarks of cancer proposed by Hanahan and Weinberg in 2000 (Hanahan and Weinberg 2000) comprised of six biological capabilities. During the tumor development; these hallmarks are acquired by various multistep processes. These hallmarks form a consolidating principle for streamlining the complexities of cancer. They include sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis. These prominent hallmarks are achieved by two enabling characteristics (Hanahan and Weinberg 2011).

The first enabling characteristic is the development of genomic instability in cancer cells. It generates random mutations and chromosomal rearrangements. These rare genetic changes can coordinate hallmark capabilities. A second enabling characteristic is the inflammation of premalignant and malignant tumor lesions that are driven by the immune system. Some of the immune cells also serve to promote tumor progression through various pathways.

Conceptual understanding of cancer in the recent years has added two emerging hallmarks to the existing list (Hanahan and Weinberg 2011). The capability to modify, or reprogram, cellular energetics in order to most effectively support neoplastic proliferation is functionally independent of the six core hallmarks of cancer. This reprogrammed energy metabolism is an emerging hallmark. Cancer cells also avoid immunological destruction particularly by T and B lymphocytes, macrophages, and natural killer cells, although the immune system offers a significant barrier to tumor formation and progression in humans. Nevertheless, highly immunogenic cancer cells may successfully evade immune destruction by turning off components of the immune system. This tumor immunoevasion is another emerging hallmark, whose designation as a core hallmark is yet to be established. Tumors also exhibit an additional facet of complexity on top of cancer cells. Tumors contain a range of recruited, presumably normal cells that contribute to the hallmark traits by creating the tumor microenvironment. Epigenetic aberrations can also be factors in both the cancer cells and the tumor-associated stroma. MicroRNAs have been shown to be involved in various tumor phenotypes (Garzon, Marcucci et al. 2010). However, the functions of the majority of microRNAs in our cells that dysregulate in expression during cancer are poorly understood.

1.2. Pancreatic Cancer

Our research group has strong interest in studying the cancer of pancreas. Pancreatic cancer is a devastating and fatal disease. Early detection tests for pancreatic cancer are unsuccessful and patients with a localized tumor will not have recognizable symptoms. This result in late diagnosis of the disease after the tumor metastasizes to other organs. Pancreatic ductal adenocarcinoma (PDAC) is the most lethal and common type of pancreatic cancer. It is usually diagnosed at an advanced stage and is resistant to therapy (Ryan, Hong et al. 2014). The mortality rates for patients with other gastrointestinal malignancies have decreased consistently during the past 30 years. However, there has been no significant change in survival rates of patients with PDAC. Currently, more than 80% of patients suffer relapse of the cancer after resection of the tumor tissue (Garrido-Laguna and Hidalgo 2015). The following sections detail the epidemiology, risk factors, biological and genetic characteristics, of pancreatic cancer with respect to PDAC phenotype.

1.2.1. PDAC is a Highly Aggressive Type of Pancreatic Cancer

As mentioned above, PDAC contributes the major proportion of pancreatic cancer cases. A characteristic feature of PDAC is an abundant accumulation of stromal cells. This suggests their possible role in the development and progression of pancreatic cancer. Pancreatic stellate cells (PSCs) are predominantly responsible for producing collagenous stroma. The stroma exhibits cellular elements like immune cells, endothelial cells and neural cells in addition to extracellular matrix proteins. There is strong evidence that indicates significant interactions between PSCs and tumor cells. *In vitro and in vivo* reports suggest that these interactions promote local tumor growth and distant metastasis of pancreatic cancer (Apte, Xu et al. 2015). Stromal cells affect the delivery of oxygen and nutrients to the tumor by reducing vascularization. Thus stroma causes the hypoxic zones that promote metabolic

adaptation in the tumor cells. The aggressive nature of PDAC is partly attributed to this metabolic reprogramming which is further influenced by environmental constraints (Olivares and Vasseur 2015).

1.2.2. Epidemiology and Risk Factors of PDAC

The early diagnosis of PDAC is very rare in patients younger than 40 years of age. The median age at diagnosis of PDAC is 71 years. The worldwide incidence of pancreatic cancer (comprising 85% PDAC) ranges from 1 to 10 cases per 100,000 people 1. It is the eighth or fourth leading cause of death from cancer in men and the ninth or fourth leading cause of death from cancer in women worldwide or in the Western world, respectively (Ryan, Hong et al. 2014). The risk factors and genetic syndromes associated with pancreatic cancer are shown in Table 1.

Variable	Approximate Risk
Risk factors	
Smoking	2–3
Long-standing diabetes mellitus	2
Nonhereditary and chronic pancreatitis	2–6
Obesity, inactivity, or both	2
Non–O blood group	1–2
Genetic syndromes and associated genes — %	
Hereditary pancreatitis (PRSS1, SPINK1)	50
Familial atypical multiple mole and melanoma syndrome (p16)	10–20
Hereditary breast and ovarian cancer syndromes (BRCA1, BRCA2, PALB2)	1–2
Peutz–Jeghers syndrome (STK11 [LKB1])	30–40
Hereditary nonpolyposis colon cancer (Lynch syndrome) (MLH1, MSH2, MSH6)	4
Ataxia–telangiectasia (ATM)	Not Available
Li–Fraumeni syndrome (P53)	Not Availabla

Table 1: Risk factors and inherited syndromes associated with pancreatic cancer (based on the original data set of Ryan et al., 2014). Values associated with risk factors are expressed as relative risks, and values associated with genetic syndromes are expressed as lifetime risks, as compared with the risk in the general population.

It is estimated that 5% to 10% of pancreatic cancers are inherited. Cigarette smoking accounts for around 20% of pancreatic cancers. Individuals who smoke have a twofold higher risk of developing pancreatic cancer than nonsmokers. The incidence of pancreatic cancer increases with a family history of this disease and chronic pancreatitis (CP), diabetes, obesity, and high alcohol consumption. Genetic syndromes such as Lynch syndrome also increase the risk of pancreatic cancer. Chronic infection with Hepatitis B virus, Hepatitis C virus, or *Helicobacter pylori* also make an individual susceptible for PDAC. Recent evidence also suggests that consumption of red or processed meat could slightly increase the risk.

1.2.3. Biological Characteristics of PDAC

Pathologic features of PDAC include a high rate of activating mutations in KRAS, precursor lesions progression, invasion and metastasis, an extensive stromal aggregation resulting in a reduced vascularization and hypoxic microenvironment, reprogramming of cellular metabolism, and evasion of tumor immunity (Feig, Gopinathan et al. 2012).

Extensive studies on molecular pathology and genome data have established a model of the PDAC progression. There are microscopic premalignant pancreatic lesions associated with the pancreatic ducts, which harbor similar mutations of PDAC. These lesions are called pancreatic intraepithelial neoplasia (PanIN). There is a stepwise progression of PanINs from low grade to high grade in types 1, 2, and 3, accumulating genetic alterations (Figure 3). Low grade PanINs, the type 1 lesions, are columnar cells and readily detectable in disease-free pancreas. High grade PanINs are papillary and lose polarity undergoing mitosis. They are detected in the pancreas with established PDAC and also in the tumor-free pancreas from individuals with familial predisposition to PDAC (Canto, Hruban et al. 2012). Approximately 90% of PanINs of all grades have KRAS mutations (Kanda, Matthaei et al. 2012). High grade PanINs specifically associates with the mutational inactivation of the CDKN2A, p53, and

SMAD family member 4 (SMAD4) tumor suppressors. These data suggest that KRAS mutations contribute to the inception of tumor (Hustinx, Leoni et al. 2005).

Intraductal papillary mucinous neoplasms (IPMNs) are another type of precursor lesions to pancreatic cancer. Even though IPMNs are asymptomatic, they are associated with an increased risk of invasive PDAC. IPMNs arising from the main pancreatic duct are more malignant than those arising from the ductal branches. IPMNs also harbor KRAS mutations (Wu, Matthaei et al. 2011).

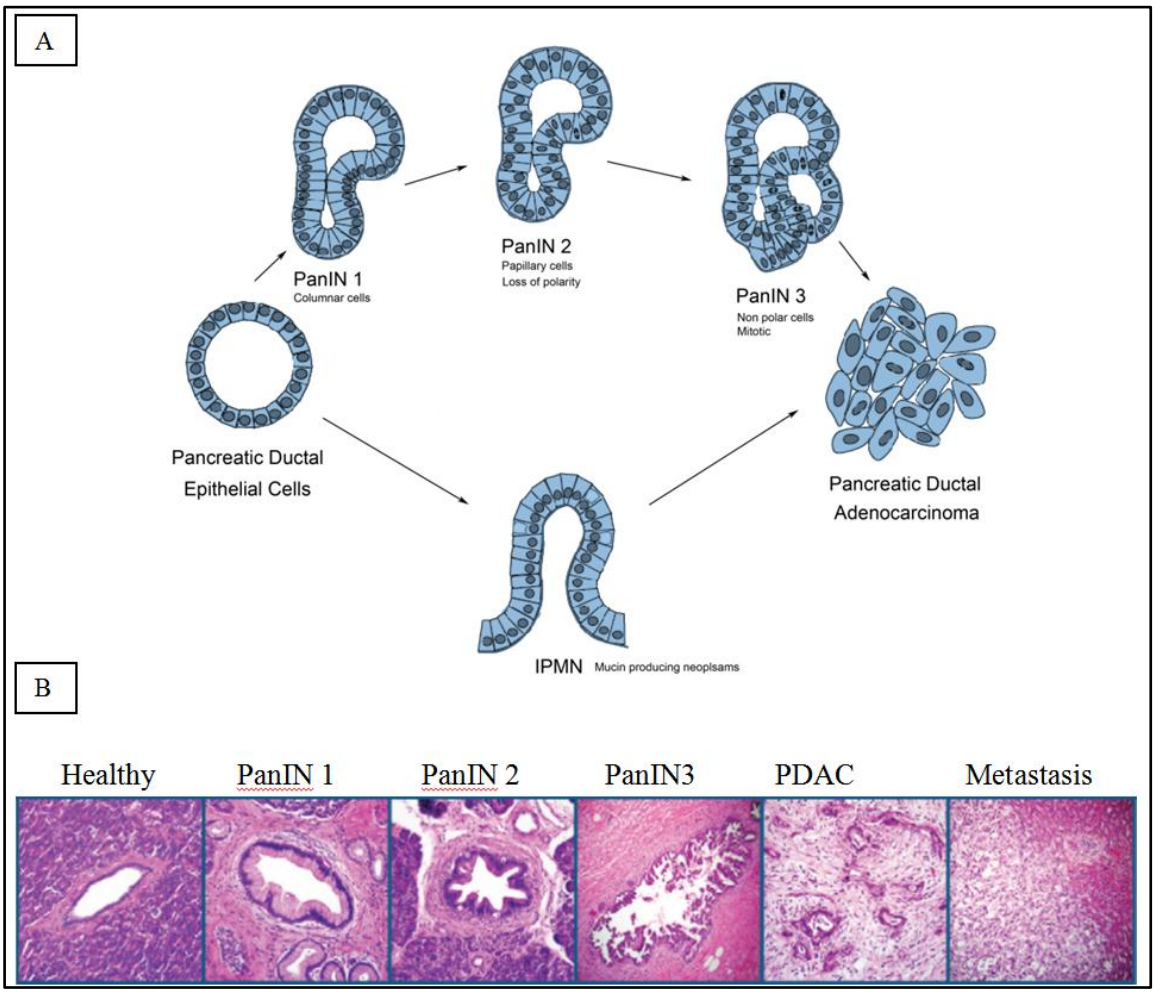


Figure 3: A: Precursor lesions leading to pancreatic ductal adenocarcinoma (redrawn based on the original illustration by Skrypek and Seuningen, 2010). **B:** Histopathology of healthy pancreas leading to metastatic PDAC (Han and Von Hoff 2013).

1.2.4. Genetics of PDAC

Cancer is fundamentally caused by germline and acquired somatic mutations in oncogenes. Large scale sequencing of tumor tissues has led to the understanding of genetic events that drive pancreatic cancer. The pancreatic tumors have been deeply sequenced, providing greater insights into the somatic mutations (Table 2) (Jones, Zhang et al. 2008, Jiao, Shi et al. 2011, Wu, Jiao et al. 2011, Wu, Matthaei et al. 2011). These data may eventually lead to better clinical management, early detection and treatment of pancreatic cancer. The sequencing of PDAC tissues revealed four genes, *KRAS*, *p16*/*CDKN2A*, *TP53* and *SMAD4* that are somatically altered in more than 50% of the cases (Jones, Zhang et al. 2008). These genes are briefly described below:

Tumor Type	Gene(s)	Prevalence Of The Alteration	Comment
Acinar cell carcinoma	APC	15%	
	CTNNB1 (beta-catenin)	5%	
Invasive ductal adenocarcinoma	**p16/CDKN2A**	**95%**	
	TP53	**75%**	
	KRAS	**95%**	**KRAS mutations occur early, and may be a target for early detection**
	SMAD4	**55%**	**SMAD4 loss associated with poor prognosis and widespread disease**
	MLL3, TGFBR2, FBXW7, ARID1A, AIRID2, and ATM	**<5%**	**Some of these, such as ATM, may be therapeutically targetable**
IPMN	PIK3CA	10%	
	p16/CDKN2A	Dependent on histologic grade	
	KRAS	80%	
	RNF43	75%	RNF43 is a marker of mucin-producing tumors because it is present in both IPMNs and MCNs
	GNAS	60%	GNAS is a marker of IPMNs. GNAS and/or KRAS mutations are present in >95% of all IPMNs
	TP53	Dependent on histologic grade	Higher grade lesions
	SMAD4	Dependent on histologic grade	Higher grade lesions

Table 2: Genetic alterations in common neoplasms of the pancreas (redrawn based on the original data from Wolfgang et al., 2013).

***KRAS*:** It is activated by point mutation in 95% of invasive PDAC (Hruban, Vanmansfeld et al. 1993, Jones, Zhang et al. 2008). A small GTPase protein coded by the *KRAS* gene plays an important role in cell signaling via the mitogen-activated protein kinase (MAPK) pathway. The point mutations in *KRAS* can be easily identified, as they target 3 codons (codons 12, 13, and 61). Loss of KRAS expression results in massive cell death and cell cycle arrest which in turn leads to rapid tumor regression (Slater, Langer et al. 2010). This suggests that *KRAS* mutations could form the basis for gene-based diagnostics to detect early PDAC phenotype (Shi, Fukushima et al. 2008).

***p16/CDKN2A*:** In approximately 95% of pancreatic cancers, a tumor suppressor gene on chromosome 9p called *p16/CDKN2A* gene is inactivated (Jones, Zhang et al. 2008). The loss of p16 function in pancreatic cancer promotes unrestricted cell growth. Hence the protein product of the *p16/CDKN2A* gene, p16, plays an important role in the regulation of the cell cycle.

***TP53*:** It is a tumor suppressor gene on chromosome 17p. In 75% of pancreatic cancers *TP53* is inactivated (Jones, Zhang et al. 2008). This gene codes for the p53 protein that plays an important role in responses to cellular stress, activation of DNA repair mechanisms, growth arrest and apoptosis. Since p53 mediates a number of important cell functions, mutation of the *TP53* gene results in PDAC through the loss of p53 function.

***SMAD4*:** It is the fourth major gene that is somatically targeted in pancreatic cancer. It acts as a tumor suppressor gene on chromosome 18q (Hahn, Schutte et al. 1996). The protein product of the *SMAD4* gene, Smad4, plays an important role in the transforming growth factor beta (TGFβ) cell signaling pathway. Mutation of *SMAD4* results in poor prognosis of pancreatic cancer (Blackford, Serrano et al. 2009, Iacobuzio-Donahue, Fu et al. 2009).

There are several other genes in addition to these four major genes that undergo lower frequency somatic mutations in pancreatic cancer. They include genes such as – *MLL3*, *TGFBR2*, *FBXW7*, *ARID1A*, *AIRID2*, and ataxia–telangiectasia mutated gene *(ATM)* (Jones, Zhang et al. 2008, Roberts, Jiao et al. 2012). Cancers which have genetically inactivated *ATM* could be sensitive to radiation damage and to poly(ADP) ribose polymerase (PARP) inhibitors (Williamson, Kubota et al. 2012).

In addition to DNA changes, there are a number changes in gene expression in pancreatic cancer. Overexpressed genes could be interesting in a clinical perspective since they could be used not only as therapeutic targets but also as biomarkers for diagnosis of PDAC (Argani, Iacobuzio-Donahue et al. 2001, Hassan, Bera et al. 2004, Harsha, Kandasamy et al. 2009).

The molecular characteristics of invasive PDAC are useful to study other lesions in the pancreas such as PanINs. As mentioned earlier, they are precursor lesions to the infiltrating PDAC. PanINs are extensively studied for genetic changes contributing to the phenotype. PanINs have similar genetic makeup as that of PDAC invasive genotype (Maitra, Fukushima et al. 2005). *KRAS* and *p16*/*CDKN2A* mutations occur early in PanINs with low- to intermediate-grade dysplasia. In PanINs with high-grade dysplasia and in invasive cancer, *TP53* and *SMAD4* mutations are late events. These findings firmly establish PanINs as noninvasive precursors to invasive pancreatic cancer.

Thus, analyses of molecular changes in pancreatic cancer at the DNA, RNA and protein identified specific genetic alterations (Table 2). Although there is such wealth of information about molecular genetics of pancreatic cancer, there are no real studies on a possible effect of integrated pathogenic foreign DNA sequences that could potentially trigger carcinogenesis especially in PDAC. Tumor viruses have been known to integrate into the host cell genome and transform the cell to a malignant phenotype. To provide a better understanding of tumor viruses the following sections provide a brief description to the field of tumor virology.

1.3. Tumor viruses

Cancer has been well recognized as a collection of related diseases. Currently, there are diverse theories on mechanisms that cause cancer. However, virus infection as one of the etiologies was not considered until the late 20th century. Although there was substantial evidence for an involvement of viruses in inflammation and cancer, these hypotheses did not receive due credit by the scientific community. This has now changed. A recent report of the International Agency for Research on Cancer (IARC) estimates that approximately 16% of the new cancer cases are attributed to infection, the majority of which caused by viruses (Stewart and Wild 2014). The burden of infection is probably even significantly higher in less developed countries.

1.3.1. Brief History of Tumor Virology

Since its inception, the research on the association of tumors and viruses has provided groundbreaking concepts for the causes of human cancer. In 1911 Peyton Rous first identified an avian virus that induced tumors in chickens; the significance of his findings was appreciated by the scientific community after 40 years. In the 1930s and later in the 1950s, Richard Shope and John Bittner, and Ludwik Gross, respectively identified viruses that cause tumors in mammals. This created interest in the scientific community to investigate such similar associations between viruses and human cancers. Consequently, in the 1960s and 1970s, the first human tumor viruses – EBV, hepatitis B virus (HBV), and human papillomaviruses (HPV) – were identified to play an important role in causation of the disease. Later in the 1980s and 1990s, the feasibility to use advanced technologies led to the identification of few more human tumor viruses: human T-cell leukemia virus type 1 (HTLV-1), hepatitis C virus (HCV), and Kaposi's sarcoma associated herpesvirus (KSHV) (Javier and Butel 2008).

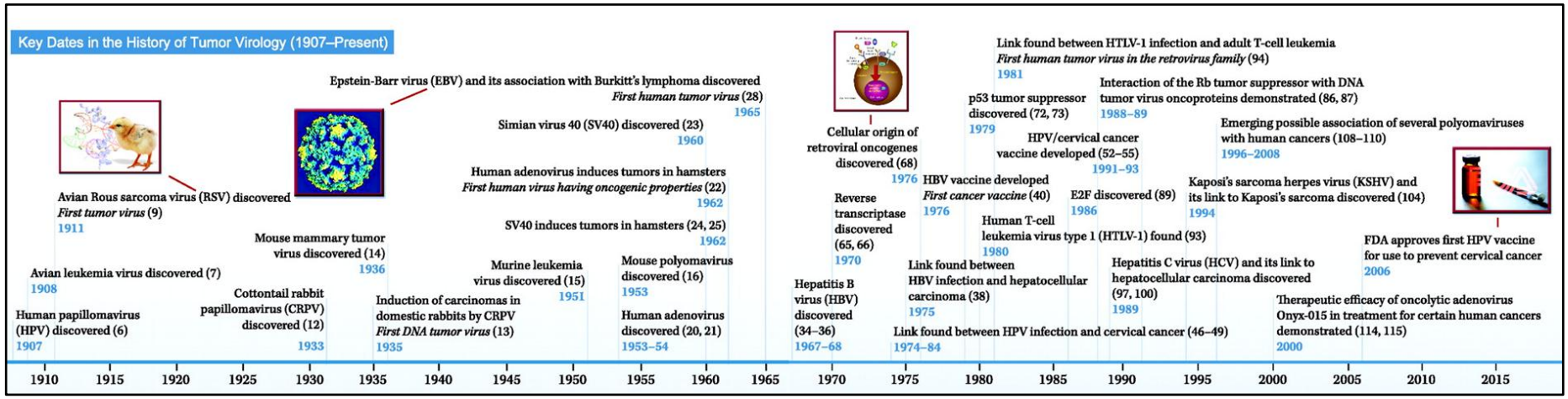


Figure 4: Timeline of advances in tumor virology (updated and adapted from the original illustration by Javier & Butel, 2008).

1.3.2. Tumor Viruses Do Not Necessarily Follow Koch's Postulates

Already in the early 19th century, Jakob Henle assumed that microorganisms are the causative agents of infectious diseases (Henle 1840). In 1884, his student Robert Koch as well as Friedrich Loeffler formulated four postulates to establish a causal relationship between a pathogen and a disease (Koch 1884, Loeffler 1884). These postulates offer a general guideline to identify and isolate the pathogens. Koch's postulates have also influenced scientists to examine pathogenesis from a molecular point of view. The four postulates are:

1. The microorganism must be found in abundance in all organisms suffering from the disease, but should not be found in healthy organisms.
2. The microorganism must be isolated from a diseased organism and grown in pure culture.
3. The cultured microorganism should cause disease when introduced into a healthy organism.
4. The microorganism must be reisolated from the inoculated, diseased experimental host and identified as being identical to the original specific causative agent.

It was soon recognized that some pathogens like Epstein - Barr virus (EBV), even though they did not fulfill all the criteria of the postulates, were responsible for disease. EBV is a common infection that can cause a rare cancer called Burkitt's lymphoma. In general, EBV cannot be isolated as pure cultures *in vitro* or used to re-infect susceptible hosts (Koch 1942, Fredricks and Relman 1996). However, Werner and Gertrude Henle, and colleagues showed a unique property of EBV that upon infection EBV immortalizes primary B cells (Henle, Diehl et al. 1967). Later oncogenes of EBV were characterized (Ernberg and Klein 2007) and additional EBV-associated tumors were described (Saemundsen, Purtilo et al. 1981). Currently, a number of viruses are an accepted cause of cancer, irrespective of their

fulfillment of criterion in Koch's postulates. Therefore, although Koch's postulates are historically important and provide a guideline for microbiologic diagnosis, fulfillment of all four postulates is not mandatory to demonstrate disease causality particularly in relevance with tumor virology. Hence, Harald zur Hausen (zur Hausen 1999) proposed alternative criteria for defining a causal role for an infection in cancer:

(i) Epidemiological plausibility and evidence that a virus infection represents a risk factor for the development of a specific tumor.
(ii) The consistent presence and persistence of the genome of the microbe in cells of the tumor.
(iii) The stimulation of cell proliferation following transfection of the genome (or portions of it) in corresponding tissue culture cells.
(iv) The demonstration that the genome of the agent induces proliferation and the malignant phenotype of the tumor.

The subsequent sections of this chapter exclusively address mechanistic insights of human tumor viruses in causing various types of cancers.

1.3.3. Direct and Indirect Viral Tumorigenesis

Viral cancer agents can be broadly classified into two categories: direct carcinogens and indirect carcinogens. Direct carcinogens express viral oncogenes that are directly involved in the transformation of healthy normal cells into a tumor. Direct viral carcinogens are usually present in cancer cells, expressing at least one active transcript to mediate tumor transformation. Direct carcinogenesis is observed with HPV, Merkel cell polyomavirus (MCV), and KSHV related cancers. Knockdown studies with the loss or down-regulation of viral proteins, ultimately resulting in the loss of host cancer viability (Steele, Cowsert et al.

1993, Tan and Ting 1995, Goodwin and DiMaio 2000, Dirmeier, Hoffmann et al. 2005, Godfrey, Anderson et al. 2005, Wies, Mori et al. 2008, Houben, Shuda et al. 2010) support direct involvement of viruses in causing cancer. Indirect carcinogens cause chronic inflammation by active infection leading to mutations that are tumorigenic to the healthy normal cells (Parsonnet 1999, Zur Hausen 2001). Indirect carcinogens include 'hit-and-run' viruses which are responsible for inflammation and thus, initiation of tumorigenesis, but the viral genes are lost as the tumor develops. Good examples of this type that are under investigation are gamma herpesviruses such as EBV (Ambinder 2000), (Stevenson, May et al. 2010).

HBV, HCV and HTLV-I do not fit precisely into either of the above two categories. HBV and HCV induce chronic cell death and regeneration after prolonged liver cirrhosis leading to hepatocellular carcinoma (HCC) (Seeger and Mason 2000, Mason, Liu et al. 2010, Tsai and Chung 2010). Although in the majority of HBV-related cancers HBV sequences are integrated into the genomes of tumor cells, it still remains to be investigated if HCC cell proliferation is dependent on stable expression of HBV (or HCV) genes (Seeger and Mason 2000).

1.3.4. Dysregulation of the Cell Cycle Machinery by Viral Oncogenes

Viral induced oncogenesis shows hallmarks such as aberrant cell proliferation and disruption of cell cycle checkpoints (Figure 5) (Sherr 1996, Vermeulen, Van Bockstaele et al. 2003, O'Nions and Allday 2004, Kumar, Saha et al. 2010). Tumor viruses replicate in the cells using host machinery and transform healthy normal cells to actively proliferate (O'Nions and Allday 2004, Kumar, Saha et al. 2010). Viral oncoproteins have a range of mechanisms to dysregulate cell cycle by evading checkpoints (Figure 5) (O'Nions and Allday 2004, Kumar, Saha et al. 2010). Viral oncogenes induce aberrant entry to S phase using interesting

strategies to avoid cell cycle arrest by the G1 -S and/ or G2 -M checkpoints, (Figure 5) (Vermeulen, Van Bockstaele et al. 2003). There is strong selection to maintain viral genes that can initiate tumorigenesis, as diverse viruses show remarkable similarity in mechanisms to target the same tumor suppressor pathways. Virus oncogenes transform host cells to immortalization by inactivating tumor suppressors like pRb and p53. While most of the human tumor viruses encode oncoproteins that target pRb and p53, the diversity of mechanisms with which they do so is large (Subramanian, Knight et al. 2002). Tumor viruses also target telomerase reverse transcriptase (TERT) (Hwang, Lee et al. 2003, Lee, Chen et al. 2007, Palermo, Webb et al. 2008, Pantry and Medveczky 2009, Stimson, Wood et al. 2009), cytoplasmic PI3K–AKT–mTOR (Smith, DeWitt et al. 2000), nuclear factor-κB (NF-κB) (Needleman, Turk et al. 1986, Levy 1997, Cho, Baek et al. 2001, Helt and Galloway 2003), β-catenin (also known as CTNNB1) (Marks, Furstenberger et al. 2007) and interferon signaling pathways (Reddy, Hirose et al. 2000), that have roles in tumorigenesis.

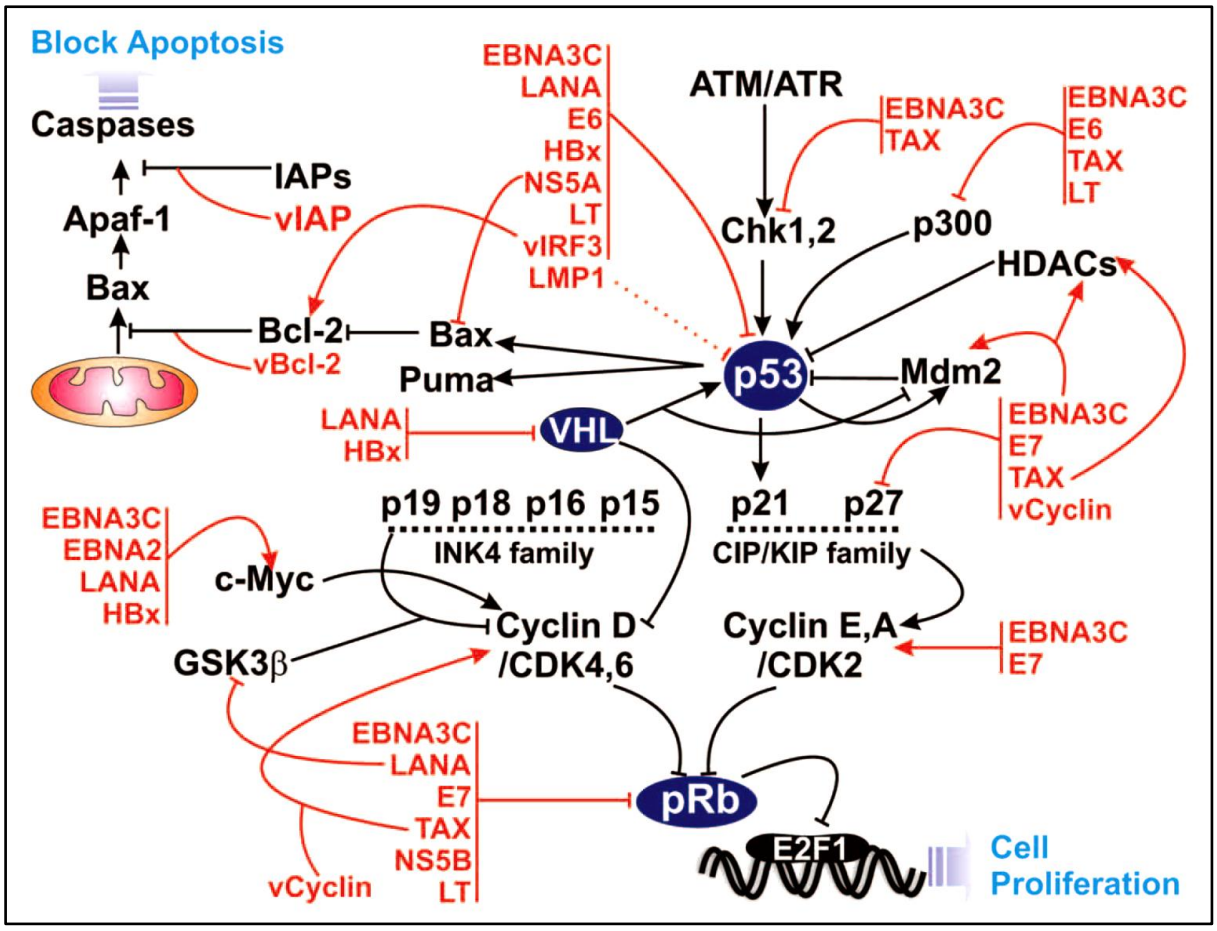


Figure 5: Viral-oncoproteins (red) dysregulate host components (black) of cell cycle and apoptotic pathways (adapted from the original illustration by Saha, Kaul, Murakami, & Robertson, 2010).

The identification of host factors that are targeted by viral oncogenes provides deep insights to the mechanisms through which viruses disrupt cell cycle. Genome-wide approaches (De Luca et al., 2003), and microarray-based analysis (Alazawi et al., 2002; Berger et al., 2002; Vasseur et al., 2003), have already provided some novel information on functional interactions between viral oncogenes and host cellular components.

1.3.5. Tumor Viruses Maintain Latency in Host Cells

Human tumor viruses cause persistent latent infections. They generally do not replicate to form active infectious virus particles in tumors cells. EBV, HBV, HTLV-1, HPV16 & 18,

HCV, KSHV, and MCV, all have the capacity to form virions and transmit between cells. However, these viruses are generally latent within tumors with a diminished or completely stopover virus replication (Dittmer and Krown 2007). Latency by a human tumor virus promotes immune evasion. It allows the virus to avoid being recognized by the immune system. During latency, viral protein expression is turned off and the virus camouflages within the host cell to escape cell-mediated immune recognition. The virus exists as a naked nucleic acid, often integrated or as an episome, relying on host cell machinery to replicate later. The discovery of EBV in Burkitt's lymphoma, mentioned earlier, also provided evidence that the tumor cells can indeed harbor viral DNA in episomes without being transmissible.

Viruses which productively replicate often initiate cell death, a situation called as the cytopathic effect (CPE). This would explain the connection between virus latency and tumorigenesis. CPE is often nonspecific innate immune response of cells to viral infection rather than a virus-induced effect. When viruses recover from latency and switch to active virion production, it triggers innate immune signaling as the cellular DNA damage repair system generates virus associated molecular patterns from various viral nucleic acids (Belanger, Gravel et al. 2001, Efklidou, Bailey et al. 2008, Thurau, Marquardt et al. 2009). The innate immune response is further amplified by activation of toll-like receptor and interferon signaling (Desaintes, Demeret et al. 1997). These immune responses kill infected cells that are undergoing virus replication by lysis completing the lytic replication. Hence, most of the human tumor viruses maintain latency after infection and cause cancer by deregulating host cellular processes.

1.3.6. Viral microRNAs in cancer

MicroRNAs (miRNAs) are broad class of small, silencing effector, noncoding RNA molecules which, in a sequence specific manner, negatively regulate gene expression. The initial discovery of small interfering (si)RNAs in virus infected plants, ultimately led to the identification of miRNAs (Hamilton and Baulcombe 1999). The miRNAs serve as genetic switches in cells, turning on and off the expression of a number of genes usually in a well-orchestrated process. However, several microRNAs are aberrantly expressed in cancer cells. MicroRNAs have long half-life and hence serve as diagnostic markers for cancer (Habbe, Koorstra et al. 2009, Matthaei, Wylie et al. 2012, Wan, Shen et al. 2012).

As mentioned earlier in the section 1.3, viruses are one of many factors that have every potential to cause proliferation of normal cells and transform them into malignant cells. Viruses develop interesting strategies through the course of evolution to surpass the immune system of the host, preventing infected cells from becoming apoptotic. The opinions on RNA silencing pathways as antiviral strategies mediated by miRNA in plants and insects has completely changed after the identification of miRNAs from EBV and herpesviruses. It is now clear that many viruses can hijack host RNA silencing machinery by producing their own miRNAs. Viral-encoded miRNAs can silence the host genes in both cis and trans manner. It helps in two ways, precise expression of viral genomes, and modification of host gene expression. In this chapter, the current information on viral miRNAs is very briefly reviewed in relation with cancer.

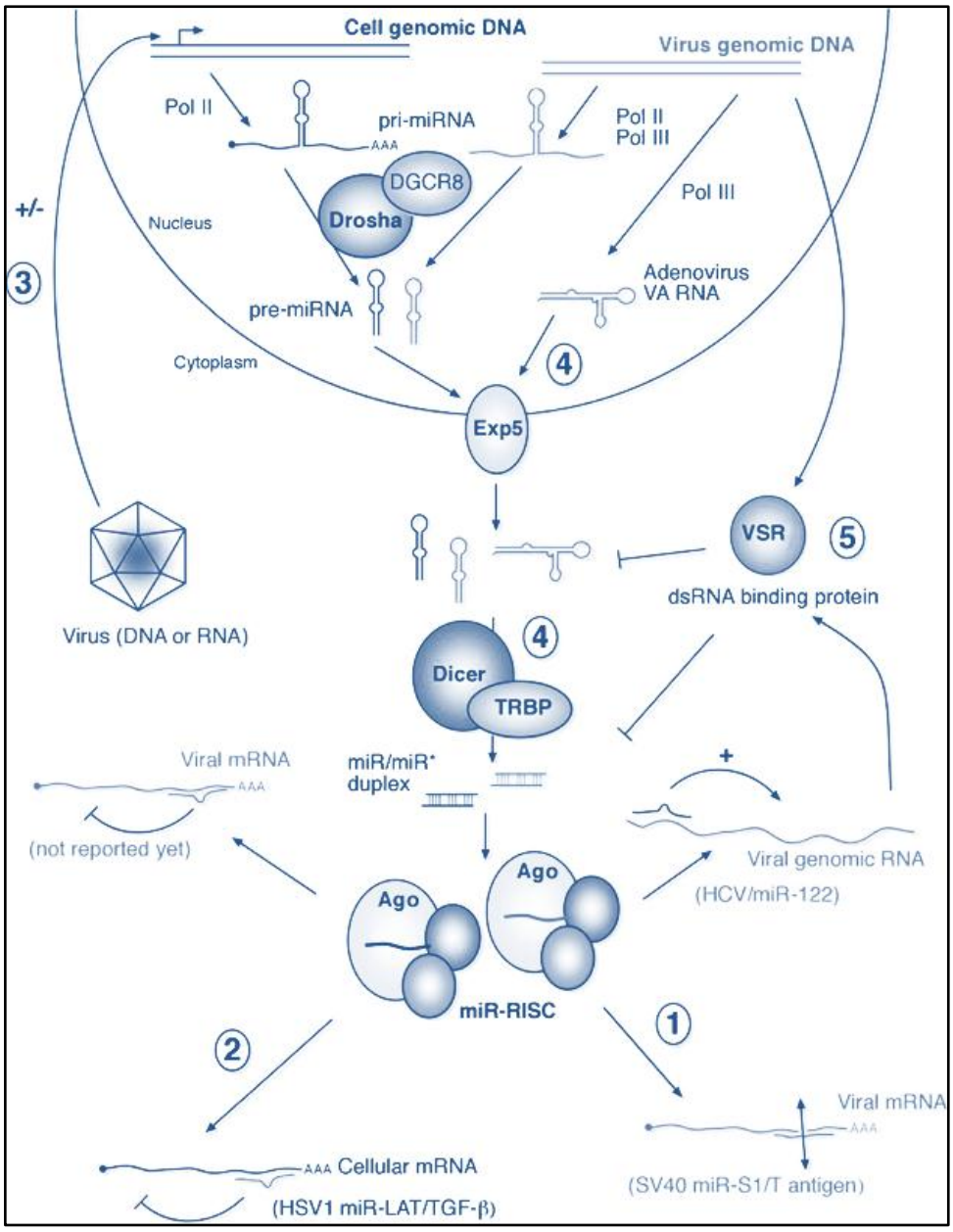


Figure 6: Viral miRNA possible modes of action (edited and redrawn based on the original illustration by Pfeffer & Voinnet, 2006). Individual steps in this figure are briefly discussed below.

(1) Viral miRNAs act in a cis manner to regulate virus host genome expression and contribute to latency. This also helps the virus to evade innate or adaptive immune system of the host. The latent infections ultimately lead to cancer development. (2) They can also act in a trans manner to cause deregulation of translation or breakdown of cellular messenger RNAs (mRNAs) that are involved in tumor suppression. (3) Interestingly, viruses also initiate

transcriptional and/or post-transcriptional dysregulation of host miRNA expression. (4) Viruses generate dsRNA molecules or pre-miRNA which compete with host miRNAs and affect cellular signaling to transform healthy cell to a tumor phenotype. (5) Viruses also produce suppressors of host RNA silencing machinery which bind to host dsRNA and interfere with cellular miRNA processor and effector complexes, leading cancer development. These recent findings provide new insights into the role of viruses and their miRNAs in cancer development.

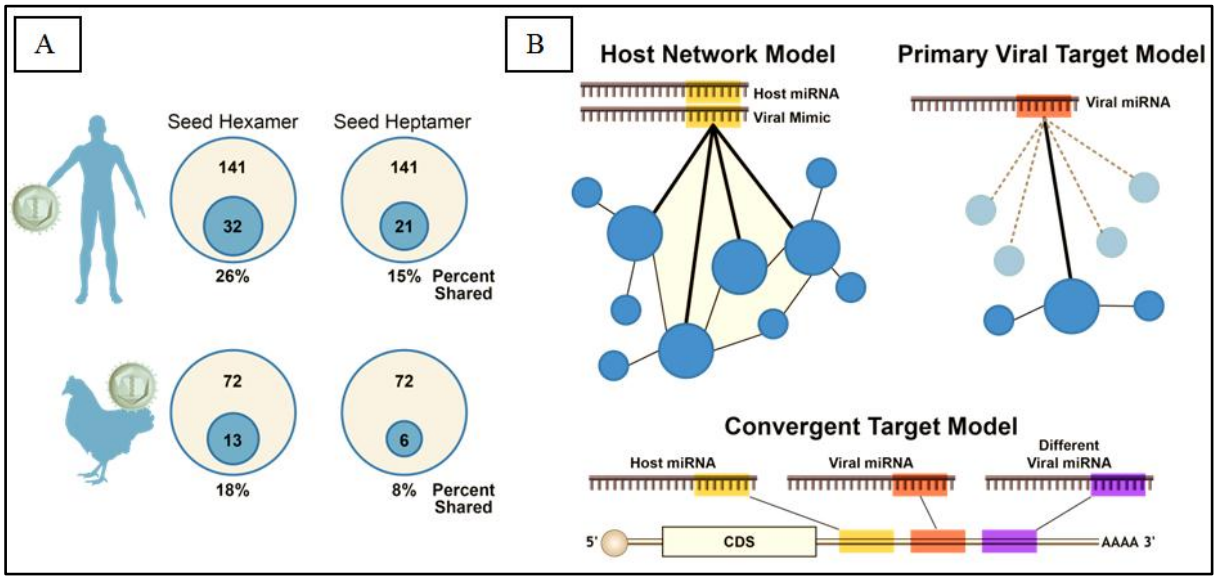


Figure 7: A. Viral miRNAs have a seed sequence homology to their host miRNA. **B.** Models of gene targeting my viral miRNAs (edited and adapted from the original illustration by Kincaid & Sullivan, 2012)

A minority of viral miRNAs, called analogs, share seed sequence identity with cellular miRNAs. In a previous studies, the miRBase version 18 annotated mature viral miRNAs for humans, and chickens, were compared with their respective host miRNAs for identity in nucleotides hexamer or heptamer seed sequence (Kincaid and Sullivan 2012). The inner circles in the figure 7A represent the number of viral miRNAs with a similarity to host seed sequence out of the total viral miRNAs. In the network models of figure 7B, analogous viral

miRNAs function as host miRNAs through seed sequence similarity, thereby targeting host genes. The binding sites for the host miRNA are conserved and allow the viral miRNA to target multiple cellular genes. This model of targeting is called host network model. The primary target model suggests that some viral miRNAs may evolve to target only one or a set of transcripts through new binding sites that are not conserved for host miRNAs. In the convergent target model, host and viral miRNAs target the same gene through different binding sites.

Many studies show evidence that viral miRNAs are involved in initiation and progression of cancer (Lovat, Valeri et al. 2011). Oncogenic pathways that viruses are known to be associated with include cell growth, proliferation, angiogenesis, genetic instability, and evasion of apoptosis. Regulation of the MAPK signaling pathway by its miRNAs is an essential for KSHV (Qin, Feng et al. 2011). The closely related herpes viruses target PI3K pathway to affect cell survival and growth leading to cancer (Vara, Casado et al. 2004, Buchkovich, Yu et al. 2008). Both pathways potentially can inhibit KSHV infection (Lambert, Shahrier et al. 2007). The PI3K pathway can also enable infected cells to withstand viral induced stress (Buchkovich, Maguire et al. 2008). DNA viruses can initiate the MAPK pathway by affecting epidermal growth factor receptor (ErbB) signaling pathway resulting in tumor growth or differentiation (Tzahar, Moyer et al. 1998). These increasing evidences suggest that viral miRNAs may play a role in the oncogenesis.

1.4. Representational Difference Analysis (RDA)

RDA is a subtractive nucleic acid hybridization technique (Lisitsyn, Lisitsyn et al. 1993, Lisitsyn, Lisitsina et al. 1995). Generally, in subtractive hybridization techniques, a DNA population called the "driver" is hybridized in higher proportions with another DNA population called the "tester". Driver and tester are similar but not identical with each other. RDA differed from earlier subtractive hybridization techniques. The genomic complexity was significantly reduced in RDA by the use of "representations" of the tester and driver DNA populations, obtained through their restriction digestion. This modification is essential for RDA to isolate the "target" DNA sequences called the "difference products" (DPs), that are present only in the tester but not in the driver.

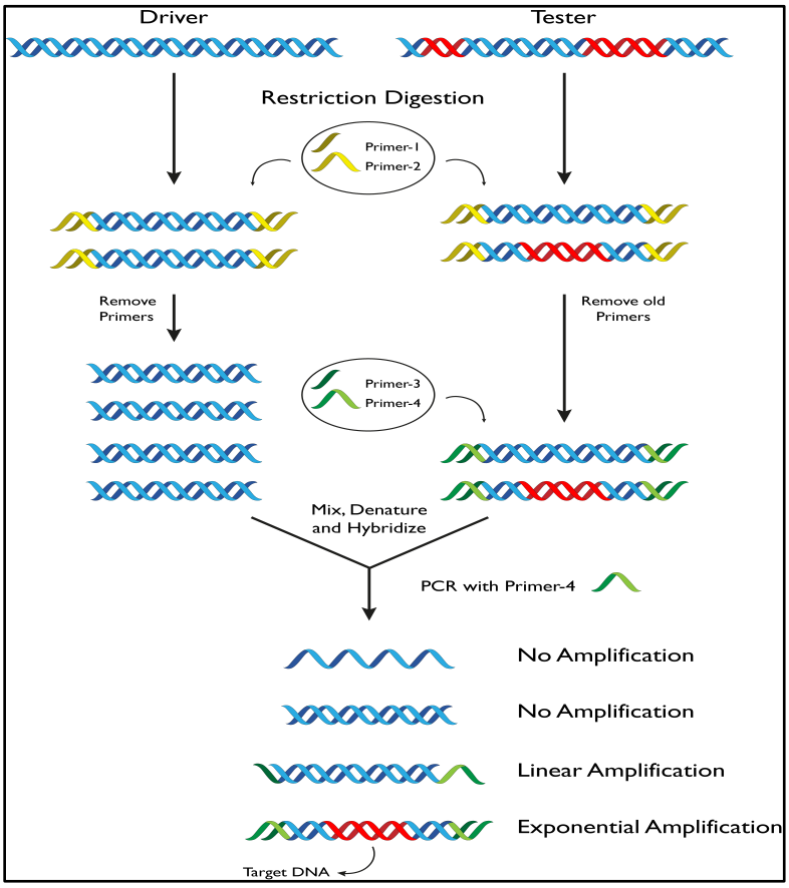


Figure 8: Basic schematic of representational difference analysis (RDA) procedure.

RDA begins with the digestion of the driver and tester DNA with restriction endonuclease followed by mixing with high driver to tester ratio. This mixture is then denatured and allowed hybridize to form three types of DNA duplexes: homo-hybrids of tester–tester and driver–driver, and tester– driver hetero-hybrids. During hybridization, increased driver to tester ratio provides a lesser chance for a tester sequence to bind to its tester complement than to bind to its driver complement. The unique "difference products" (DPs) present in the tester, thus will only hybridize with its complement from the tester DNA population. These unique target tester–tester homo-hybrids will then be cloned selectively to sequence and identify the DPs DNA information. The higher proportions of driver DNA drive the complementary tester sequences out from the pool of tester DNA, thus the name "driver". The DNA population which is used to test against driver with RDA is thus called a "tester". Strauss *et al.* improved RDA technique by performing multiple rounds of hybridization. This modification enriches the DPs present in tester that had been amplified in the previous round of hybridization, through second-order kinetics of self-association (Straus and Ausubel 1990).

1.4.1. RDA to Study Cancer Genetics

RDA is used to identify genetic aberrations, in a tumor genome by comparing it with the genome of healthy cells (Lisitsyn, Lisitsyn et al. 1993, Lisitsyn, Lisitsina et al. 1995). RDA could successfully identify homozygous deletions in tumors from a strong background of heterozygous deletions. The BRCA2 and PTEN tumor suppressor genes were identified for homozygous deletions found by RDA (Schutte, Dacosta et al. 1995, Li, Yen et al. 1997). The position of the DPs in the chromosome-specific yeast artificial chromosome clone arrays was provided by the use of genomic RDA (Zeschnigk, Horsthemke et al. 1999). The differentially expressing genes between tumor and healthy cells can be analyzed by RDA using cDNA as

the starting material instead of DNA (Hubank and Schatz 1994). High-throughput analysis of multiple representations is possible by employing cDNA-RDA during microarray hybridization (Welford, Gregg et al. 1998). Finally, isolation of sequences that are differentially methylated between healthy and tumor cells could also be analyzed using methylated CpG island amplification (MCA) coupled with RDA (Toyota, Ho et al. 1999).

Hence, we used RDA as our experimental approach to isolate the integrated pathogenic viral sequences in PDAC patients genomic DNA. This technique successfully provided us the DPs which were further analyzed using next generation sequencing (NGS). The results of these interesting studies are shown in the following sections.

2. RESULTS

2.1. Representational Difference Analysis (RDA)

Pancreatic cancer is one of the deadliest diseases with a mortality that is nearly close to the incidence. There have been numerous attempts to study genetic perturbations in the cells that lead to pancreatic cancer. Majority of these studies employ next generation sequencing techniques to identify gene amplifications, gene deletions, single nucleotide polymorphisms, mutations, and changes in gene expression. However, rare events such as integration of pathogenic virus sequences are, less and rarely studied in pancreatic cancer. Whole genome sequencing of tumors employs size selection of genomic libraries and subsequent nucleic acid purification prior to sequencing. These are although essential steps, may result in a poor identification of rare events like viral integration in the tumor genome. In the present study, RDA has significantly addressed this problem by enriching these rare events, or the difference products (DPs), by PCR amplification. The enriched DPs are then identified and characterized using next generation sequencing technologies. Even though a robust technique, RDA is highly sensitive to DNA contamination resulting in false positives. Hence, RDA was optimized and performed under sterile laminar hood.

2.1.1. Optimization of RDA with Pilot Samples

To establish the protocol for RDA, a small pilot study was conducted. In this study, human placental DNA was used as driver. Tester sample was generated by adding bacterial plasmids to the driver. The hypothesis was that the RDA between driver and tester should isolate and enrich the bacterial plasmid sequences as difference products. Considering the scope of this pilot study, TA cloning of the DPs was performed and cloned recombinant colonies were single read sequenced instead of using NGS platform.

2.1.1.1. Generation of Pilot Representations

The driver and tester genomic samples were digested with restriction endonuclease. The adapters were used to ligate to the digested genomic samples. The adapters were used as the primers for PCR to generate the corresponding representations. DNA smear on the gel increased in intensity reaching a plateau with increasing cycle number. The plateau was observed as a shift of the product toward the larger sizes or to smaller sizes with a decrease in quality. Optimal number of cycles (25) for each representation was selected, avoiding this plateau phase. The figure below shows that, with increasing PCR amplification cycle number, the efficiency as well as sensitivity of amplification decreases.

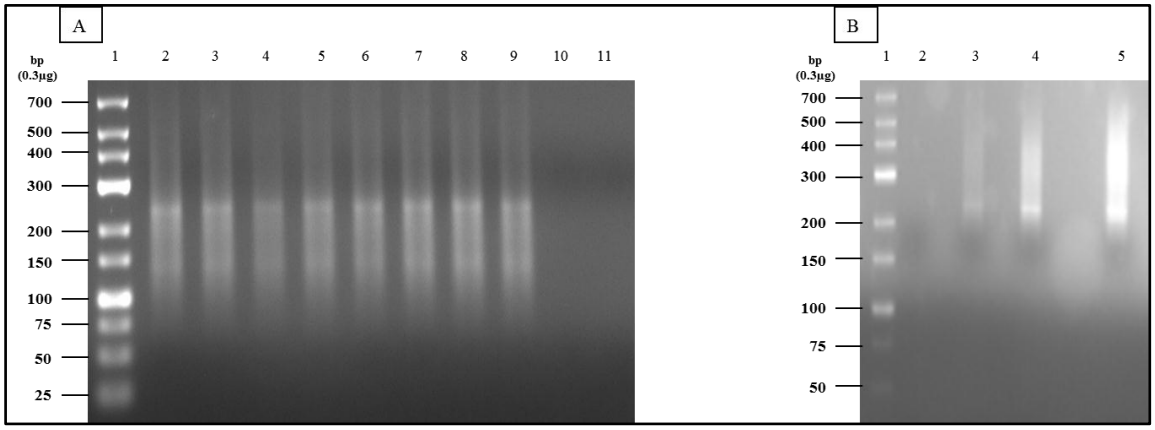


Figure 9: Representations of driver (human placental DNA) and tester (mixture of human placental DNA and bacterial plasmids). **A:** PCR was performed under sterile laminar hood using the R-Bam-24 adapter as a primer. Lanes 1 represents low range DNA ladder; 2, 3, 4 and 5 are driver representations generated by 20, 25, 30 and 35 cycles of PCR amplification respectively; 6, 7, 8 and 9 are tester representations generated by 20, 25, 30 and 35 cycles of amplification respectively; 10 and11 are PCR negative controls for 30 and 35 cycles of amplification respectively. **B:** PCR was performed outside sterile laminar hood for negative controls. With the increase in amplification cycle, this PCR was positive for DNA. Lane 2, 3, 4 and 5 represent 20, 25, 30 and 35 amplification cycles respectively.

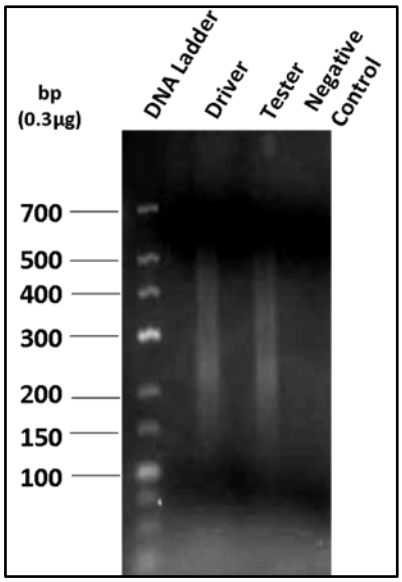


Figure 10: Representations of driver (human placental DNA) and tester (mixture of human placental DNA and bacterial plasmids) generated by selecting 25 PCR amplification cycles with R-Bam-24 primer.

2.1.1.2. First Round of Subtractive Hybridization

After generation of driver and tester representations by 25 cycles of amplifications each, the tester was prepared for subsequent subtractive hybridization to driver. To achieve this, the tester representation with R-Bam-24 adapters was subjected to DpnII digestion and new adapters J-Bgl-12 and J-Bgl-24 were ligated. The success of this ligation was checked for the same 25 cycles of amplification to maintain uniformity of the tester representation. However, the adapters from the driver were removed prior hybridization to the processed tester. PCR of the first round hybridization mixture with J-Bgl-24 resulted in the amplification of DP1 with visible bands on the agarose gel. Different hybridization temperatures were used for the first round to optimize the subtraction (data not shown). The least hybridization temperature at which non-specific hybridization and amplification does not occur was selected for further rounds of subtraction.

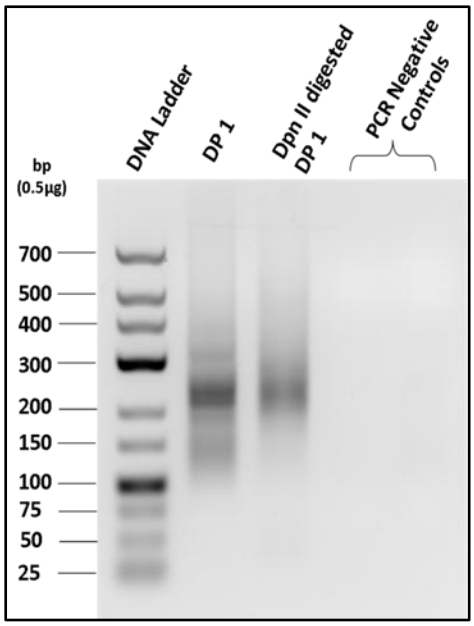


Figure 11: First round of RDA driver and tester (pilot study) produced DP1 with visible bands of amplification. To ensure that the DP1 was amplified with J-Bgl-24, the DP1 was further digested using DpnII. The third lane represents the digested DP1. The shift of the bands towards the smaller size confirms the DP1 amplification post hybridization.

2.1.1.3. Enrichment of Difference Products

The DP1 obtained after the first round of RDA is prepared for subsequent enrichment of sequences present only in tester but not in driver. The DP1 was digested with DpnII and new adapters, N-Bgl-12 and N-Bgl-24, were ligated. The enrichment of differential sequences is possible by using DP1 in place of tester during the second round of hybridization. Using the same hybridization temperature optimized during the first round of hybridization, the second round of RDA was performed. The DP2 showed enrichment of differential sequences. The third round of RDA was performed using DP2 in place of tester against driver. Since the complexity of the samples is less, three rounds of RDA were sufficient to isolate unique bacterial sequences present in the tester. The highly enriched DP3 was successfully purified and cloned using TOPO TA cloning system and the recombinant colonies were single read sequenced.

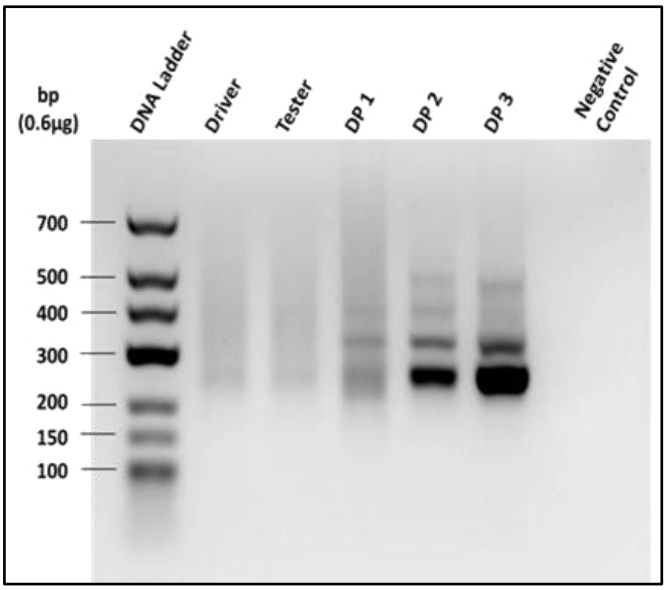


Figure 12: Three rounds of RDA. The difference products of each round of subtraction DP1, DP2 and DP3 were shown beside the tester representation used for the subtraction against driver representation. The PCR negative control showed no DNA signal.

2.1.1.4. Sequence Analysis of Difference Products

The recombinant colonies obtained by cloning DP3 were selected for single read sequenceing. The sequencing data from 96 colonies was used for the analysis. The sequences were aligned to human reference genome from NCBI using BLAST algorithm. All sequences showed a high similarity to the bacterial cloning vectors confirming that the RDA has specifically enriched unique target sequences present only in the tester but not in the driver. Since the target sequences were the bacterial plasmid sequences that were added to the human placental DNA, the driver, subtractive hybridization isolated and amplified these sequences from the driver. Further, there were no similarity hits with human genome confirming no false were picked up by the PCR amplifications or nucleic acid purification procedures.

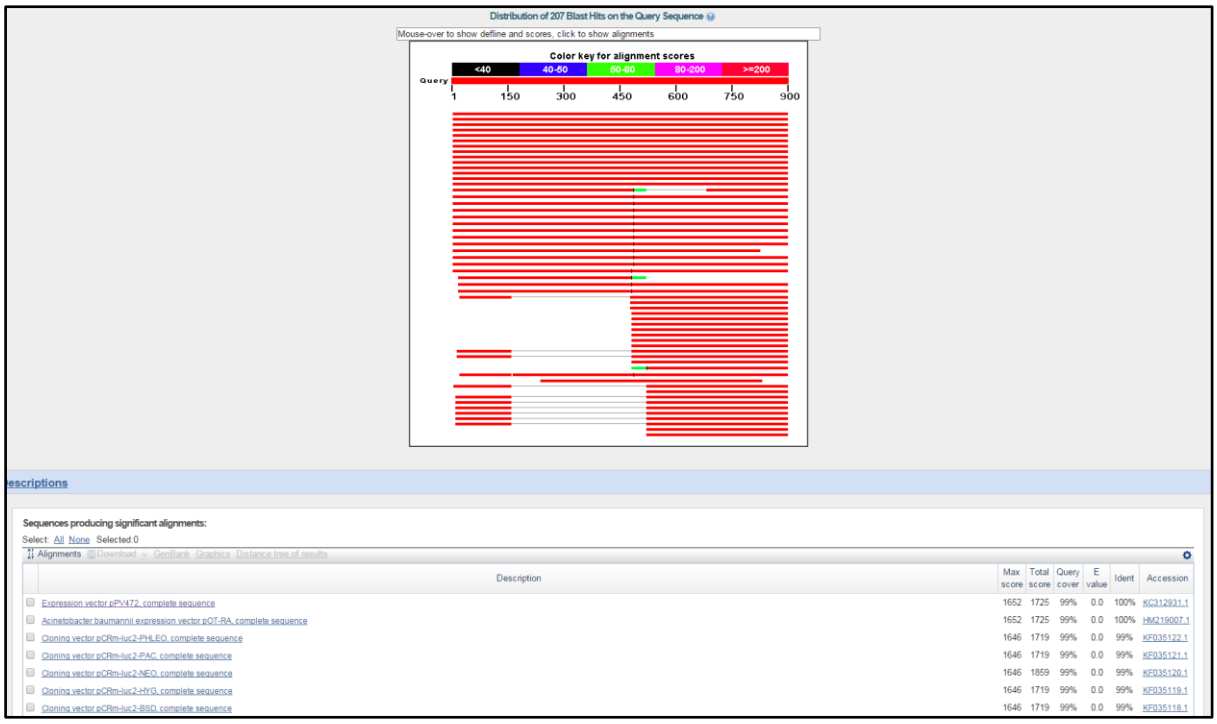


Figure 13: The sequenced DPs (representative) were used to align to all genomic databases by BLAST. There were no matches with human genome. The DPs were identical to the bacterial plasmid sequences shown at the bottom panel of the figure.

2.1.2. RDA between healthy normal tissue DNA and PDAC tissue DNA

After optimizing the RDA technique with pilot samples, the next step was to use RDA to isolate DPs from PDAC patients DNA. A pool of each, healthy controls and PDAC patients genomic DNA, were analyzed using this technique. Equal amount of DNA from each set of individual healthy samples were added to prepare a normal driver sample. Similarly, tester sample was prepared by adding equal amount of DNA from PDAC patients genomic DNA. The pooling of samples facilitates two functions; it increases diversity among each set of samples being hybridized and increases the coverage of hybridizing-difference products between driver and tester. As mentioned earlier in the chapter 1.4, the RDA is highly prone is contamination if performed on a daily lab bench environment. Hence, care was taken not to contaminate the DNA samples, especially after the restriction digestion. After hybridization

between driver and tester, the PCR second order kinetics using the adapters J-Bgl-24 and N-Bgl-24 as primers pulled difference products from the tester and were enriched during subsequent rounds of RDA.

2.1.2.1. Generation of Healthy and PDAC Representations

The pooled normal healthy control DNA samples and PDAC DNA samples were digested with restriction enzyme and after adapter ligation, a representation of the startup samples were generated using R-Bam-24 as primer. As in pilot studies, an optimal number of amplification cycles were selected in which, both driver and tester representations look similar on an agarose gel. The second lane in the figure below shows no signal of DNA contamination in the negative control.

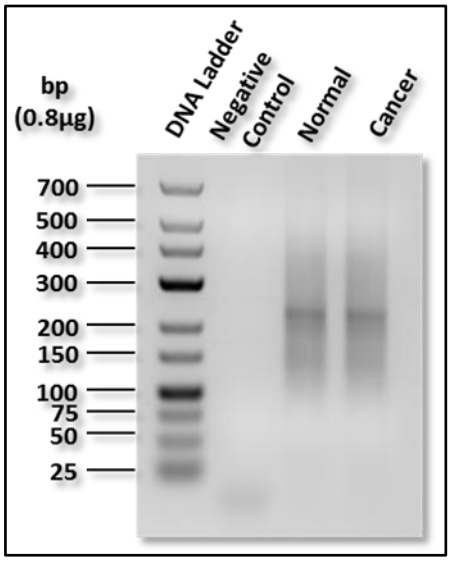


Figure 14: Representations of driver and tester generated from normal healthy and PDAC patients genomic DNA. The amount of DNA ladder loaded was indicated in parenthesis to scale the signal intensity in normal and cancer representations.

2.1.2.2. First Round of Hybridization between Healthy and Tumor Representations

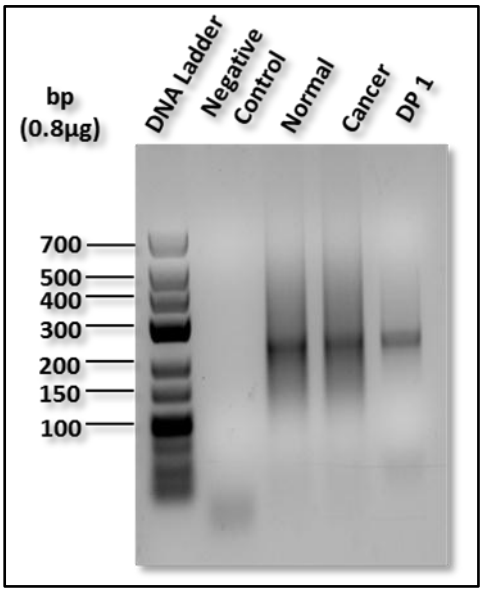


Figure 15: First round of RDA between healthy driver and PDAC tester samples. The amount of DNA ladder loaded was indicated in parenthesis to scale the signal intensity in normal and cancer representations.

The driver was produced in excess by pooling multitude of PCR representations and purifying it from the polymerase and free dNTPs. The driver is then prepared for the first round of hybridization by digesting the adapters using restriction enzyme. After digestion, two clear populations of DNA were seen on the agarose gel (data not shown). The population with larger size DNA is the driver alone region and the population with smaller size DNA contains the adapters. The driver was extracted from the gel and was hybridized with processed tester. The tester was prepared by removing the R-Bam adapters and ligating with J-Bgl adapters. The tester population for the first round of hybridization was produced using J-Bgl-24 as primer in the subsequent PCR. The tester was also purified of polymerase and free dNTPs to ensure optimal hybridization with driver. The difference product obtained after the first round RDA lies between the size range of driver and tester representation. The clean negative control was crucial at this step as contamination from external sources will be picked up easily during hybridization.

2.1.2.3. Enrichment of DPs between Healthy and Tumor Representations

The DP1 obtained in the first round of RDA between healthy and the tumor representation, was isolated from the gel to change the adapters and prepare for subsequent rounds of hybridization. In addition to the number of rounds of RDA performed in relatively simpler pilot study, one more round of RDA was performed between healthy driver and PDAC tester as the number and complexity of samples is much higher than the previous pilot study. The genomic sequences present only in the PDAC samples were enriched with every round of RDA. The DPs were diluted and hybridized against the driver in different ratios before hybridization. This step not only increased the competition of similar sequences between present in both samples but also amplified the target sequences exponentially.

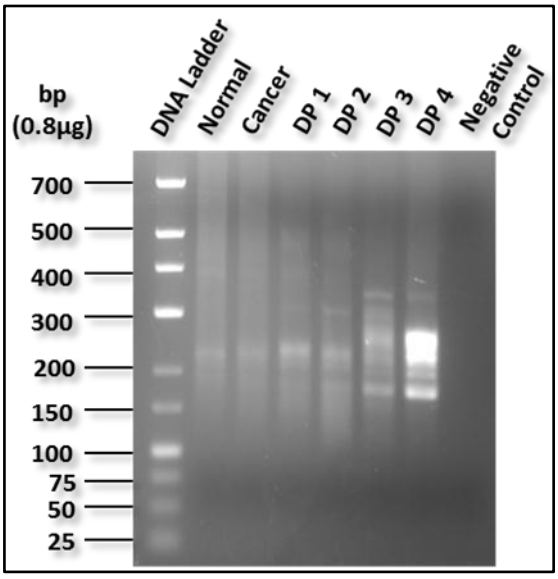


Figure 16: Four rounds of RDA producing corresponding DPs. The differential genomic sequences between driver and tester are enriched with each round of hybridization.

2.1.2.4. Bioanalyzer High Sensitivity DNA Assay

Considering the enrichment factor of differential sequences, DP3 and DP4 are selected to further analyze with Bioanalyzer high sensitivity assay. The y-axis in the figure below represents florescence units (FU) and the x-axis represents the DNA fragment length, as a peak in the area of the graph, in base pairs. The differential target sequences were significantly enriched by amplification as can be seen by the increased FU in the DP4 than DP3. The peaks with similar number of base pairs appear in both DP3 and DP4 further confirming the enrichment of same fragments from previous rounds of RDA.

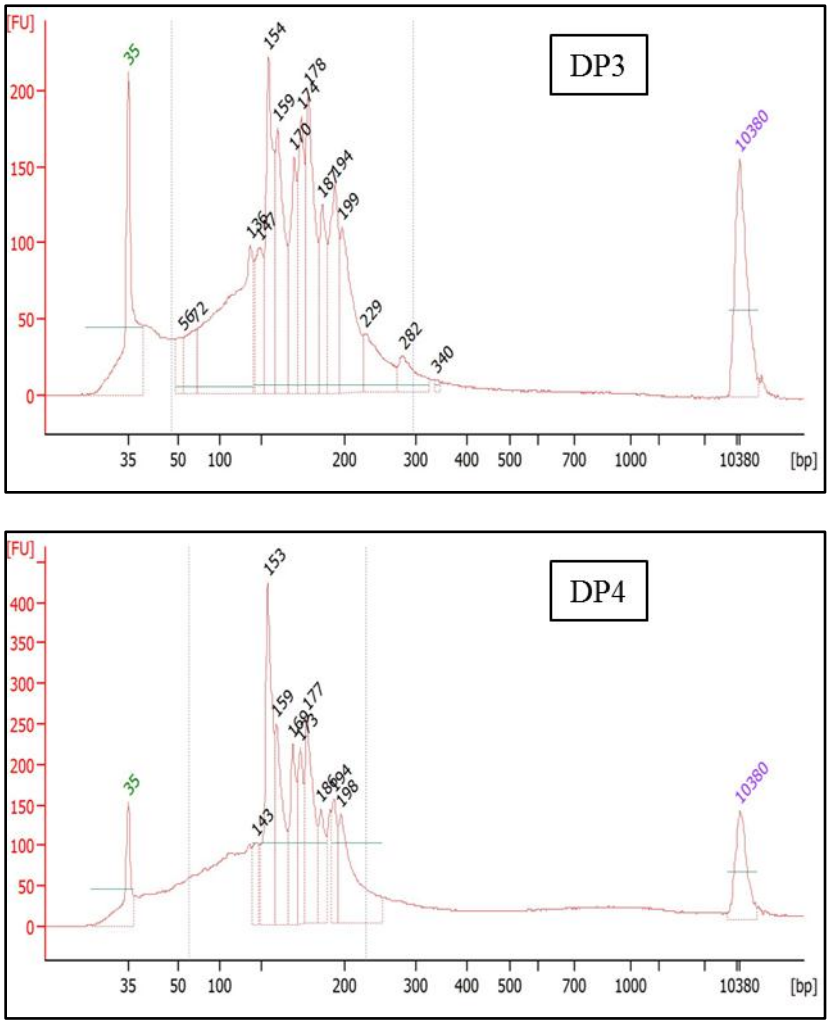


Figure 17: Electropherogram of DP3 and DP4 by analyzing them with high sensitivity DNA assay on Bioanalyzer. First and last peaks in the graphs represent loading markers.

2.1.2.5. Next Generation Sequencing (NGS) of DPs

During sequencing, the first four bases of the reads are critical to assign the read clusters by the Illumina platform. The DP3, and DP4, were amplified using adapter as primer. This resulted in all DNA fragments of the DPs with identical primer nucleotides at the ends. Hence it was necessary to randomly digest DPs with ds-DNA fragmentase for 10min before sequencing. The DP3, and DP4, were then successfully paired end sequenced.

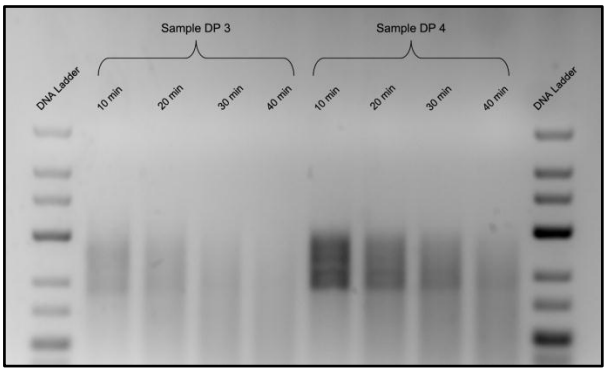
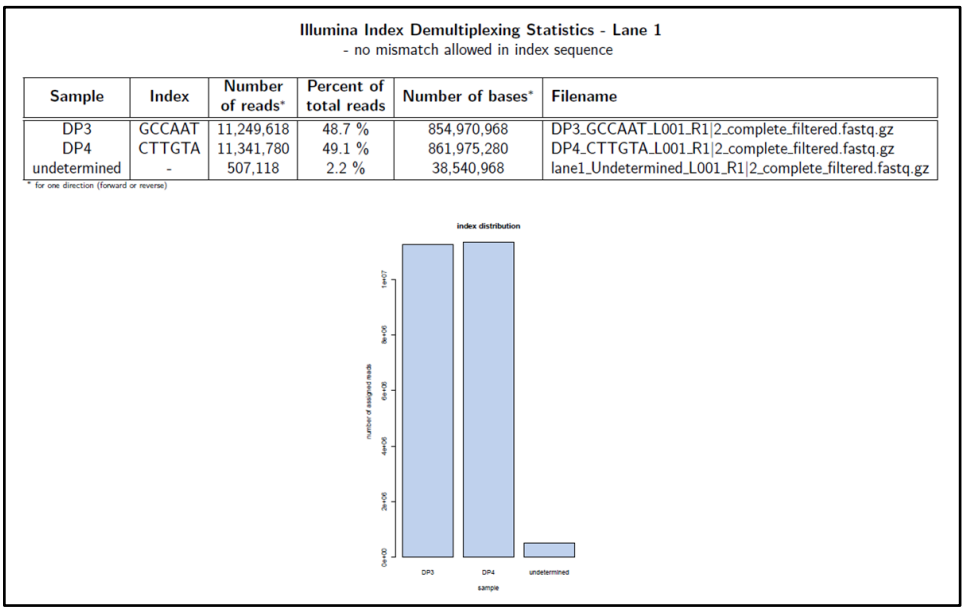


Figure 18: dsDNA fragmentase digestion of DPs at different time points. 10 min time point was selected since it provided lesser degradation of DPs than higher time points.



Illumina Index Demultiplexing Statistics - Lane 1
- no mismatch allowed in index sequence

Sample	Index	Number of reads*	Percent of total reads	Number of bases*	Filename
DP3	GCCAAT	11,249,618	48.7 %	854,970,968	DP3_GCCAAT_L001_R1|2_complete_filtered.fastq.gz
DP4	CTTGTA	11,341,780	49.1 %	861,975,280	DP4_CTTGTA_L001_R1|2_complete_filtered.fastq.gz
undetermined	-	507,118	2.2 %	38,540,968	lane1_Undetermined_L001_R1|2_complete_filtered.fastq.gz

* for one direction (forward or reverse)

Figure 19: Extent of sequencing DP3 and DP4 on a single lane.

After the quality control, lllumina sequencing of the two samples showed that the level of sequence duplication is higher in DP4 than DP3. This further confirms that RDA enriched the sequences by amplification without bias. We used two samples on a single lane during Illumina paired end sequencing. In a diverse library most sequences will occur few times. A low level of duplication may indicate a very high level of coverage of the target sequence, but a high level of duplication is more likely to indicate enrichment of relatively few target sequences as expected between two samples of same species.

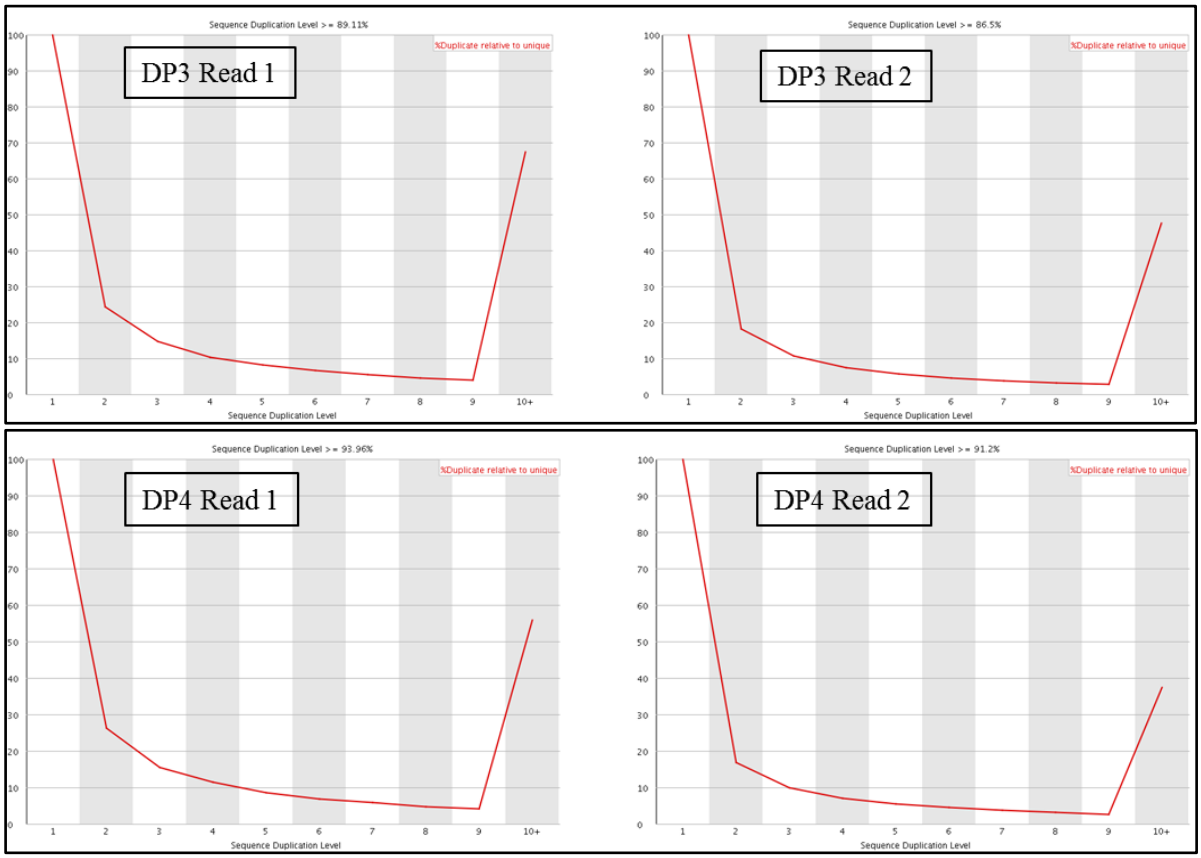


Figure 20: Sequence duplication level of both paired end reads of DP3 and DP4.

2.1.2.6. In silico Analysis of the NGS Data from DPs

The NGS data obtained after sequencing DP3 and DP4 from our core facility is analyzed using a modified version of the pipeline previously designed for identifying viral miRNAs in PDAC samples (figure 22). The raw data was filtered for the reads without sequencing adaptors. The cleaned reads were aligned with reference human genome from NCBI allowing few mismatches. As expected more than 90% of the reads aligned with human genome, since there would be few nucleotide changes between the driver and tester that would be picked up by the RDA owing to its sensitivity. The reads that were not matched with human genome were then aligned with Meleagrid herpesvirus-1 (MeHV-1), also called Herpes Virus of Turkeys (HVT), genome (we already had strong evidence for the presence of this in PDAC by the miRNA content analysis). The reads that are aligned with HVT specific genes are further analyzed.

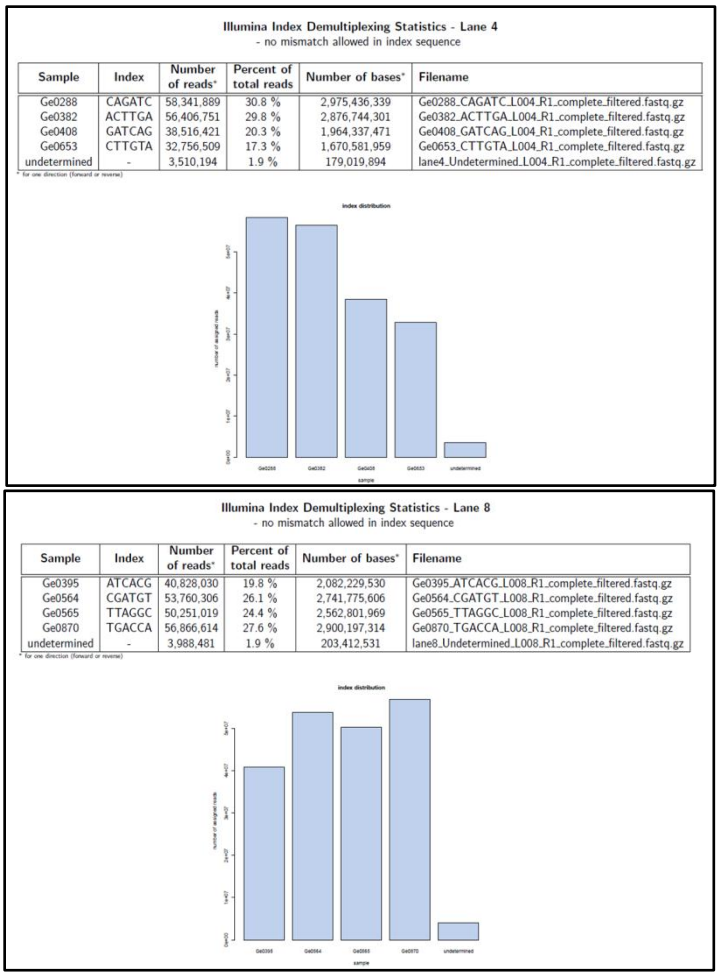
RDA of Healthy and Tumor Samples
NGS Analysis of DPs – Alignment against HVT

HVT Specific Genes	DP3	DP4
HVT001	No	No
HVT002	No	No
HVT003	No	No
HVT004	No	No
HVT072	No	Yes
HVT073	No	No
HVT074	Yes (< 10)	Yes (<10)
HVT075	Yes (~ 150)	Yes (~ 150)
HVT084	No	No

Table 3: Homology of DP3 and DP4 reads with HVT specific genes. HVT075 is representing significantly higher number of reads than other HVT specific set of genes.

2.2. MicroRNA Sequencing of PDAC

During the course of the project, a new process became available by which the identification of viruses in tumor samples became possible directly, namely by using the sequence analysis of the miRNA content of cells. Because of the limited number of molecules, this analysis form does not need any selection and amplification step in order to achieve a sufficiently enough sensitivity to find viral microRNAs. Upon the identification of viral microRNAs, the sequence of the relevant virus could be looked for selectively within the human DNA from NGS analysis of RDA. Hence, we sequenced miRNA from eight PDAC patients. We sequenced four samples per lane using Illumina MiSeq platform.



Illumina Index Demultiplexing Statistics - Lane 4
- no mismatch allowed in index sequence

Sample	Index	Number of reads*	Percent of total reads	Number of bases*	Filename
Ge0288	CAGATC	58,341,889	30.8 %	2,975,436,339	Ge0288_CAGATC_L004_R1_complete_filtered.fastq.gz
Ge0382	ACTTGA	56,406,751	29.8 %	2,876,744,301	Ge0382_ACTTGA_L004_R1_complete_filtered.fastq.gz
Ge0408	GATCAG	38,516,421	20.3 %	1,964,337,471	Ge0408_GATCAG_L004_R1_complete_filtered.fastq.gz
Ge0653	CTTGTA	32,756,509	17.3 %	1,670,581,959	Ge0653_CTTGTA_L004_R1_complete_filtered.fastq.gz
undetermined	-	3,510,194	1.9 %	179,019,894	lane4_Undetermined_L004_R1_complete_filtered.fastq.gz

* for one direction (forward or reverse)

index distribution

Illumina Index Demultiplexing Statistics - Lane 8
- no mismatch allowed in index sequence

Sample	Index	Number of reads*	Percent of total reads	Number of bases*	Filename
Ge0395	ATCACG	40,828,030	19.8 %	2,082,229,530	Ge0395_ATCACG_L008_R1_complete_filtered.fastq.gz
Ge0564	CGATGT	53,760,306	26.1 %	2,741,775,606	Ge0564_CGATGT_L008_R1_complete_filtered.fastq.gz
Ge0565	TTAGGC	50,251,019	24.4 %	2,562,801,969	Ge0565_TTAGGC_L008_R1_complete_filtered.fastq.gz
Ge0870	TGACCA	56,866,614	27.6 %	2,900,197,314	Ge0870_TGACCA_L008_R1_complete_filtered.fastq.gz
undetermined	-	3,988,481	1.9 %	203,412,531	lane8_Undetermined_L008_R1_complete_filtered.fastq.gz

* for one direction (forward or reverse)

index distribution

Figure 21: Extent of sequencing miRNA from PDAC patients.

2.2.1. Digital MicroRNAome Subtraction

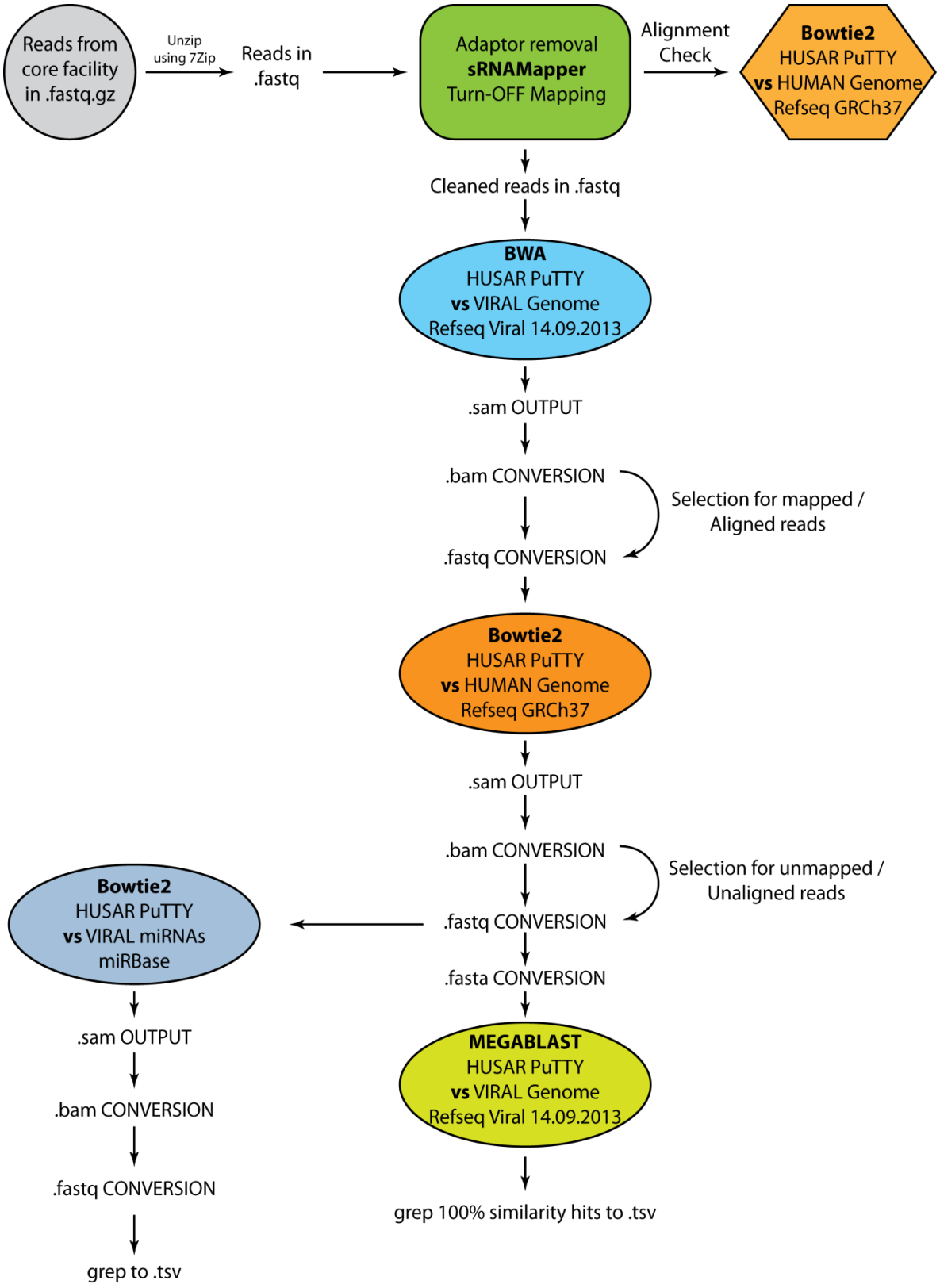


Figure 22: Pipeline designed to analyze the miRNA content of PDAC for virus sequences.

The miRNA sequencing data from the core facility is analyzed for viral sequences. After cleaning the raw data for sequencing adapters, the alignment with human genome showed good similarity. This step ensured us that most of the adapters are removed from the data. The cleaned reads are now aligned with viral genome to select sequences that have viral origin. The reads that matched with viruses could also be present in humans as both species coevolved. Hence, the virus matched reads now are aligned with human reference genome. This resulted in digital subtraction of the miRNA from human genome. The reads that did not align with human genome are uniquely virus. These reads are now aligned with viral miRNAs from miRBase.

2.2.2. Identification of hvt-miR-H14-3p in PDAC

The miRNA sequencing study showed us many unique viral miRNAs present in PDAC patients. However, we identified a top candidate miRNA that is present in higher number of reads than other viral miRNAs. This miRNA is from MeHV-1 (also known as HVT). The hvt-miR-H14-3p was one of the few miRNAs that also shares sequence homology with two human miRNAs, hsa-miR-221-3p and hsa-miR-222-3p. Apart from hvt-miR-H14-3p, we also identified few other miRNAs from this virus in the PDAC patients. However, hvt-miR-H13 was the only other miRNA from this virus that showed considerable number of reads in the PDAC in comparison to other miRNAs and was next to the top candidate.

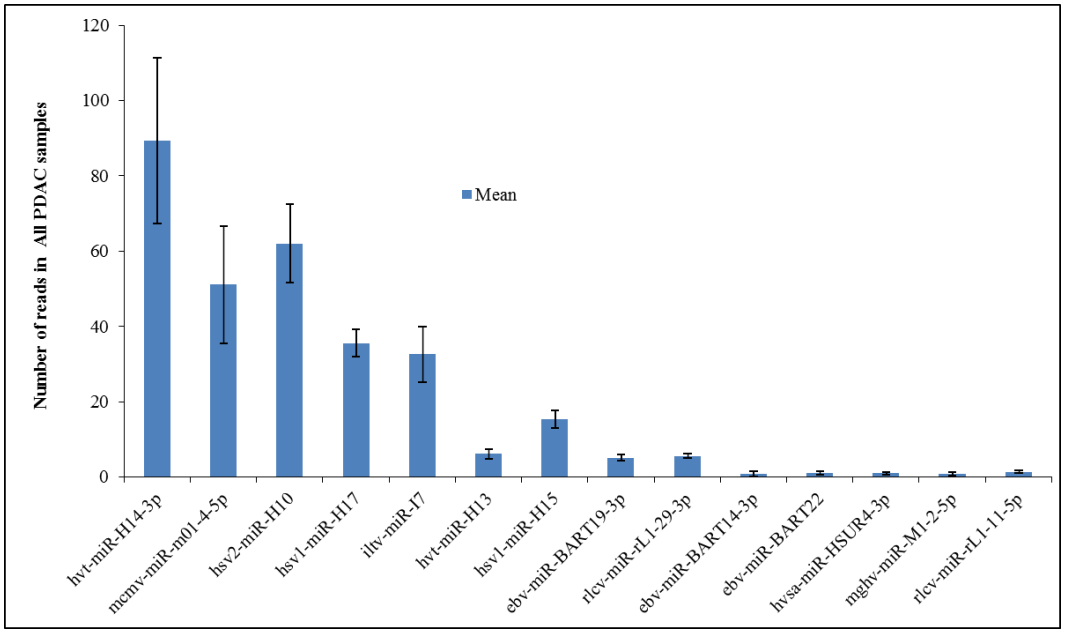


Figure 23: Number of reads of all viral miRNAs identified through digital miRNAome subtraction. The graph represents average number of miRNA reads from eight PDAC samples after digital subtraction. Hvt-miR-H14-3p is the top candidate for our study.

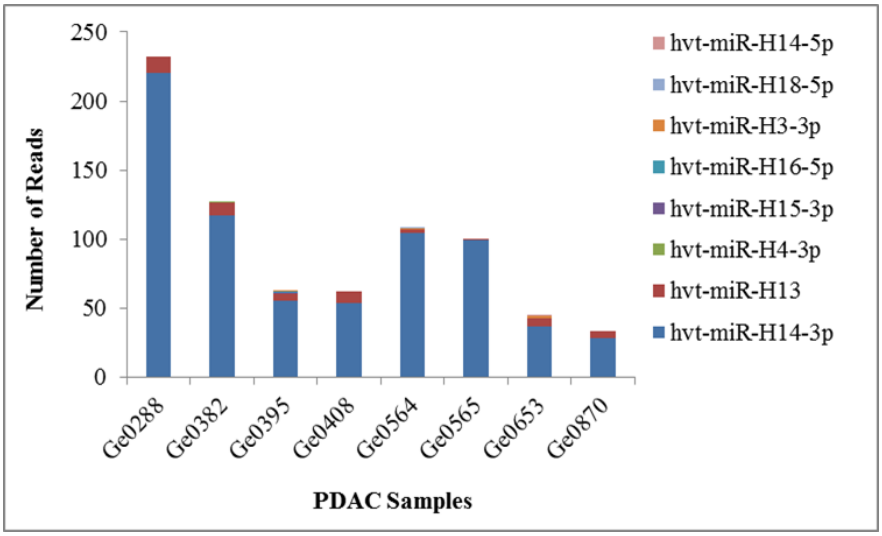


Figure 24: Number of reads of different miRNAs from Meleagrid herpesvirus-1 (herpesvirus of turkeys) identified through digital miRNAome subtraction. Hvt-miR-H14-3p is having highest number of reads in PDAC in comparison to other miRNAs from this virus. It also shares a 100% seed sequence homology with two human miRNAs upregulated in PDAC.

Hvt-miR-	H14-3p	H13	H4-3p	H15-3p	H16-5p	H3-3p	H18-5p	H14-5p
Ge0288	220	12	NA	NA	NA	NA	NA	NA
Ge0382	117	9	1	NA	NA	NA	NA	NA
Ge0395	55	5	NA	1	1	1	NA	NA
Ge0408	54	8	NA	NA	NA	NA	NA	NA
Ge0564	104	3	NA	NA	NA	1	1	NA
Ge0565	99	1	NA	NA	NA	NA	NA	NA
Ge0653	37	6	NA	NA	NA	1	NA	1
Ge0870	28	5	NA	NA	NA	NA	NA	NA

Table 4: The number of different miRNA reads from HVT identified in eight PDAC samples by digital miRNAome subtraction.

2.3. Hvt-miR-H14-3p is Upregulated in PDAC and CP

As mentioned earlier, hvt-mir-H14-3p resembles two human miRNAs from mir-221 family in sequence. There are many studies that provide strong evidence that hsa-miR-221 and 222 are highly upregulated in PDAC. Since both the miRNAs, from HVT and humans, share sequence homology which also extends to their seed sequence, we thought to analyze the expression levels of hvt-miR-H14-3p in PDAC and CP in comparison with healthy normal controls. We used a human miRNA, RNU44, as a reference gene for normalization of expression levels. Hvt-miR-H14-3p was significantly upregulated in tumors and inflammation condition in comparison with healthy controls. For the quantitative reverse transcription PCR (RT-qPCR), we used miRNA specific stem loop primers that converted hvt-miR-H14-3p or RNU44 into a cDNA sequence. The cDNA was further quantified using miRNA sequence specific hydrolysis probes by qPCR and digital PCR. Multiple negative controls were performed to ensure absence of false positive quantification.

2.3.1. Quantitative Real Time PCR Analysis

The hydrolysis probes for miRNA quantification from Life technologies use a sequence specific stem loop primer to convert target miRNA to cDNA from total RNA. Since the copy number of the viral miRNA is very low, there was significant standard deviation between technical triplicates of qPCR. Further, the quantification cycle (Cq) values were also very high for the viral miRNA. However, there was no amplification in multiple negative controls. Hence, the amplification observed in viral miRNA sample set was highly specific to the genes of interest. The hvt-miR-H14-3p was significantly upregulated in PDAC and CP than in healthy normal controls. Further, there was no significant difference in hvt-miR-H14-3p relative fold change between PDAC and CP. This suggests that viral miRNA has a role in the progression of PDAC from inflammation.

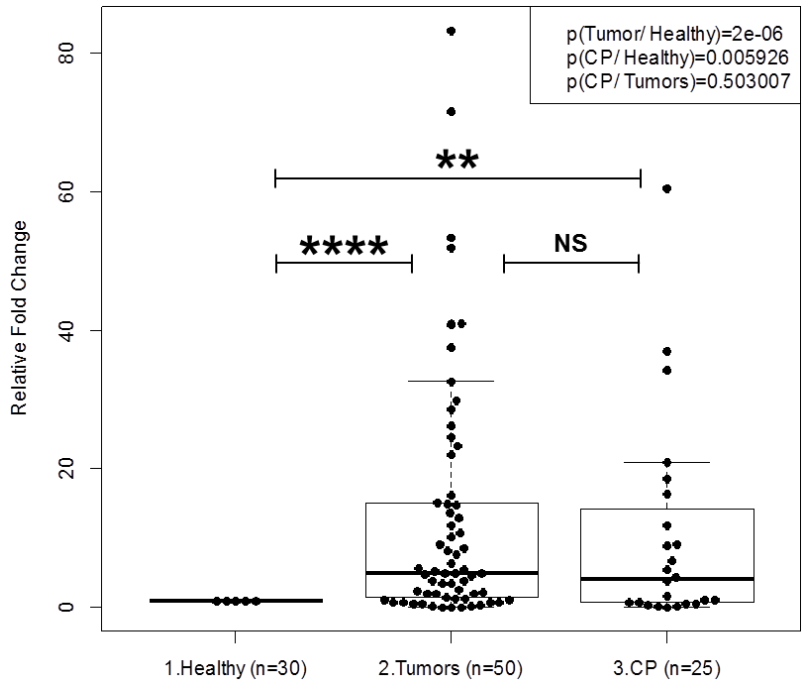


Figure 25: Relative fold change of hvt-miR-H14-3p between healthy normal controls and PDAC, and CP, normalized to controls. The numbers in parenthesis represent sample size. Significant upregulation of viral miRNA is observed between controls, and PDAC and CP. However, no significant relative fold change is seen between PDAC and CP.

2.3.2. Digital PCR Analysis

The viral miRNA is present in very low copy numbers across all samples used in this study. This made the quantification of hvt-miR-H14-3p theoretically challenging as the standard deviation between technical triplicates was high. Hence a more sensitive and robust quantification was employed to compare the viral miRNA at the level of copy number per microliter of sample using digital PCR. Two independent digital PCR platforms were used to validate the previous qPCR observation. The Bio-Rad platform generated approximately 14,000 to 18,000 droplets per sample and measured the copy number of target sequence per microliter of sample. The RainDance platform generated approximately 4 million droplets per sample and quantified the amplification of target gene per droplet. Both platforms emphasized the previous observation with qPCR that hvt-miR-H14-3p was significantly upregulated in PDAC than in healthy controls.

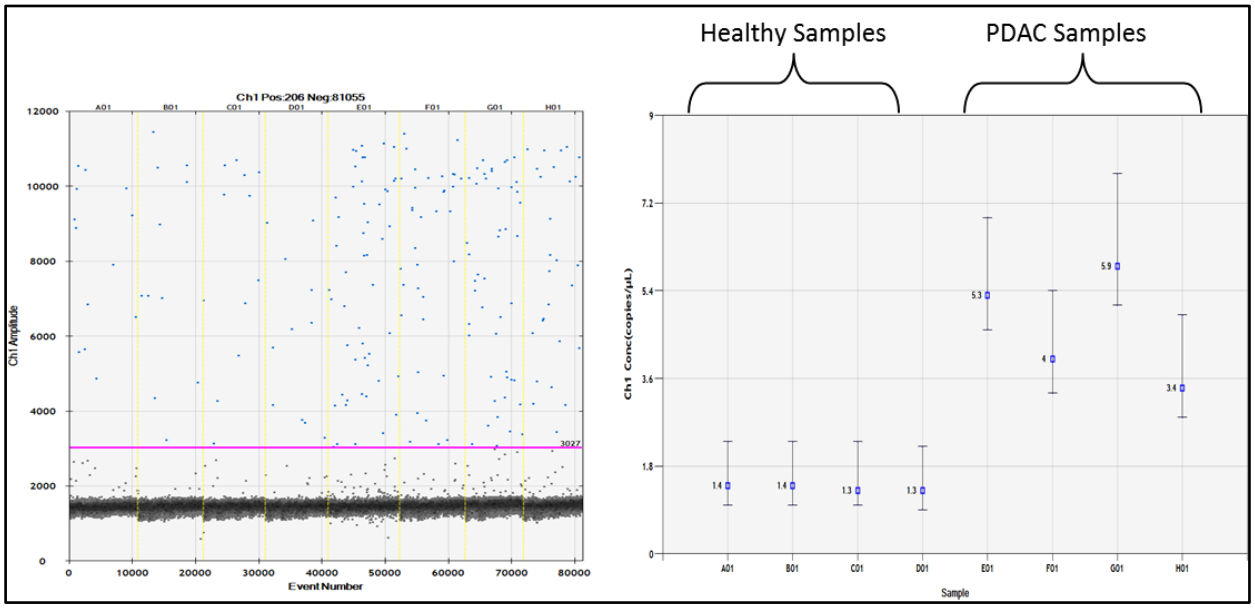


Figure 26: Copy number variation of viral miRNA between healthy controls and PDAC samples. The number of positive and negative events per well are shown on the right side part of the figure. The values in the chart area of the left part represent copy numbers of viral miRNA per microliter. The bars represent 95% CI.

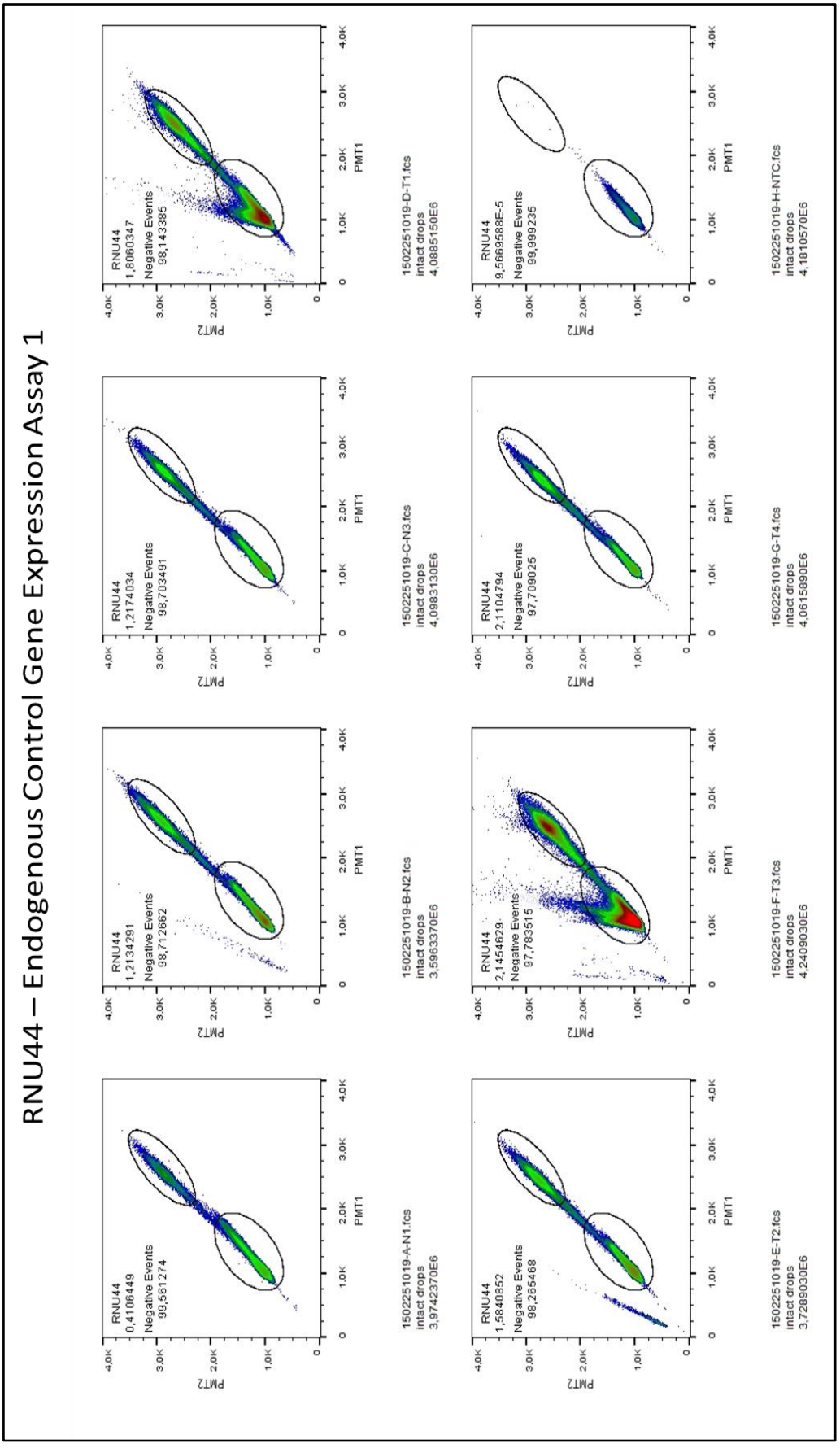


Figure 27A: Reference gene RainDance digital PCR amplification profile generated through assay number 1

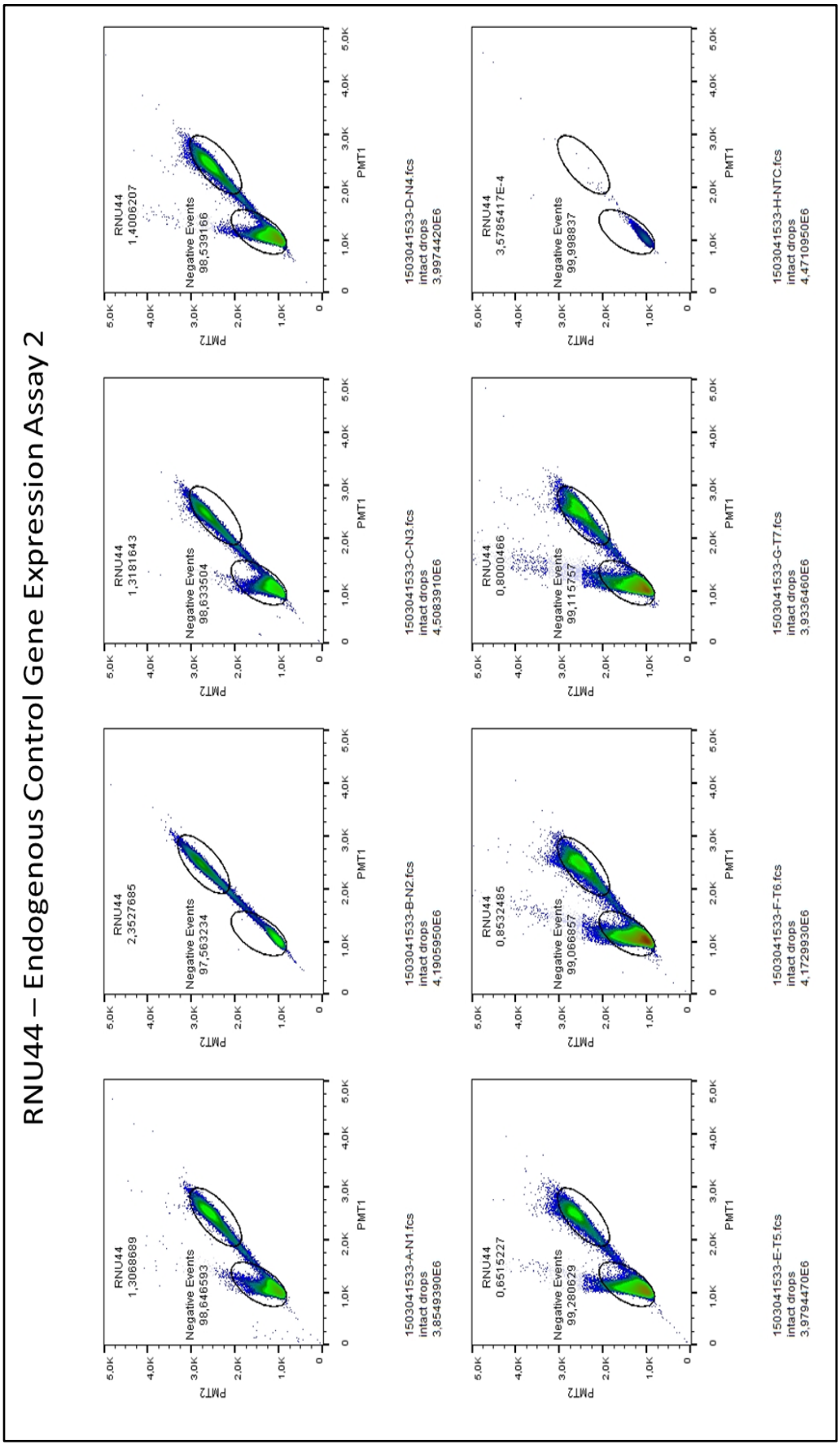


Figure 27B: Reference gene RainDance digital PCR amplification profile generated through assay number 2

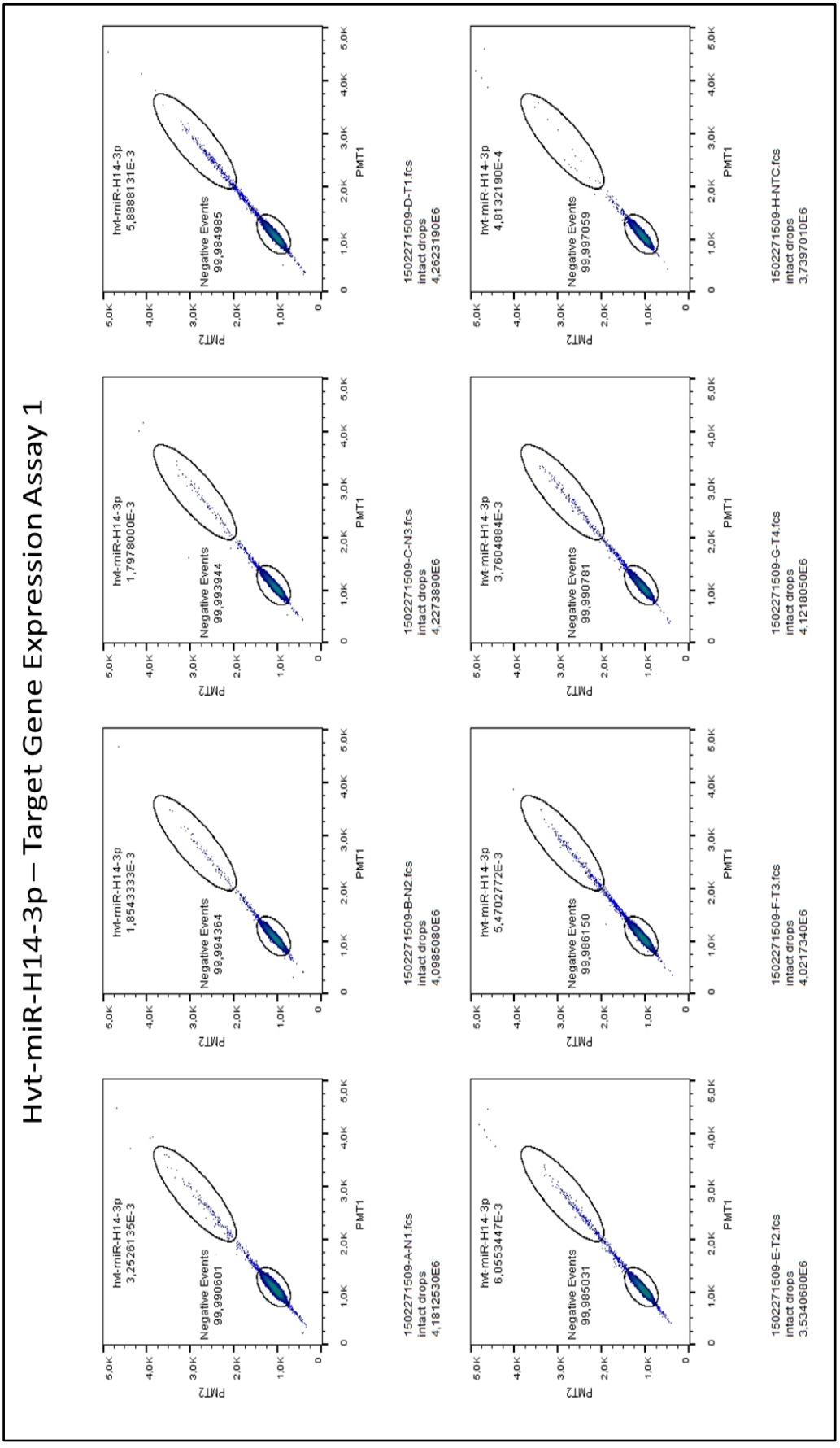


Figure 27C: Hvt-miR-H14-3p RainDance digital PCR amplification profile generated through assay number 1

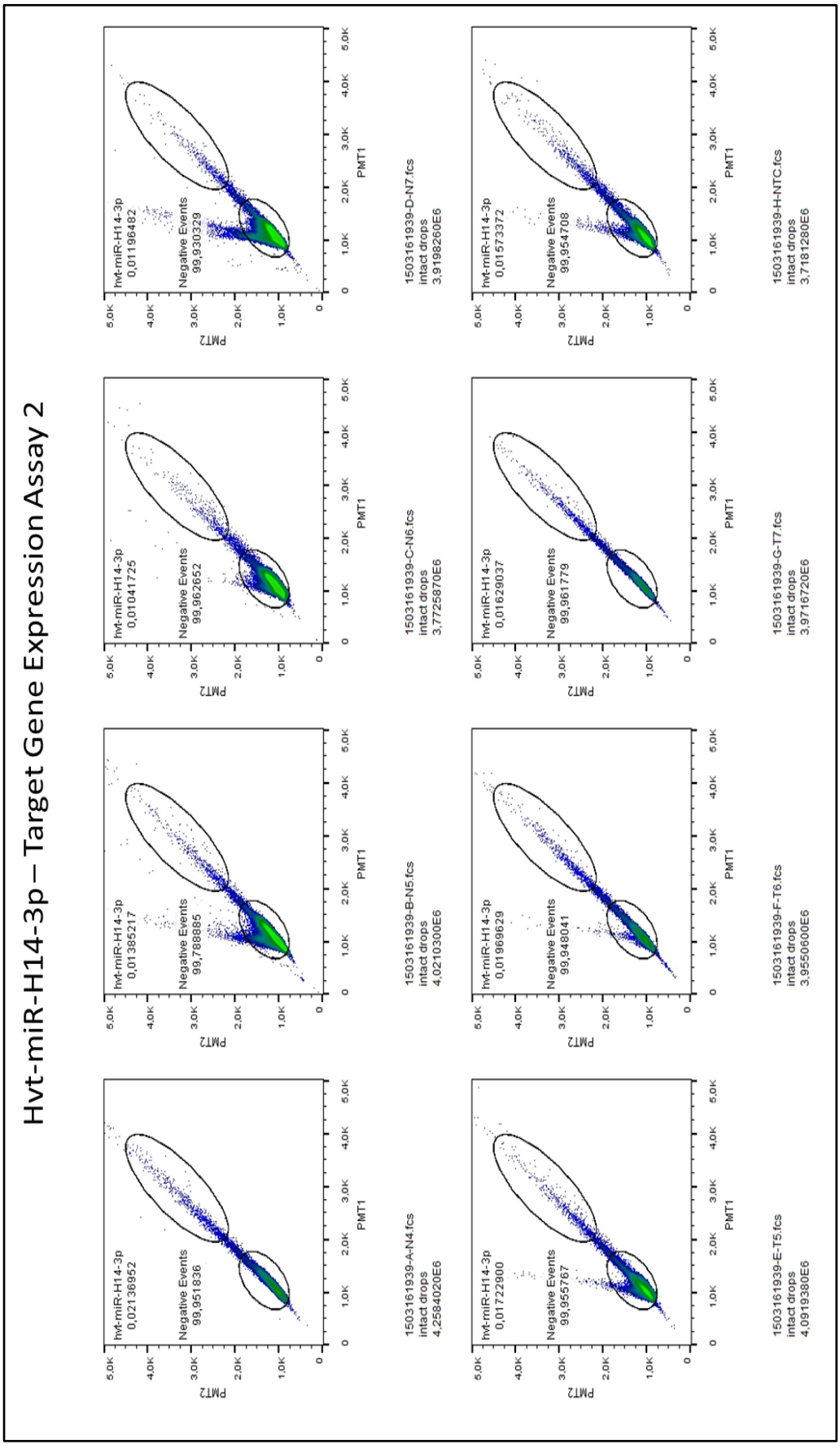


Figure 27D: Hvt-miR-H14-3p Rain Dance digital PCR amplification profile generated through assay number 2

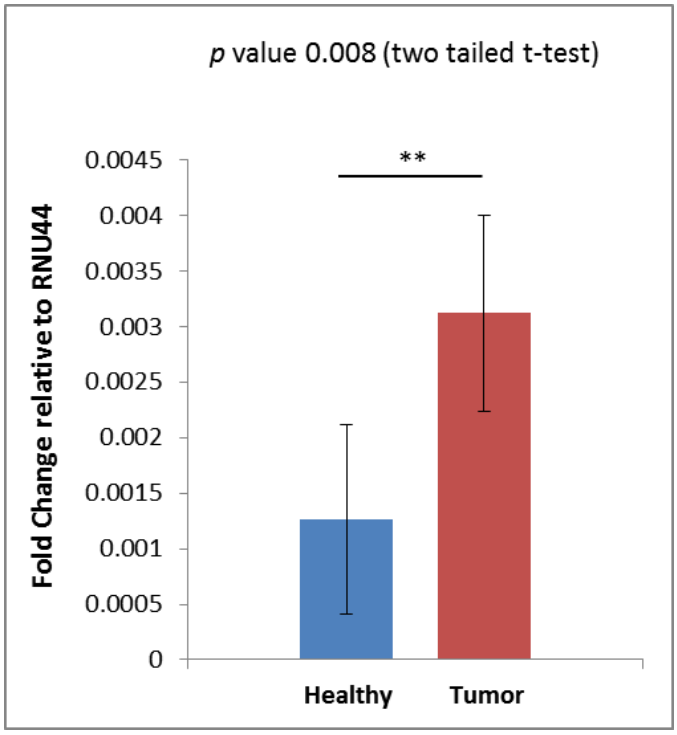


Figure 28: Relative quantification of hvt-miR-H14-3p between healthy controls and PDAC using RainDance digital PCR platform. The observation in qPCR experiments were further validated in this study. Viral miRNA showed a significant upregulation in PDAC than in healthy control.

2.4. Functional Analysis of Hvt-miR-H14-3p *in vitro*

The previous section validated the qPCR observation that hvt-miR-H14-3p is significantly upregulated in PDAC and CP. The following sections will provide functional analysis of hvt-miR-H14-3p performed *in vitro*. Over expression studies were performed in different pancreatic cancer cell lines. Messenger RNA target identification was also performed and the down regulation of the target upon overexpression of viral miRNA was validated in cell line system. These functional studies have strongly suggested the role of hvt-miR-H14-3p in invasion and metastasis of PDAC cell lines. To describe briefly, we transfected the cells with viral miRNA mimic and negative control. We used these populations of PDAC cell lines to study migration, invasion and proliferation simultaneously. The results of each of these assays are shown in the following sections.

2.4.1. Hvt-miR-H14-3p Overexpression in Metastatic PDAC Cell Lines

As mentioned in the previous section 1.2.1, the PDAC is a very aggressive disease. Local and distant organ metastasis is a significant concern for PDAC patients. Hence, overexpression of hvt-miR-H14-3p using second generation mimic systems was performed in order to investigate modulation of metastasis by viral miRNA. The figure below convincingly shows that overexpressing hvt-miR-H14-3p significantly increases the metastatic PDAC cell migration.

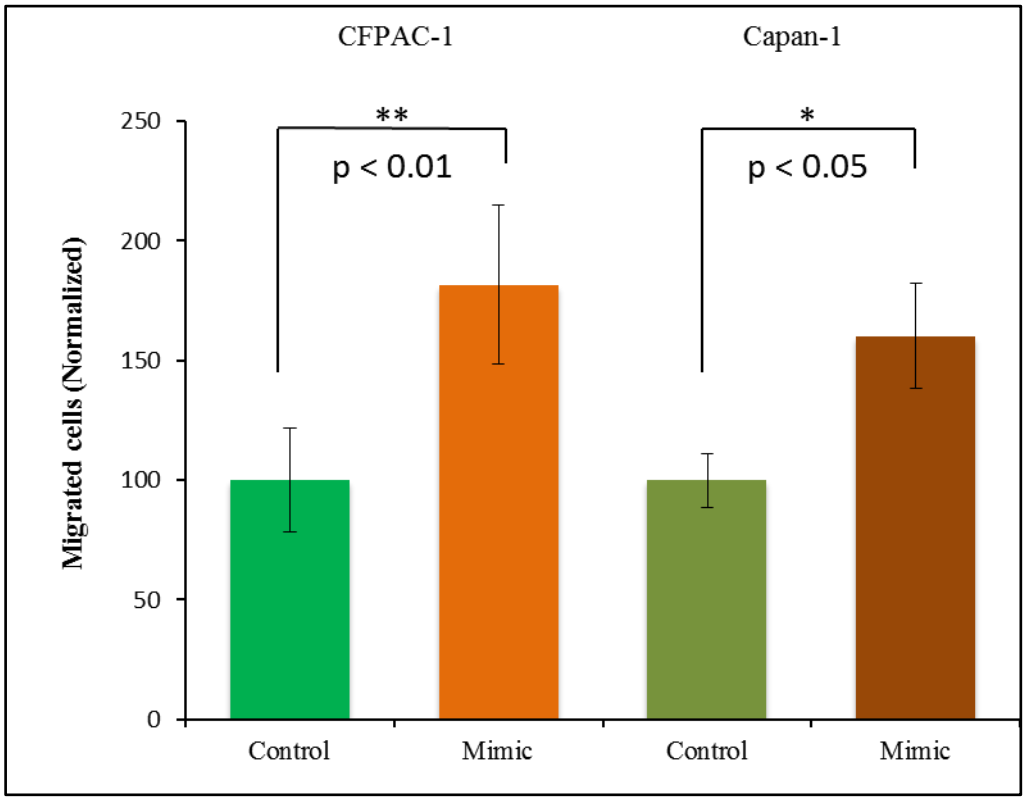


Figure 29: Hvt-miR-H14-3p significantly increases migration of metastatic PDAC cell lines. It is shown in this figure that CFPAC-1 and Capan-1 migrate respectively around 80 and 60 fold more than their controls. The results are represented as means and SD from three independent experiments.

2.4.2. Hvt-miR-H14-3p Overexpression in Non-Metastatic PDAC Cell Lines

In the previous section the overexpression of viral miRNA increased migration of metastatic PDAC cell. This observation was crucial as it suggests the role of hvt-miR-H14-3p in metastasis. Invasion is a very important phenotypic character of a metastatic cell. Hence to follow up on the results on migration, non-metastatic cells were overexpressed with hvt-miR-H14-3p and their invasion was studied in Matrigel chambers. The non-metastatic primary tumor cell lines showed significant increase in invasion in comparison with controls.

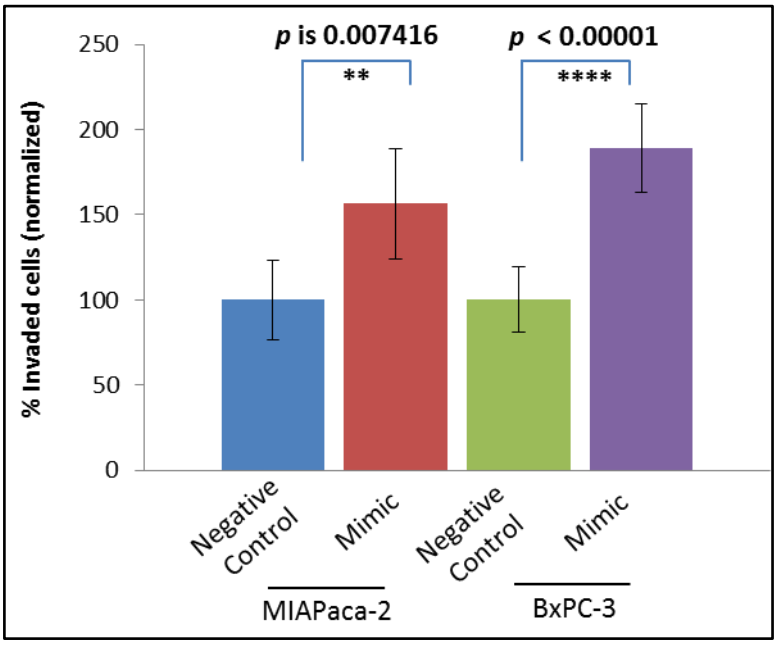


Figure 30: This graph shows the percentage of invaded cells upon overexpression with viral miRNA mimic and negative control. The invasion of BxPC-3 is more than MIAPaCa-2 upon viral miRNA overexpression. This effect is also in correlation with the inherently aggressive nature of the BxPC-3, in that this cell line shows an early invasion than MIAPaCa-2. The results are represented as means and SD from three independent experiments.

2.4.3. Hvt-miR-H14-3p Overexpression and PDAC Cell Proliferation

The previous sections emphasized that the overexpression of viral miRNA, hvt-miR-H14-3p, promoted significant increase in migration and invasion in metastatic and non-metastatic PDAC cell lines respectively. To further validate these interesting observations, we investigated the proliferation of cells overexpressed with viral miRNA. The migrated and invaded cells that were overexpressed with viral miRNA, interestingly, showed no change in proliferation. This further validates that the significant increase in migration and invasion of *viral miRNA overexpressed PDAC cells* happened not because of proliferation but due to the real phenotype change.

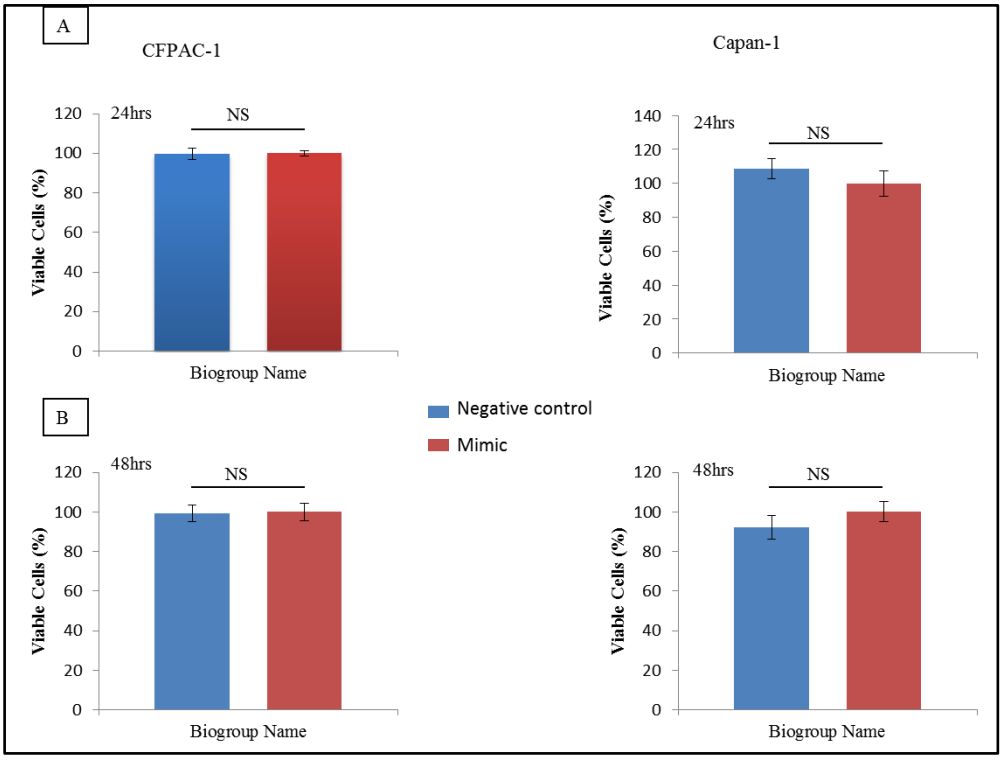


Figure 31: The graphs shown here represent percentage viability of the cells in assay. **A:** Metastatic cell lines, CFPAC-1 and Capan-1 show no change in viable cells. This assay was performed with CellTiter Glo that measures ATP. **B:** Since cells need at least 48hr of culture to express proteins, the second time point was at 48hr and in this assay amount of proteins in the cells is measured using SRB. The results are represented as means and SD from three independent experiments.

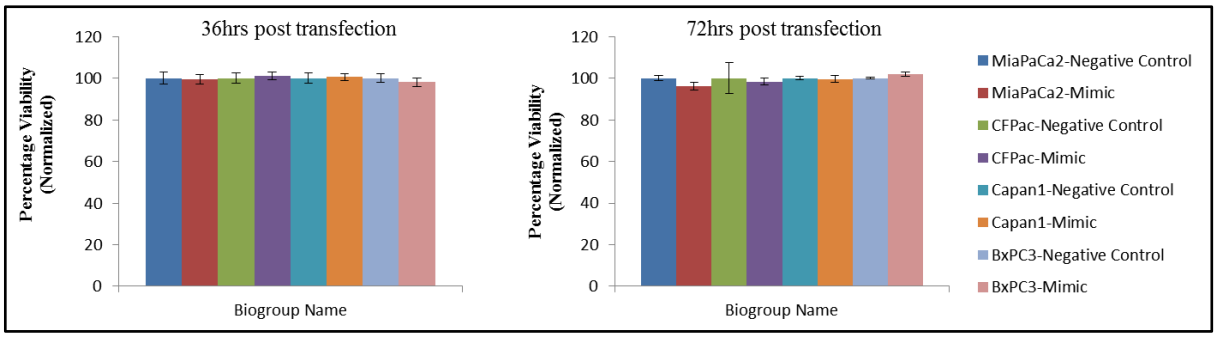


Figure 32: As a follow up of the previous experiment, all previous PDAC cell lines, in which increased migration and invasion is observed upon viral miRNA overexpression, were used to study proliferation change at two farther time points 36hr and 72hr. All of which did not show any significant change in proliferation at any time point. The results are represented as means and SD from three independent experiments.

2.4.4. Hvt-miR-H14-3p Down-regulates p27 Expression

As mentioned in the earlier sections, the viral miRNA, hvt-miR-H14-3p, shares a great degree of sequence homology to hsa-miR-221 and 222 including a 100% identical seed sequence, the primary hypothesis was that viral miRNA may target same genes as that of human miRNAs. To this end, we over expressed the viral miRNA and performed a Western blot to investigate the down regulation of the top candidate mRNA target of human miRNAs, p27. This experiment showed that the viral miRNA can target p27 and down regulate the levels of P27 in PDAC cell line system. This observation was validated also using a dual luciferase expression system. The dual luciferase detection works on the principle that miRNA binds to the 3'UTR of targets. If wild type p27 3'UTR binds to viral miRNA, it results in down regulation of luciferase signal. If the 3'UTR of p27 is mutated, it can no longer bind to the viral miRNA and luciferase signal remains similar to that of the constitutively expressing empty vector control. This assay validates the specificity between viral miRNA and its target, p27.

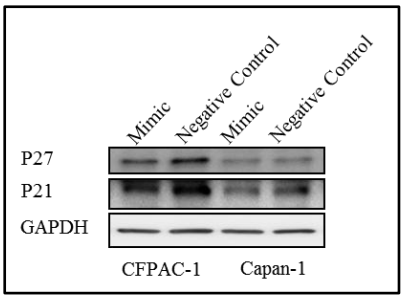


Figure 33: Western blot for P27 target, and P21 downstream affected protein, shows deregulation after over expressing viral miRNA than to negative controls in CFPAC-1 and Capan-1 PDAC cell lines. The result is representative of three independent experiments.

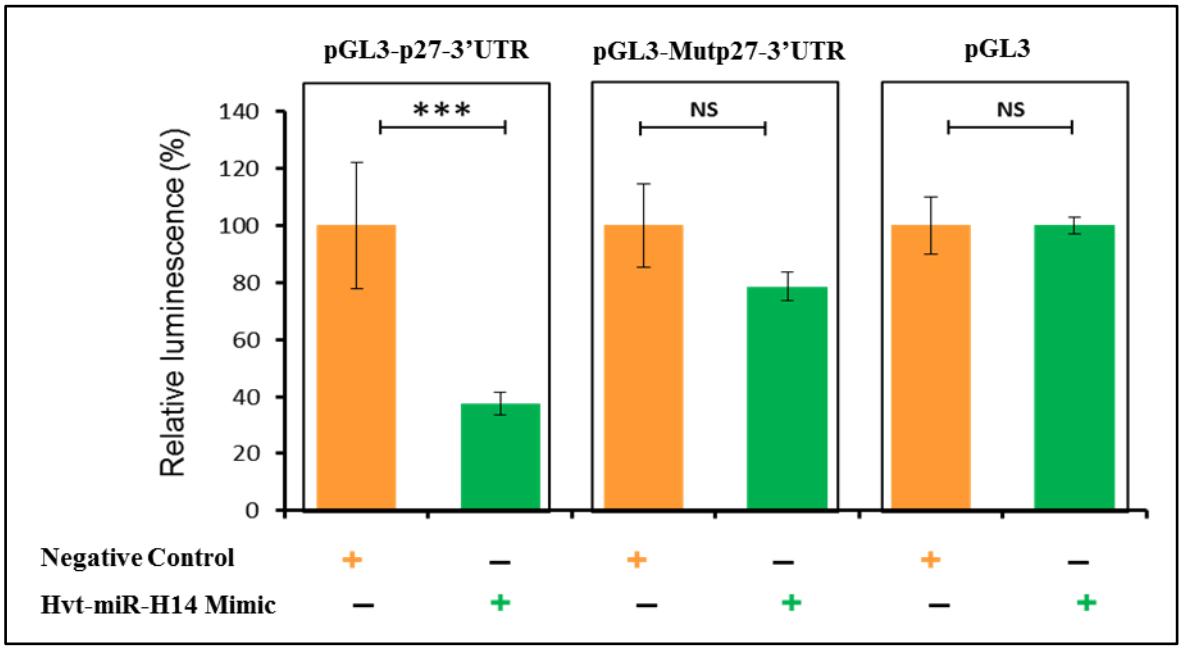


Figure 34: Dual luciferase assay shows that viral miRNA, hvt-miR-H14-3p, specifically targets 3'UTR of p27 while leaving mutated p27 3'UTR unaffected. Hvt-miR-H14-3p and negative control were co-transfected with the indicated luciferase constructs. Relative luciferase activity is the ratio between firefly luciferase and renilla control luciferase, adjusted to 100%. The results are represented as means and SD from three independent experiments.

2.5. *In vivo* Validation of hvt-miR-H14-3p Phenotype Using Xenograft Mice

In vitro culture systems used previously are helpful to investigate specific phenotypes of the viral miRNA. However, these systems are simplifications that does not account for the complexity or heterogeneity of PDAC as well as the tumor microenvironment. Hence, *in vivo* studies in nude mice with human PDAC cell lines (xenografts) were performed to understand the behavior of PDAC cells overexpressed with viral miRNA.

2.5.1. Generation of Stable PDAC Cell Lines Expressing hvt-miR-H14-3p

Lentivirus was used for stable integration of sequence of interest in PDAC cell lines. The desired population of lentivirus was produced by cloning hvt-mir-H14 or non-targeting scramble control sequence in lentivirus cloning vector. To monitor tumor progression, luciferase gene under the control of CMV promoter was also cloned in both populations. Using this pair of lentivirus population, two pairs of stable cell lines were produced with BxPC-3 and MIAPaCa-2. A mice PDAC cell line, K-8484, was also used for prospective studies in orthotopic tumor models. To validate our observations on the role of viral miRNA on invasion, we performed the simultaneous invasion and proliferation assays to check if the stable cell lines replicate these findings as that of *in vitro* systems. Similar to our previous observations, stable cell lines showed identical data that emphasizes that the significant increase in invasion of PDAC cells, overexpressed with viral miRNA, however, without significant change in proliferation.

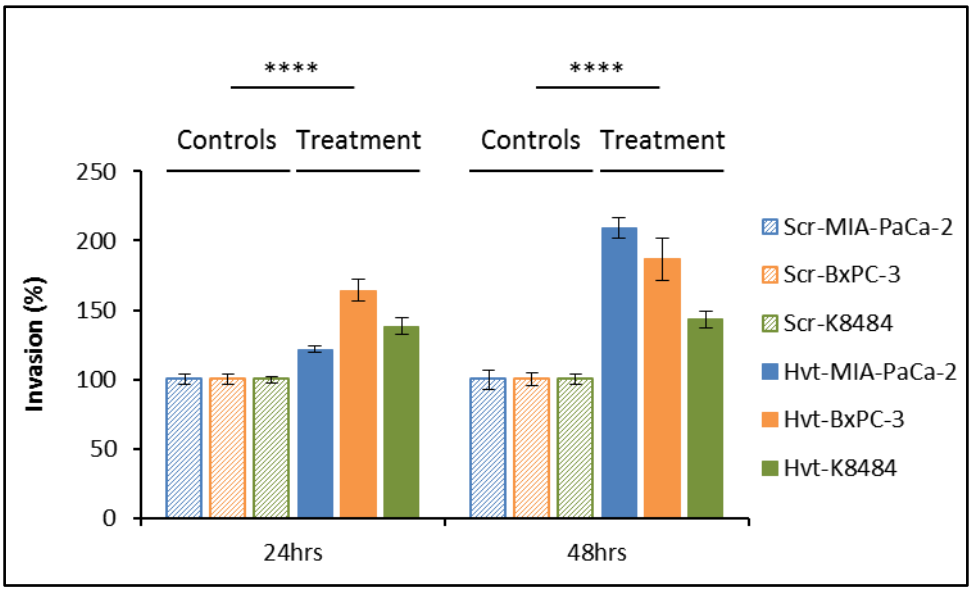


Figure 35: Stable cell lines show significant increase in invasion over time post transduction. Both non-metastatic cell lines invade significantly more than their controls. This graph shows percentage of invasion normalized to negative controls at 100%. The results are represented as means and SD from three independent experiments.

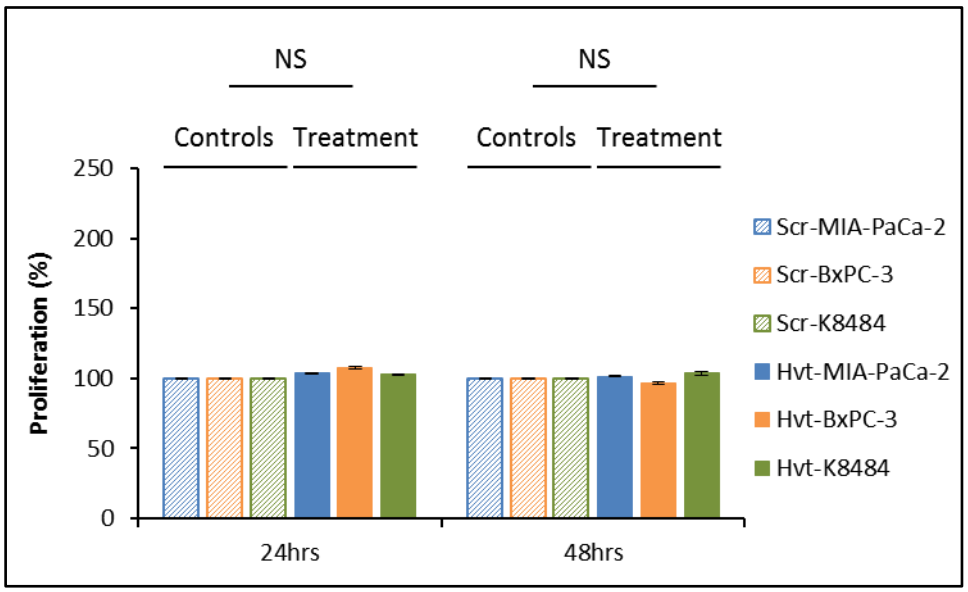


Figure 36: Stable cell lines show no significant change in proliferation over time post transduction. This graph shows percentage of viable cells normalized to negative controls at 100%. The results are represented as means and SD from three independent experiments.

2.5.2. Tumor Progression in Mice Injected with hvt-miR-H14-3p Stable Cells

The stable cell lines of BxPC-3 and MIAPaCa-2 expressing hvt-mir-H14 or non-targeting scramble control sequence were subcutaneously injected into the right flank of the mice on the abdomen in four different respective groups. Bioluminescence (BLI) and tumor volumes were measured after the tumor has reached a reached a significant size in mice. With our preliminary data, the cohort of mice with BxPC-3 stable cell lines expressing hvt-mir-H14 showed a significant increase in tumor progression within 3 weeks in comparison with its scramble control cohort also effecting survival of mice.

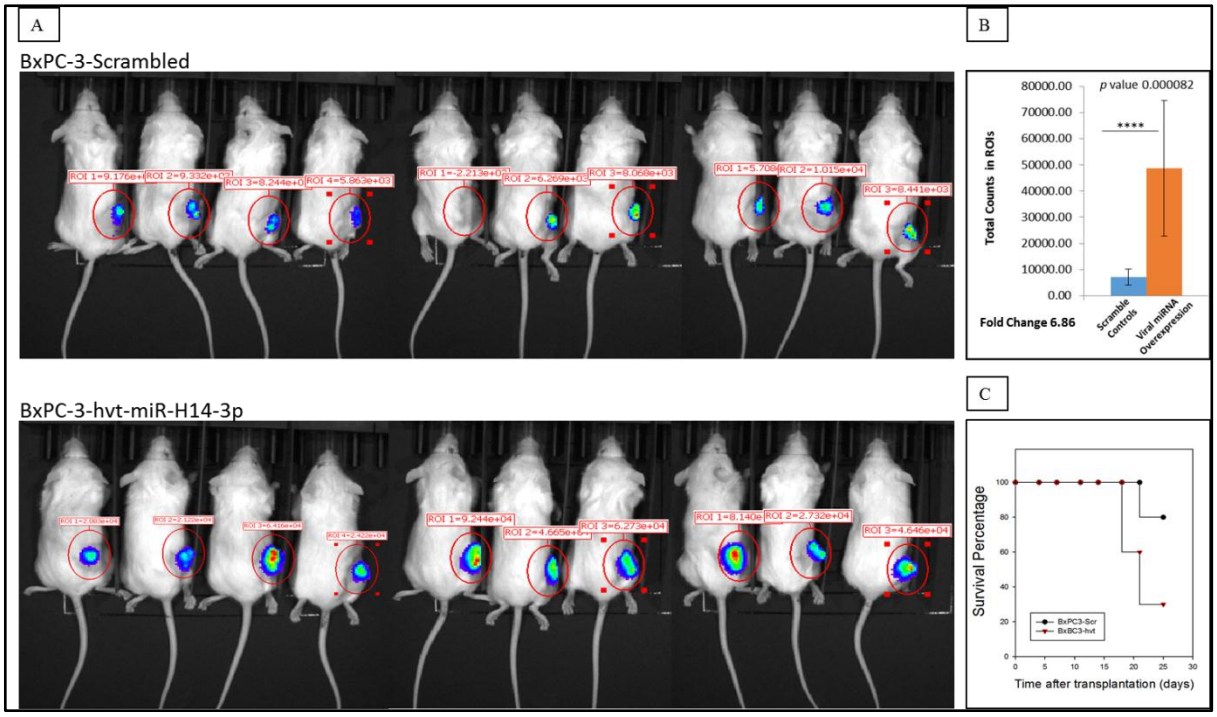


Figure 37: The group of mice expressing viral miRNA, hvt-miR-H14-3p, showed significant increase in tumor progression in comparison with the group of scramble control mice. **A:** The BLI measurements taken by anesthetizing mice and injecting them with luciferin. The luminescence corresponds to tumor progression. **B:** The bioluminescence is quantified in terms of total photons in the region of interest (ROI). The graph clearly shows that the viral miRNA significantly increases the tumor progression in mice expressing viral miRNA in their tumors in comparison with scramble controls. **C:** Post transplanting transduced cells, mice survival significantly reduced with increase in tumor progression in viral miRNA cohort than in control group.

3. DISCUSSION

In order to understand the importance of the findings in this project, it is essential to discuss the relation between genomic integration of viruses and carcinogenesis. In the following sections, a brief review of carcinogenic mechanisms mediated by sequence integration of three major tumor viruses, HBV, HPV, and EBV, are discussed to declutter various pathways that are affected, or could be involved in, tumor initiation or progression.

3.1. Integration of Pathogenic Virus Sequences in Cancer Genome

Various sequencing platforms have proved effective for discovering novel viruses and virus–tumor associations; however, these methods are at infancy to be helpful for comprehensive patient cohorts. Gene fusion events that happen recurrently include HPV in *RAD51B*, *ERBB2* and in the 13q22.1 intergenic region harboring the *LINC00393* lncRNA. Coevolution of host and viral genes provide important information about papillomavirus oncogenes function. Integrations are usually associated with altered gene expression (Peter, Stransky et al. 2010). Viral integration is usually the result of induction of local genomic instability. However, integration could also happen by pre-existing instability. Analysis of host transcriptome aberrations caused by oncoproteins can progress the identification and prioritization of oncogenes and pathways (Rozenblatt-Rosen, Deo et al. 2012). Abundant enrichment of low copy viral integration events from PDAC by RDA here enabled us to identify the virus and analyze host cell function in relation to the viral infection and viral gene expression at a previously undetectable scale and level of detail.

Circular viral DNA or tandem arranged genomes could also act as intermediates for integration. These integrations involve preferential activation of the cohesive end region analogous with retrovirus genome integration (Varmus and Swanstrom 1985). Virus DNA

can also integrate randomly along cancer genome, but random integration follows preferential loss of DNA in the cohesive end region. Tsurimoto et al. in 1987 proposed that three tandem arranged complete HBV genomes can be spread in integrated form in a hepatoblastoma cell lines (Tsurimoto, Fujiyama et al. 1987).

Viruses that are involved in carcinogenesis include several DNA viruses such as KSHV, MCV, EBV, HPV, HBV and SV40, and RNA viruses like HTLV-1 and, HCV. Incidence of cancer is also associated with integration of the viral genome for several oncogenic viruses, including HBV, HPV and MCV. Additionally, evidence is not clear that integration of EBV is a mechanism for carcinogenesis. The incidence of virus integration depends on the amount of DNA damage; because integration requires double strand breaks in both the host and virus DNA (Pett and Coleman 2007). Also, one of the major causes of DNA damage is oxidative stress, which can be triggered by viruses and enhanced by exogenous factors.

3.1.1. Strategies for Virus Sequence Integration

3.1.1.1. Hepatitis B Virus (HBV)

Hepatocellular carcinoma (HCC) is the one of the lethal form of cancer leading third in cause of cancer deaths worldwide (El-Serag and Rudolph 2007), and more than 70% of HCC is due to either HBV or HCV infection (Perz, Armstrong et al. 2006). HBV is a partially double-stranded DNA hepadnavirus with majority of retroviral features. The risk of HCC increases 5 to 15 fold in chronic HBV carriers compared with the healthy virus free individuals (El-Serag and Rudolph 2007), and there are evidences that show the integration of the HBV genome into the cellular genome is present in over 85%–90% of HBV-related HCCs. However, non-tumor tissues of patients with chronic HBV infections also carry the integrated form of HBV in their genome. Persistent HBV infection leads to integration of the HBV genome into hepatocytes thus leading to HCC (Brechot, Pourcel et al. 1980, Shafritz, Shouval et al. 1981).

HBV integration causes elevated expression of several cellular oncogenes, such as TERT, mixed-lineage leukemia 4 (MLL4). MLL4 is a part of the ASC-2 complex implicated in the p53 tumor suppressor pathway (Lee, Kim et al. 2009) and CCNE1 (encoding cyclin E1) (Sung, Zheng et al. 2012). HBV integration is further associated with early onset of HCC (Sung, Zheng et al. 2012). It was also show in previous studies that the integrated HBV sequences that encode HBx and/or truncated envelope pre-S2/S proteins are found in majority of HCC cases (Tsai and Chung 2010). As mentioned in the previous section, genetic instability would result, if events such as integration of the HBx sequence into host DNA happen in HCC. Mechanisms including the inactivation of the UV-damage DNA binding protein resulting in aberrant nucleotide excision repair, inactivation of p53-dependent apoptosis, repression of p53-mediated gene transcription (Lee and Rho 2000) and, cell cycle regulation, DNA repair and tumor suppression (Kremsdorf, Soussan et al. 2006). HBx also trans-activates several signaling pathways connected to carcinogenesis, including those mediated by protein kinase C, JAK/STAT and PI3K (Feitelson and Lee 2007, Feitelson, Reis et al. 2009). Further, HBx also promotes carcinogenesis by upregulating TGF-β expression in HCC tissue (Yoo, Ueda et al. 1996). TGF-β inhibits hepatocyte proliferation during liver regeneration (Nakamura, Tomita et al. 1985, Braun, Mead et al. 1988, Fausto and Mead 1989). During liver cirrhosis, TGF-β also stimulates extracellular matrix protein production by hepatocytes (Czaja, Weiner et al. 1989, Nakatsukasa, Nagy et al. 1990). A truncated envelope pre-S2/S protein that is frequently found in HCC samples originates from the deletion of the preS2 region of the S2/S protein during HBV integration (Tai, Suk et al. 2002). Pre-S2/S truncated product by trans-activates several cellular genes such as *c-myc*, *c-fos* and *c-Ha-ras* promoting malignant transformation (Schluter, Meyer et al. 1994, Luber, Arnold et al. 1996). Interestingly, the pre-S mutant large surface antigens are present in the endoplasmic reticulum. This strategy evades detection by the host immune system. This

protein can also trigger ER stress that induces oxidative DNA damage and thus promotes genomic instability (Wang, Huang et al. 2006). The cell cycle progression and proliferation of hepatocytes is induced by the pre-S, by upregulating COX-2 and cyclin A(Wang, Huang et al. 2006).

The integration rate of HBV DNA into the host genome is significantly increased in the presence of DSBs (Bill and Summers 2004). Integration of HBV into the human genome occurs at susceptible sites in the genome that are involved in regulation of cell signaling, proliferation and viability (Murakami, Saigo et al. 2005). Common targets of HBV integration include human cyclin A2 (Wang, Zindy et al. 1992), the PDGF receptor, calcium signaling-related genes, mixed lineage leukemia encoding genes, 60S ribosomal protein genes (Murakami, Saigo et al. 2005), human telomerase reverse transcriptase (hTERT) (Horikawa and Barrett 2001) and the retinoic acid receptor β (Yaginuma, Kobayashi et al. 1987). The next-generation sequencing data allowed not only to determine the integration sites, but also to specify DNA damage, and to identify the mutations that contribute to carcinogenesis (Barzon, Lavezzo et al. 2011, Li and Mao 2013). These studies have shown that HBV integrates into TERT, MLL4, CCNE1 and ANGPT1 (encoding angiopoietin 1) (Fujimoto, Totoki et al. 2012, Jiang, Jhunjhunwala et al. 2012, Sung, Zheng et al. 2012). They also identified mutations in CTNNB1 (encoding β-catenin), IRF2 (encoding interferon regulatory factor 2), TP53, ARID2 (subunit of the polybromo- and BRG1-associated factor chromatin remodeling complex (Yan, Cui et al. 2005), a tumor suppressor gene) (Li, Zhao et al. 2011) and ARID1A (encoding a component of the SWI/SNF chromatin remodeling complex) (Li, Zhao et al. 2011, Guichard, Amaddeo et al. 2012, Huang, Deng et al. 2012). ARID2 mutations are correlated with HCV-associated HCC (Li, Zhao et al. 2011), ARID1A is also involved in HCC invasion and metastasis (Huang, Deng et al. 2012). IRF2 mutations

result in hyperploidy and high genomic instability (Guichard, Amaddeo et al. 2012). TP53 silencing mutations result in an deregulated p53 pathway (Guichard, Amaddeo et al. 2012).

Altogether, the process of HBV integration induces significant genetic aberrations, including chromosomal deletions, translocations, transcripts fusion, DNA duplication, and finally genomic instability (Guerrero and Roberts 2005, Feitelson and Lee 2007). All these alterations lead to overexpression of oncogenes, deregulation of tumor suppressor genes and aberrant microRNA profile (Feitelson and Lee 2007).

3.1.1.2. Human Papillomaviruses (HPVs)

HPV is circular, double-stranded DNA virus. There are more than 100 different genotypes of HPV (Calleja-Macias, Kalantari et al. 2005). Few of them are classified as either high risk or low risk casual agents for cervical cancer (Parkin and Bray 2006, Stewart and Wild 2014). High-risk HPVs integrate into the host genome causing cancer lesions (Arends, Buckley et al. 1998, Scheurer, Tortolero-Luna et al. 2005). More than 90% of cancerous lesions in the uterine cervix are HPV DNA positive (zur Hausen 1991, Clifford, Rana et al. 2005). 90% of all cervical cancer cases are attributable to HPV16, HPV18, HPV31 and HPV33. Majority of invasive cancers contain integrated HPV-16. This suggests that integration is an important for invasion and tumor progression (Hopman, Smedts et al. 2004, Guo, Sneige et al. 2007). The generation of reactive oxygen species and reactive nitrogen species during lesion formation, creates double stranded breaks in both the viral and host DNA. This allows HPV to integrate into the human genome. HPV integration sites are distributed randomly throughout the host genome, without a single region predominating (Wentzensen, Vinokurova et al. 2004). High-risk HPV integration occurs within or adjacent to host oncogenes (Wentzensen, Ridder et al. 2002, Ferber, Thorland et al. 2003, Thorland, Myers et al. 2003, Wentzensen, Vinokurova et al. 2004). This integration enables E6 and E7 (oncogenes) overexpression, which promotes

malignant transformation. HPV-16 integration disrupts E2 gene since the E2 ORF is a preferential site of integration. This results in significant increase of E6 and E7 oncoproteins expression levels as E2 viral regulatory protein losses control over negative feedback (von Knebel Doeberitz 2002, Pett and Coleman 2007). However, Low-risk types are usually present as episomal fraction. The levels of E6 and E7 oncogene expression are low in episomal HPV16. Further, during latency, transcriptional activity of integrated HPV-16 DNA is down regulated by E2 proteins from the episomal form (Bechtold, Beard et al. 2003, Herdman, Pett et al. 2006, Pett, Herdman et al. 2006, Pett and Coleman 2007, Hafner, Driesch et al. 2008). The cell proliferation, anti-apoptotic effects, and inflammation, combined with low-level expression of the E6 and E7 oncogenes encoded by the episomal HPV, contribute to tumor lesion progression (Williams, Filippova et al. 2011).

E6 and E7 expression affects a number of signal transduction pathways. E6 is involved in the rapid degradation of p53 and activates hTERT, while E7 inactivates pRB (zur Hausen 2000). Further, E6 binds to and inhibits interferon regulatory factor-3 transcriptional activity. This strategy helps HPV to evade antiviral response from host immune system (Ronco, Karpova et al. 1998). E7 mediate the activation of cyclin E and cyclin A required for malignant transformation (Zerfass, Schulze et al. 1995). Interactions between the E6 and E7 proteins also leads to cellular immortalization (Band, Zajchowski et al. 1990, Munger and Phelps 1993).

Altogether, the elevated expression of oncoproteins from integrated forms of HPV deregulates cellular proliferation, blocks apoptosis and increases genomic instability, all of which contribute to cellular transformation.

3.1.1.3. Epstein-Barr Virus (EBV)

EBV is a double-stranded DNA herpesvirus that is majorly associated with Burkitt's lymphoma (Epstein, Achong et al. 1964). Burkitt's lymphoma has three clinical variants namely, the endemic, sporadic, and immunodeficiency associated variants. EBV is detected in majority of endemic Burkitt's lymphoma (zur Hausen, Schulte-Holthausen et al. 1970). In contrast, EBV is rarely detected in the sporadic Burkitt's lymphoma (Xing and Kieff 2007). EBV-Burkitt's lymphoma is common in individuals with low levels of efficient T-cells (Kieff and Rickinson 2007).

Chromosomal translocations are common in Burkitt's lymphoma. These translocations place MYC oncogene under the control of the Ig heavy chain or one of the light-chain loci inducing MYC deregulation. This leads to the Burkitt's lymphoma (Polack, Hortnagel et al. 1996, Kovalchuk, Qi et al. 2000, Li, Van Calcar et al. 2003). EBV also displays transformative abilities. EBV induces B-cell transformation by dergulating cellular gene transcription and upregulating cell-signaling pathways (Young and Rickinson 2004). Like HPV low-risk strains, EBV usually persists in an episomal state with copies of circular DNA. In addition, EBV integration into host genome is rare and does not contribute to Burkitt's lymphoma. However, in several Burkitt's lymphoma cell lines, such as IB4, BL-36, BL-60 and BL-137 EBV integration has been evidenced (Matsuo, Heller et al. 1984, Delecluse, Bartnizke et al. 1993, Popescu, Chen et al. 1993, Wolf, Pawlita et al. 1993, Jox, Rohen et al. 1997). Integrated, episomal and linear copies of EBV DNA can coexist in Burkitt's lymphoma cells (Delecluse, Bartnizke et al. 1993). EBV integration in the host genome induces chromatin instability in the host cell genome (Jox, Rohen et al. 1997). This genome instability can in turn result in the loss of host tumor suppressor genes, such as BACH2 promoting carcinogenesis (Takakuwa, Luo et al. 2004).

3.2. Integration of Herpes Virus of Turkeys sequences in PDAC Genome

NGS-based methods, as the one employed in this thesis work, provide a very efficient method to map viral integration sites. Half of the mapped HPV16 integration sites directly target human cellular genes (Xu, Chotewutmontri et al. 2013). These studies suggest that, in some cancers, the insertional mutagenesis of the host genome plays a crucial role in tumorigenesis (Pett and Coleman 2007). However, many studies using NGS focused on identifying viral genotype and viral load, rather than on identifying integration sites (Conway, Chalkley et al. 2012, Meiring, Salimo et al. 2012). Hence our study presents a unique scenario of identifying both, the virus – MeSHV-1 or HVT and its integration sites, in pancreatic cancer.

HVT was first isolated in 1969 and is antigenically related to Marek's disease virus (MDV) (Gibbs, Nazerian et al. 1984). HVT is ubiquitous and non-pathogenic for chickens. Hence, HVT is routinely used in the poultry industry to vaccinate chickens against MDV. HVT infections are persistent in chickens hence provide long-lasting immunity (Purchase, Okazaki et al. 1972, Witter and Solomon 1972, Lee 1980, Fabricant, Calnek et al. 1982, Calnek and Witter 1991). The epidemiological features of natural HVT infection in turkeys include, early appearance of infection, rapid spread throughout the flock and as mentioned earlier, its persistence. Data from previous studies indicate that both HVT and the virulent MDV share many epidemiological features (Witter and Solomon 1971). These studies suggested that the rapid spread of HVT infection within turkey flocks is by horizontal transmission mediated by environmental factors such as air and dirty isolation cages acting as natural reservoirs for the virus. HVT can also transmit similar to MDV, in that, it can form into enveloped viral particles on skin of turkey flocks (Witter and Solomon 1971). Hence, the natural process in chickens such as shedding can transmit HVT. However, pathogenicity of HVT with regard to tumor induction is not reported to date, at least in humans.

4. CONCLUSION

In the present study, we observed the first clues on HVT sequence integration events in PDAC and further validation is being carried out on these findings. However, we identified multiple miRNAs from HVT through NGS analysis in PDAC patients. One of the miRNAs, hvt-miR-H14-3p, is also shown to be upregulated both, in PDAC and CP patients by quantitative real time PCR. Functional analysis of HVT miRNA later showed that it promotes tumor invasion and progression both, *in vitro* in PDAC cell lines, and *in vivo* in nude mice. Interestingly, this effect of invasion and migration on PDAC cell lines is accompanied by without significant change in proliferation. These observations suggest a plausible role of HVT in the aggressive nature of pancreatic cancer.

Hence this study represents the first documented evidence of HVT pathogenicity with regard to pancreatic cancer in humans. Furthermore, subtractive hybridization followed by NGS analysis showed that a putative uncharacterized protein highly specific to HVT is involved in high number of integration events in PDAC genome (table 4). Further analysis of this region showed that the DP3 and DP4 reads were highly specific to this region across the databases. These findings suggest a strong role of the HVT075 gene into PDAC genome integration and tumor progression. According to our present working hypothesis, this gene might have similar implications as that of BARF1 from EBV or Pre-S2/S truncated protein of HBV or E6/E7 of HPV. We are also working to identify very specific cellular mRNA targets for the HVT miRNA. We have already showed that hvt-miR-H-14-3p, like its homologous human miRNA has-miR-221, targets cellular p27. Thus the knowledge of HVT mediated carcinogenesis, and the network of pathways involved in the transition from initial infection to the pancreatic cancer provide clues on prophylactic and therapeutic strategies, which may ultimately reduce the risk of HVT-mediated PDAC.

5. MATERIALS AND METHODS

5.1. Materials

Equipment

Name	Manufacturer
Centrifuge 5415D	Eppendorf, Germany
Centrifuge 5810R	Eppendorf, Germany
NanoDrop ND 1000 Spectrophotometer	Thermo Scientific, USA
PCR Thermocycler PTC200 MJ Research	BioRad, USA
pH-Meter MP 230 Mettler Toledo	Mettler Toledo, Germany

Lab ware

Name	Manufacturer
96-well culture plate	Greiner Bio-One, Germany
Cell culture microplate,96 well, F-bottom	Greiner Bio-One, Germany
Cell culture Petri dishes 96 x 20 mm	Corning Life Science, USA
Eppendorf Safe-Lock tubes 1.5 mL and 2 mL	Eppendorf, Germany
Falcon 15 mL and 50 mL	Greiner Bio-One, Germany
Latex gloves	Latex, Blossom Mexpo, USA
Lazy-L-Spreaders	Sigma-Aldrich, USA
Microtiter cover film	Nunc, Germany
Nitril gloves	Nitril, Microflex, Austria
Parafilm PM 996	Fisher Scientific, USA
Microplates 96 and 384 well for PCR	Greiner Bio-One, Germany
Sterile filter 500 mL	Nalgene, USA

Chemicals

Name	Manufacturer
1,2-Bis(dimethylamino)ethane (TEMED)	Roth, Germany
2-Ethanesulfonic acid (HEPES)	Roth, Germany
Acrylamide/Bisacrylamide	Roth, Germany
Agar	Roth, Germany
Agarose	Invitrogen, USA
Ammoniumpersulfate (APS)	Roth, Germany

Ampicillin (sodium salt)	Genaxxon, Germany
BSA (bovine serum albumin)	Sigma-Aldrich, Germany
Disodium phosphate (Na2HPO4)	Roth, Germany
DL-Dithiothreitol (DTT)	Sigma-Aldrich, Germany
Ethanol, absolute	Sigma-Aldrich, Germany
Ethidium bromide	AppliChem, Germany
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich, Germany
O' GeneRulerTM 100 bp DNA marker MBI	Fermentas, St. Leon-Rot
Glycerol	Sigma-Aldrich, Germany
Imidazole	Roth, Germany
Isopropyl-β-D-thiogalactopyranoside (IPTG)	Roth, Germany
Kanamycin	Genaxxon, Germany
Methanol	VWR, Germany
Monosodium phosphate (NaH2PO4)	Roth, Germany
Oligonucleotides	Biomers, Germany
Oligonucleotides	Sigma-Aldrich, Germany
Protein-Marker: Broad Range	NEB, Germany
QIAGEN's 10xPCR buffer	Qiagen, Germany
Sodium chloride (NaCl)	Sigma-Aldrich, Germany
Taq DNA Polymerase	Qiagen, Germany
Tris-base	Sigma-Aldrich, Germany
Tris-HCl	Sigma-Aldrich, Germany
Triton X 100	Gerbu, Germany
TRIzol® Reagent	Invitrogen, Germany
Tryptone/Peptone	Roth, Karlsruhe, Germany
Tween 20	Sigma-Aldrich, Germany
Yeast extract	Gerbu Biotechnik, Germany

Kits

Name	Manufacturer
Qiagen QIAprep Spin Miniprep Kit	Qiagen, Germany
Qiagen QIAquick PCR Purification Kit	Qiagen, Germany
Dual-Luciferase® Reporter Assay System	Promega, Germany
5 Prime PCR Purification Kit	5 Prime, Germany
PureLink PCR purification Kit	Life Technologies

Buffers and media

Name	Composition
0.5 M EDTA pH 8.0	Dissolve 186.1 g Na2EDTA·2H2O in 800 mL dH2O.adjust pH to 8.0 with NaOH (~20 g of NaOH pellets). EDTA will dissolve at pH 8.0. Adjust volume to 1 liter with dH2O. Sterilize by autoclaving and store at room temperature.
1 M Ethanolamine	60.5 ml ethanolamine in 1000 mL
1 M NaH2PO4 (monobasic)	138 g NaH2PO4·H2O in sufficient H2O to make a final volume of 1 L
1 M Tris-HCl pH 6–8	12.1 g Tris base in 100 mL H2O, adjust pH with concentrated HCl
10 M NaOH	40 g NaOH in 100 ml H2O
1M HEPES-NaOH pH 7.5	Dissolve 238.3 g HEPES in 1 L H2O. Use NaOH pellets to adjust pH to 7.5. Start with about 5.5 g NaOH pellets
1M Na2HPO4 (dibasic)	142 g of Na2HPO4 in sufficient H2O to make final volume 1 L
6xGel-loading buffer	25 mg bromophenol blue and 4 g sucrose, make up volume to 10 mL with dH2O, store at 4°C
Ethidium bromide solution	0.5 µg/mL final concentration
Laemmli	buffer 30.1 g Tris base, 144.2 g glycine, 50 mL SDS (20%), add 1 L dH2O
LB Agar	LB-medium + 1.5% (w/v) agar
LB Medium (1 Liter)	10 g Tryptone/Pepton, 5 g yeast extract, 10 g NaCl, pH 7.2
PBS 10x (1 Liter)	80 g NaCl, 2 g KCl, 26.8 g Na2HPO4, 2.4 g KH2PO4, pH 7.4
PBST 1x (1 Liter)	1x PBS/0.05% Tween-20
TBS 10x (1 Liter)	50 mM Tris, 150 mM NaCl with HCl, pH 7.5
TBST	1x TBS/0.05% Tween-20
Transfer buffer	150 mM glycine, 25 mM Tris base, 20% ethanol
3x EE hybridization buffer	30 mM EPPS (pH 8.0), 3 mM EDTA

5.2. Methods

5.2.1. Representational Difference Analysis

The protocol for RDA was adapted from (Frohme and Hoheisel 2006). An overview of the procedure was presented in a schematic in the figure 38. Pool of 10 genomic samples from normal healthy individuals is use as a driver and a pool of 10 genomic DNA samples from PDAC patients is used as tester. Briefly, DNA from the two pools was digested with *DpnII* and then ligated with R-Bam-12 and R-Bam-24 oligomers for PCR-amplification separately. To establish a similar library of driver and tester representations, many PCRs with different number of cycles ranging from 20 to 30 cycles of amplification were performed using R-Bam-24 as primer. 25 cycles of amplification generated a uniform representation for driver and tester. As the driver is used in relatively higher amounts than the tester, 96 PCRs of driver and 8 PCRs of tester were done with 25 cycles of amplification. All the reactions of driver and tester were pooled separately and DNA was isolated. Both the driver and tester DNA were digested again by *DpnII* and purified to remove the R-oligomers. New J-Bgl-12 and J-Bgl-24 oligomers were ligated to tester DNA alone. Now, driver and tester were mixed in a 100:1 ratio and denatured for a first round of subtractive hybridization. A difference product (DP1) was obtained and prepared for a second round of RDA (driver:tester = 800:1) by changing the J-adaptors to N-Bgl-12 and N-Bgl-24 oligomers. The DP2 so obtained was again prepared for a third round of RDA (driver:tester = 400,000:1) to achieve DP3. Similarly, DP3 was prepared for fourth round of RDA with a driver:tester ratio 8,000,000:1. The difference products, DP3 and DP4 were digested with dsDNA fragmentase and paired end sequenced on Illumina platform. For pilot studies with human placental DNA and bacterial plasmids, the DP3 was cut out from the gel; the DNA was extracted and cloned into vectors using TOPO TA cloning system. Recombinant clones were selected for single-read sequencing.

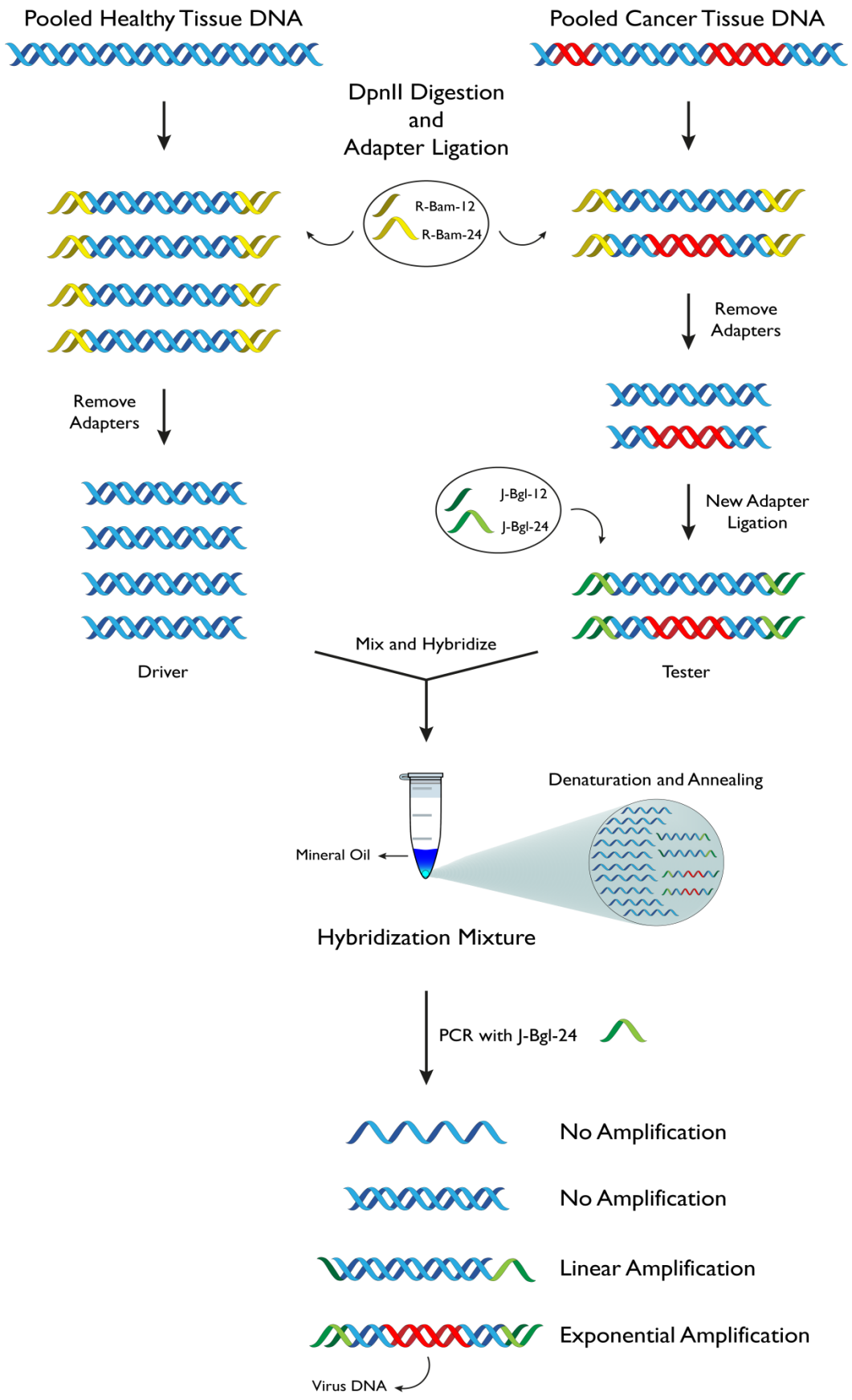


Figure 38: Schematic of the representation of RDA procedure.

5.2.2. DPs and MicroRNA Sequencing

The DPs of RDA, and total RNA samples from PDAC patients, were submitted to our core facility at German Cancer Research Center (DKFZ), Heidelberg. The samples have are quality checked by the core facility using Bioanalyzer high sensitivity assays. DPs were sequenced using Illumina paired-end sequencing platform. The miRNAs were sequenced using Illumina MiSeq platform. The quality control data is provided in the digital supplementary information.

5.2.3. NGS Analysis of Sequencing Data

The data from DPs was cleaned to remove sequencing adapters. The filtered reads were aligned with human reference genome from NCBI RefSeq GRCh37. The sequences that were not aligned with human genome were selected and aligned with MeHV-1 complete genome from NCBI. The paired end sequencing reads that matched with the MeHV-1 genome were selected and sorted according to the p value, percentage identity, overlap length, mismatches, and gaps. The regions of the MeHV-1 where there are majority read alignments were compiled to map the reads against virus genome. The MeHV-1 specific set of genes were selected and checked for alignment with the reads from DPs.

For miRNAs analysis, the MiSeq reads from the PDAC patient samples were first cleaned to remove sequencing adapters using sRNA mapper by turning of the mapping function. The success of adapter removal was checked by aligning the reads with human genome from NCBI RefSeq database. Reads with an alignment percentage greater than 90% are considered to be filtered from adapters. The clean reads are then aligned with viral genome from NCBI RefSeq Viral (version: 14.09.2013) database allowing 1 mismatch and 0 gaps. The reads that are aligned with viral genome are selected and filtered for human sequences by aligning them with human genome from NCBI RefSeq human genome database. The reads that did not

align with human database but aligned with viral database aligned earlier were considered to be uniquely viral in origin or virus only sequences. These reads are then aligned with all available miRNAs from miRBase 21. The top candidates were selected and compiled for further analysis.

5.2.4. Hvt-miR-H14-3p Expression Analysis

5.2.4.1. Quantitative Real Time PCR

The RNA from healthy normal, PDAC, and CP patients was isolated using Trizol reagent from Sigma. 10ng of RNA from the samples is reverse transcribed to miRNA specific cDNA using miRVana assay system following manufacturers' protocol from Life technologies. The cDNA from all biological groups is now quantitated using hydrolysis probes from the assay system. RNU44 was used as reference gene to normalize expression of viral miRNA across the samples.

5.2.4.2. Digital PCR

The cDNA prepared as described in the previous section was also used for digital PCR with two independent platforms.

Bio-Rad QX100 digital PCR platform used the same qPCR reaction setup from life technologies and generated approximately 14,000 to 18,000 oil droplets per reaction. The quantitative data generated from each droplet per reaction was analyzed using QuantaSoft software to determine copy number of gene of interest from the samples.

RainDance platform also used the same qPCR reaction setup from life technologies but generated approximately 4 million oil droplets per reaction. The quantitative data generated

from each droplet per reaction was analyzed using RainDance Analyst software to determine relative fold change of gene expression across samples.

5.2.5. *In vitro* Experiments

5.2.5.1. Cell Culture

Different types of PDAC cell lines were cultured. MIAPaCa-2, Capan-1 and PANC-1 were cultured in DMEM containing 10% FBS, 1% antibiotic. MIAPaCa-2 was further supplemented with 1% pyruvate. BxPC-3, CFPAC-1 and Capan-2 were grown in RPMI, IMDM and McCoy medium respectively, each supplemented with 10% FBS, 1% antibiotic . All these cell lines were passaged 1:5 every fourth day. All the cells were cultured by incubation in a dedicated chamber with 5% CO_2 supply at 37^{o}C temperature. The cells were regularly tested for mycoplasma contamination and also genotyped for authenticity.

5.2.5.2. Transfection of Cells

RNA transfections were carried out in 6-well plates using siPORTTM NeoFXTM (Ambion) reagent and Lipofectamine 2000 (Invitrogrn). Reverse transfection by means of siPORT^{TM} NeoFX^{TM} involves simultaneous transfecting and plating of cells. siPORTTM NeoFXTM transfection agent and the RNA molecules are mixed and distributed on the culture plates over which the cells are overlaid. The final transfection volume in a 6-well plate is 2.5 ml of medium containing 2 x 10^5 cells per well. As the transfection complexes are stable in presence of serum, no change of medium or other precautionary measures taken in case of Lipofectamine 2000 transfection method are needed. The final concentration of the RNA molecules transfected was 10µM. After this procedure, the plates were maintained in incubation chamber at 37ºC and 5% CO2.

5.2.5.3. Migration Assays

HTS Transwell inserts were used by adding serum rich chemoattractant medium 235µL/well to the 96 well plates, followed by adding the Transwell inserts and lastly adding the 75µL serum free medium with 10,000 cells overexpressing viral miRNA or non-targeting negative control, to the inside compartment per well. An initial equilibrium period was used to improve cell attachment by adding medium to the 96 well plate wells and then to the Transwell inserts. The plate was then incubated for at least one hour at the same temperature that was used to grow the cells. The cells are then added in the fresh medium to the Transwell insert and returned to the incubation chamber. To count the number of migrated cells, both the well plate and Transwell inserts were rinsed with PBS. Trypsin was then added in the plate wells and incubated to dissociate the migrated cells in to the well plate. The detached cells were then measured by CellTiter-Glo Luminescent Cell Viability Assay in 8 technical replicates for each condition in 3 independent experiments.

5.2.5.4. Invasion Assays

BioCoat Matrigel TM invasion plates were used for the assay to check invasion of PDAC cells after overexpressing viral miRNA or non-targeting negative control. 0.1 Million cells per well were seeded in the inserts with a serum free medium. BD falcon TC companion plate was filled with 500µL of the serum rich medium. The serum rich medium in the plate wells acts as chemoattractant. 24hr post incubation, the Matrigel chambers were rinsed with PBS, placed in the companion plate wells with accutase and incubated for 10 min. This procedure allowed the invaded cells on the lower side of the Matrigel to detach and these detached cells were measured by CellTiter-Glo luminescent Cell Viability Assay in 8 technical replicates for each condition in 3 independent experiments.

5.2.5.5. Proliferation Assays

The proliferation assay was performed with PDAC cells overexpressing viral miRNA or non-targeting negative control. 10,000 cells per well were seeded in 96 well transparent and opaque plates. The cells seeded in transparent plates were assayed for protein expression by sulforhodamine B (SRB) colorimetric assay after 24hr, 48hr, 36hr, and 72hr. After incubation of cells at 37ºC with 5%CO_2, the plates were taken out and medium is discarded. 10% (w/v) TCA was added to the wells and the plates were incubated at 4ºC for 2 hours to fix the cells. TCA is then discarded and the plates were rinsed with water and dried at 37ºC for 20 minutes. 0.05% (w/v) of SRB reagent was added to the wells and the plates were incubated for 30 min at room temperature in dark. The plates were washed for 3 to 4 times with 1% (v/v) acetic acid to remove SRB reagent and then dried for 20 minutes at 37ºC. 100 mM Tris was added to the plates and the plates were shaken for 10 minutes after which, the absorbance was measured at 570 nm from the stained cells and at 650 nm from blank after which, the results were tabulated for calculating the percentage of viable cells after transfection. The cells seeded in opaque plates were assayed for ATP using CellTiter-Glo Luminescent Cell Viability Assay according to the manufacturers' protocol. All assays were performed with 8 technical replicates for each condition for 3 independent experiments.

5.2.5.6. Western Blot

10% SDS gels were used for resolving protein. 10% and 5% acrylamide/bisarcylamide were used respectively for resolving and stacking part of the gel. 0.06% (w/v) ammoniumpersulphate and 0.1% (v/v) N, N, N', N' – tetramethylethylenediamine (TEMED) were used to induce the polymerization of the gel. 10 µg of protein with rotiload loading dye were boiled together for 5 minutes and loaded into the respective slots in the gel. A prestained- protein ladder was also loaded referring to the molecular weight. Electrophoresis

of the gel was carried out for 90 minutes at 135 V and 500 mA in 1X SDS-gel tank buffer. The transfer of polypeptides from the gel to a nitrocellulose membrane was carried out by TE70 PWR semidry transfer apparatus. A sandwich model was prepared by soaking Whatman filter papers in anode buffers I, II and cathode buffer. The membrane was activated in anode buffer II. The stacking part of the gel was cut and the sandwich was assembled with the filter papers, membrane and the gel after which the semidry electrophoretic transfer was carried out for 60 minutes at 35 V and 500 mA. To detect the transferred protein, after the transfer the membrane was blocked for 1h at room temperature with the milk blocking buffer. After blocking, the membrane was incubated with the diluted primary antibody over night at 4ºC. After incubation, the membrane was washed 3 times with 1XTBST and was incubated with secondary antibody conjugated with horse radish peroxidase for 1h at room temperature. Then, the membrane was washed for 3 times with 1XTBST and protein was detected by enhanced chemiluminescence (ECL) using the ECL prime western blot detection kit. The ECL substrate was prepared according to the manufacturer's instructions and incubated on the membrane for 1 minute and the solution was drained off. Now, the membrane was kept on a clean plate inside the LAS Fujifilm 5000 machine and images were captured using a CCD camera on exposing the membrane gradually to the X- rays. Similarly, the procedure was repeated for the detection of the house keeping protein in the same membrane.

5.2.5.7. Dual Luciferase Assays

The target genes of hvt-miR-H14-3p are first observed to be down regulated at protein level by Western Blot. For validation, a dual luciferase assay system was produced. The p27 3'UTR wild type and mutant dual luciferase vectors were kindly provided by Prof. Reuven Agami from the Netherlands Cancer Institute (NKI). These vectors were transformed into suitable chemically competent bacteria, and plasmids were isolated post overnight culture.

The isolated plasmids were co-transfected with viral miRNA mimic or non-targeting negative control into CFPAC-1 cells. Empty vector is also transfected separately to normalize the luciferase expression.

5.2.5.8. Construction of Lentivirus Vectors

The desired lentiviral plasmids were selected Addgene. These plasmids include pLKO.1 - TRC control as transfer vector, pCMV-VSVG as envelope vector, and pCMV-dR8.2 dvpr as packaging vector for second generation packaging. The scramble shRNA plasmid was also ordered as non-targeting control. For *in vivo* bioluminescence tracking of tumor progression in mice, pLenti CMV Puro LUC (w168-1) vector was ordered form Addgene, to amplify the CMV promoter and firefly luciferase gene from it and clone into, both the transfer and scramble vectors. The vector maps for all plasmids used in this section can be found in the supplementary data section.

The hvt-mir-H14 shRNA construct was designed to be compatible with PLKO-1 lentiviral transfer vector system. The sense and antisense viral shRNA oligomers were designed in the way that following the annealing they generate *AgeI* and *EcoRI* restriction sites respectively in 5' and 3' of the construct. The annealing reaction of sense and antisense oligos was performed by incubation of the strands at 95°C for 5 minutes in a heating block, then slowly cooling down to room temperature overnight. The annealing reaction contains 100 µM of each of sense and antisense oligomers with 5 µL 10x NEB buffer 2 in 35 µL distilled H2O.

To clone the shRNA into the pLKO.1 transfer vector, double digestion using AgeI and EcoRI Fast digest restriction enzymes was performed. The digested plasmid was purified using 5-Prime PCR purification kit. The purified and linearized PLKO-1 plasmid was ligated with shRNA construct and transformed into chemically competent *E.coli*. To screen for the

recombinant plasmids, colony PCR using forward primer on the shRNA construct and the reverse primer on the plasmid sequence was performed. The recombinant colonies with desired sequence were cultured; plasmid was purified, and sent for sequencing to confirm the right clones. The next step was to clone the CMV-Luc gene sequence into the both PLKO-1-shRNA and scramble control vectors. The *ClaI* enzyme, that cuts the insertion site in the plasmid, is blocked by CpG and *dam* methylation. Hence, both plasmids were transformed into the K12 *E.coli*, and the plasmids were unmethylated in this strain of *E.coli*. The CMV-Luc sequence was amplified from pLenti CMV Puro LUC (w168-1) and cloned into PLKO-1 and scramble vectors using in-fusion cloning system from Invitrogen. After the ligation and transformation, positive clones were selected by colony PCR and confirmed by single read sequencing.

5.2.5.9. Production of Lentiviral Particles

Seeding cells: HEK293T were seeded in DMEM, 10% FCS, no PS, with fresh L-Glutamine (4 mM) added at 350.000-500.000 cells per well of a 6-well plate in 2 mL per well (= 17.5 x 10^4 cells/mL – 25 x 10^4 cells/mL).

Transfection: Mastermix was prepared for packaging and envelope vectors, distributed into safe lock tubes. “Plasmid plus Reagent Mix” was inncubated for 15 min at room temperature. “Lipofectamine Mix” was added to each plasmid-mix and incubated for 15 min at room temperature. “Complete Mix” was added (Lipofectamine + Plasmid Mix) to each well into cell medium and incubated for 18hr. The transfection medium was removed carefully after incubation. 2 mL DMEM with 30 % FCS and 1 % P/S was added to each well.

First viral harvest: Viral supernatant for each virus type was pooled, dead cells pelleted and kept at -80°C. The virus population is collected in 100 µL aliquots in PCR-strip-tubes (kept in a 50mL Falcon tube).

5.2.5.10. Transduction of Cells

BxPC-3 and MIAPaCa-2 cells were transduced with two lentiviral populations separately. The transduction was enhanced by the use of polybrene transduction agent. 1 million cells were seeded per well in 6 well plates. The lentivirus and added to the growth medium in a ratio of 1:50 along with polybrene. After this procedure, the plates were maintained in incubation chamber at 37ºC and 5% CO2. The cells were selected for recombinants under puromycin selection pressure.

5.2.6. *In vivo* Experiments

Tumor xenograft experiments were carried out in NOD *scid* gamma mice using BxPC-3 and MIAPaCa-2 transduced with viral miRNA or scramble negative control. 60 million lentivirus transduced PDAC cells diluted in 1x PBS in appropriate volume were mixed with equal volume of Matrigel and subcutaneously injected into the right flank of mice. The tumor size (length, width and depth) was measure using the caliper. The mice cohorts with BxPC-3 cells were injected with D-luciferin before bioluminescence (BLI) measurements by IVIS Xenogen machine. Currently, these cohorts are being sacrificed 3 weeks post inoculation and primary tumors are also being resected.

6. REFERENCES

Ambinder, R. F. (2000). "Gammaherpesviruses and "Hit-and-Run" oncogenesis." Am J Pathol **156**(1): 1-3.

Apte, M. V., Z. Xu, S. Pothula, D. Goldstein, R. C. Pirola and J. S. Wilson (2015). "Pancreatic cancer: The microenvironment needs attention too!" Pancreatology.

Arends, M. J., C. H. Buckley and M. Wells (1998). "Aetiology, pathogenesis, and pathology of cervical neoplasia." Journal of Clinical Pathology **51**(2): 96-103.

Argani, P., C. Iacobuzio-Donahue, B. Ryu, C. Rosty, M. Goggins, R. E. Wilentz, S. R. Murugesan, S. D. Leach, E. Jaffee, C. J. Yeo, J. L. Cameron, S. E. Kern and R. H. Hruban (2001). "Mesothelin is overexpressed in the vast majority of ductal adenocarcinomas of the pancreas: Identification of a new pancreatic cancer marker by serial analysis of gene expression (SAGE)." Clinical Cancer Research **7**(12): 3862-3868.

Band, V., D. Zajchowski, V. Kulesa and R. Sager (1990). "Human Papilloma-Virus Dnas Immortalize Normal Human Mammary Epithelial-Cells and Reduce Their Growth-Factor Requirements." Proceedings of the National Academy of Sciences of the United States of America **87**(1): 463-467.

Barzon, L., E. Lavezzo, V. Militello, S. Toppo and G. Palu (2011). "Applications of next-generation sequencing technologies to diagnostic virology." Int J Mol Sci **12**(11): 7861-7884.

Bechtold, V., P. Beard and K. Raj (2003). "Human papillomavirus type 16 E2 protein has no effect on transcription from episomal viral DNA." Journal of Virology **77**(3): 2021-2028.

Belanger, C., A. Gravel, A. Tomoiu, M. E. Janelle, J. Gosselin, M. J. Tremblay and L. Flamand (2001). "Human herpesvirus 8 viral FLICE-inhibitory protein inhibits Fas-mediated apoptosis through binding and prevention of procaspase-8 maturation." Journal of Human Virology **4**(2): 62-73.

Bill, C. A. and J. Summers (2004). "Genomic DNA double-strand breaks are targets for hepadnaviral DNA integration." Proc Natl Acad Sci U S A **101**(30): 11135-11140.

Blackford, A., O. K. Serrano, C. L. Wolfgang, G. Parmigiani, S. Jones, X. S. Zhang, D. W. Parsons, J. C. H. Lin, R. J. Leary, J. R. Eshleman, M. Goggins, E. M. Jaffee, C. A. Iacobuzio-Donahue, A. Maitra, J. L. Cameron, K. Olino, R. Schulick, J. Winter, J. M. Herman, D. Laheru, A. P. Klein, B. Vogelstein, K. W. Kinzler, V. E. Velculescu and R. H. Hruban (2009). "SMAD4 Gene Mutations Are Associated with Poor Prognosis in Pancreatic Cancer." Clinical Cancer Research **15**(14): 4674-4679.

Braun, L., J. E. Mead, M. Panzica, R. Mikumo, G. I. Bell and N. Fausto (1988). "Transforming growth factor beta mRNA increases during liver regeneration: a possible paracrine mechanism of growth regulation." Proc Natl Acad Sci U S A **85**(5): 1539-1543.

Brechot, C., C. Pourcel, A. Louise, B. Rain and P. Tiollais (1980). "Presence of integrated hepatitis B virus DNA sequences in cellular DNA of human hepatocellular carcinoma." Nature **286**(5772): 533-535.

Buchkovich, N. J., T. G. Maguire, Y. Yu, A. W. Paton, J. C. Paton and J. C. Alwine (2008). "Human cytomegalovirus specifically controls the levels of the endoplasmic reticulum chaperone BiP/GRP78, which is required for virion assembly." Journal of Virology **82**(1): 31-39.

Buchkovich, N. J., Y. Yu, C. A. Zampieri and J. C. Alwine (2008). "The TORrid affairs of viruses: effects of mammalian DNA viruses on the PI3K-Akt-mTOR signalling pathway." Nature Reviews Microbiology **6**(4): 265-275.

Calleja-Macias, I. E., M. Kalantari, B. Allan, A. L. Williamson, L. P. Chung, R. J. Collins, R. E. Zuna, S. T. Dunn, R. Ortiz-Lopez, H. A. Barrera-Saldana, H. A. Cubie, K. Cuschieri, L. L. Villa and H. U. Bernard (2005). "Papillomavirus subtypes are natural and old taxa: Phylogeny of human papillomavirus types 44 and 55 and 68a and -b." Journal of Virology **79**(10): 6565-6569.

Calnek, B. W. and R. L. Witter (1991). Marek's disease. Ames, Iowa, Iowa State University Press.

Canto, M. I., R. H. Hruban, E. K. Fishman, I. R. Kamel, R. Schulick, Z. Zhang, M. Topazian, N. Takahashi, J. Fletcher, G. Petersen, A. P. Klein, J. Axilbund, C. Griffin, S. Syngal, J. R. Saltzman, K. J. Mortele, J. Lee, E. Tamm, R. Vikram, P. Bhosale, D. Margolis, J. Farrell, M. Goggins and C. American Cancer of the Pancreas Screening (2012). "Frequent detection of pancreatic lesions in asymptomatic high-risk individuals." Gastroenterology **142**(4): 796-804; quiz e714-795.

Cho, J. W., W. K. Baek, S. H. Yang, J. Chang, Y. C. Sung and M. H. Suh (2001). "HCV core protein modulates Rb pathway through pRb down-regulation and E2F-1 up-regulation." Biochimica Et Biophysica Acta-Molecular Cell Research **1538**(1): 59-66.

Clifford, G. M., R. K. Rana, S. Franceschi, J. S. Smith, G. Gough and J. M. Pimenta (2005). "Human papillomavirus genotype distribution in low-grade cervical lesions: comparison by geographic region and with cervical cancer." Cancer Epidemiol Biomarkers Prev **14**(5): 1157-1164.

Conway, C., R. Chalkley, A. High, K. Maclennan, S. Berri, P. Chengot, M. Alsop, P. Egan, J. Morgan, G. R. Taylor, J. Chester, M. Sen, P. Rabbitts and H. M. Wood (2012). "Next-Generation Sequencing for Simultaneous Determination of Human Papillomavirus Load,

Subtype, and Associated Genomic Copy Number Changes in Tumors." Journal of Molecular Diagnostics **14**(2): 104-111.

Czaja, M. J., F. R. Weiner, K. C. Flanders, M. A. Giambrone, R. Wind, L. Biempica and M. A. Zern (1989). "In vitro and in vivo association of transforming growth factor-beta 1 with hepatic fibrosis." J Cell Biol **108**(6): 2477-2482.

Delecluse, H. J., S. Bartnizke, W. Hammerschmidt, J. Bullerdiek and G. W. Bornkamm (1993). "Episomal and integrated copies of Epstein-Barr virus coexist in Burkitt lymphoma cell lines." J Virol **67**(3): 1292-1299.

Desaintes, C., C. Demeret, S. Goyat, M. Yaniv and F. Thierry (1997). "Expression of the papillomavirus E2 protein in HeLa cells leads to apoptosis." Embo Journal **16**(3): 504-514.

Dirmeier, U., R. Hoffmann, E. Kilger, U. Schultheiss, C. Briseno, O. Gires, A. Kieser, D. Eick, B. Sugden and W. Hammerschmidt (2005). "Latent membrane protein 1 of Epstein-Barr virus coordinately regulates proliferation with control of apoptosis." Oncogene **24**(10): 1711-1717.

Dittmer, D. P. and S. E. Krown (2007). "Targeted therapy for Kaposi's sarcoma and Kaposi's sarcoma-associated herpesvirus." Current Opinion in Oncology **19**(5): 452-457.

Efklidou, S., R. Bailey, N. Field, M. Noursadeghi and M. K. Collins (2008). "vFLIP from KSHV inhibits anoikis of primary endothelial cells." Journal of Cell Science **121**(4): 450-457.

El-Serag, H. B. and K. L. Rudolph (2007). "Hepatocellular carcinoma: epidemiology and molecular carcinogenesis." Gastroenterology **132**(7): 2557-2576.

Epstein, M. A., B. G. Achong and Y. M. Barr (1964). "Virus Particles in Cultured Lymphoblasts from Burkitt's Lymphoma." Lancet **1**(7335): 702-703.

Ernberg, I. and G. Klein (2007). Human Herpesviruses: Biology, Therapy and Immunoprophylaxis. Cambridge, UK, Cambridge Univ. Press.

Fabricant, J., B. W. Calnek and K. A. Schat (1982). "The early pathogenesis of turkey herpesvirus infection in chickens and turkeys." Avian Dis **26**(2): 257-264.

Fausto, N. and J. E. Mead (1989). "Biology of Disease - Regulation of Liver Growth - Protooncogenes and Transforming Growth-Factors." Laboratory Investigation **60**(1): 4-13.

Feig, C., A. Gopinathan, A. Neesse, D. S. Chan, N. Cook and D. A. Tuveson (2012). "The pancreas cancer microenvironment." Clin Cancer Res **18**(16): 4266-4276.

Feitelson, M. A. and J. Lee (2007). "Hepatitis B virus integration, fragile sites, and hepatocarcinogenesis." Cancer Letters **252**(2): 157-170.

Feitelson, M. A., H. M. Reis, N. L. Tufan, B. Sun, J. Pan and Z. Lian (2009). "Putative roles of hepatitis B x antigen in the pathogenesis of chronic liver disease." Cancer Lett **286**(1): 69-79.

Ferber, M. J., E. C. Thorland, A. A. Brink, A. K. Rapp, L. A. Phillips, R. McGovern, B. S. Gostout, T. H. Cheung, T. K. Chung, W. Y. Fu and D. I. Smith (2003). "Preferential integration of human papillomavirus type 18 near the c-myc locus in cervical carcinoma." Oncogene **22**(46): 7233-7242.

Fredricks, D. N. and D. A. Relman (1996). "Sequence-based identification of microbial pathogens: a reconsideration of Koch's postulates." Clin Microbiol Rev **9**(1): 18-33.

Frohme, M. and J. D. Hoheisel (2006). Representational difference analysis: A methodology to study differential gene expression. Cell Biology: A laboratory handbook, Elsevier Academic. **1:** 113-120.

Fujimoto, A., Y. Totoki, T. Abe, K. A. Boroevich, F. Hosoda, H. H. Nguyen, M. Aoki, N. Hosono, M. Kubo, F. Miya, Y. Arai, H. Takahashi, T. Shirakihara, M. Nagasaki, T. Shibuya, K. Nakano, K. Watanabe-Makino, H. Tanaka, H. Nakamura, J. Kusuda, H. Ojima, K. Shimada, T. Okusaka, M. Ueno, Y. Shigekawa, Y. Kawakami, K. Arihiro, H. Ohdan, K. Gotoh, O. Ishikawa, S. Ariizumi, M. Yamamoto, T. Yamada, K. Chayama, T. Kosuge, H. Yamaue, N. Kamatani, S. Miyano, H. Nakagama, Y. Nakamura, T. Tsunoda, T. Shibata and H. Nakagawa (2012). "Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators." Nat Genet **44**(7): 760-764.

Garrido-Laguna, I. and M. Hidalgo (2015). "Pancreatic cancer: from state-of-the-art treatments to promising novel therapies." Nat Rev Clin Oncol.

Garzon, R., G. Marcucci and C. M. Croce (2010). "Targeting microRNAs in cancer: rationale, strategies and challenges." Nat Rev Drug Discov **9**(10): 775-789.

Gibbs, C. P., K. Nazerian, L. F. Velicer and H. J. Kung (1984). "Extensive Homology Exists between Marek Disease Herpesvirus and Its Vaccine Virus, Herpesvirus of Turkeys." Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences **81**(11): 3365-3369.

Godfrey, A., J. Anderson, A. Papanastasiou, Y. Takeuchi and C. Boshoff (2005). "Inhibiting primary effusion lymphoma, by lentiviral vectors encoding short hairpin RNA." Blood **105**(6): 2510-2518.

Goodwin, E. C. and D. DiMaio (2000). "Repression of human papillomavirus oncogenes in HeLa cervical carcinoma cells causes the orderly reactivation of dormant tumor suppressor pathways." Proc Natl Acad Sci U S A **97**(23): 12513-12518.

Guerrero, R. B. and L. R. Roberts (2005). "The role of hepatitis B virus integrations in the pathogenesis of human hepatocellular carcinoma." Journal of Hepatology **42**(5): 760-777.

Guichard, C., G. Amaddeo, S. Imbeaud, Y. Ladeiro, L. Pelletier, I. Ben Maad, J. Calderaro, P. Bioulac-Sage, M. Letexier, F. Degos, B. Clement, C. Balabaud, E. Chevet, A. Laurent, G. Couchy, E. Letouze, F. Calvo and J. Zucman-Rossi (2012). "Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma." Nature Genetics **44**(6): 694-U120.

Guo, M., N. Sneige, E. G. Silva, Y. J. Jan, D. E. Cogdell, E. Lin, R. Luthra and W. Zhang (2007). "Distribution and viral load of eight oncogenic types of human papillomavirus (HPV) and HPV 16 integration status in cervical intraepithelial neoplasia and carcinoma." Mod Pathol **20**(2): 256-266.

Habbe, N., J. B. M. Koorstra, J. T. Mendell, G. J. Offerhaus, J. K. Ryu, G. Feldmann, M. E. Mullendore, M. G. Goggins, S. M. Hong and A. Maitra (2009). "MicroRNA miR-155 is a biomarker of early pancreatic neoplasia." Cancer Biology & Therapy **8**(4): 340-346.

Hafner, N., C. Driesch, M. Gajda, L. Jansen, R. Kirchmayr, I. B. Runnebaum and M. Durst (2008). "Integration of the HPV16 genome does not invariably result in high levels of viral oncogene transcripts." Oncogene **27**(11): 1610-1617.

Hahn, S. A., M. Schutte, A. T. M. S. Hoque, C. A. Moskaluk, L. T. daCosta, E. Rozenblum, C. L. Weinstein, A. Fischer, C. J. Yeo, R. H. Hruban and S. E. Kern (1996). "DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1." Science **271**(5247): 350-353.

Hamilton, A. J. and D. C. Baulcombe (1999). "A species of small antisense RNA in posttranscriptional gene silencing in plants." Science **286**(5441): 950-952.

Han, H. and D. D. Von Hoff (2013). "SnapShot: pancreatic cancer." Cancer Cell **23**(3): 424-424 e421.

Hanahan, D. and R. A. Weinberg (2000). "The Hallmarks of Cancer." Cell **100**(1): 57-70.

Hanahan, D. and R. A. Weinberg (2011). "Hallmarks of cancer: the next generation." Cell **144**(5): 646-674.

Harsha, H. C., K. Kandasamy, P. Ranganathan, S. Rani, S. Ramabadran, S. Gollapudi, L. Balakrishnan, S. B. Dwivedi, D. Telikicherla, L. D. N. Selvan, R. Goel, S. Mathivanan, A. Marimuthu, M. Kashyap, R. F. Vizza, R. J. Mayer, J. A. DeCaprio, S. Srivastava, S. M. Hanash, R. H. Hruban and A. Pandey (2009). "A Compendium of Potential Biomarkers of Pancreatic Cancer." Plos Medicine **6**(4).

Hassan, R., T. Bera and I. Pastan (2004). "Mesothelin: A new target for immunotherapy." Clinical Cancer Research **10**(12): 3937-3942.

Helt, A. M. and D. A. Galloway (2003). "Mechanisms by which DNA tumor virus oncoproteins target the Rb family of pocket proteins." Carcinogenesis **24**(2): 159-169.

Henle, J. (1840). Von den Contagien und Miasmen und den contagiös-miasmatischen Krankheiten. Berlin.

Henle, W., V. Diehl, G. Kohn, H. Zur Hausen and G. Henle (1967). "Herpes-type virus and chromosome marker in normal leukocytes after growth with irradiated Burkitt cells." Science **157**(3792): 1064-1065.

Herdman, M. T., M. R. Pett, I. Roberts, W. O. Alazawi, A. E. Teschendorff, X. Y. Zhang, M. A. Stanley and N. Coleman (2006). "Interferon-beta treatment of cervical keratinocytes naturally infected with human papillomavirus 16 episomes promotes rapid reduction in episome numbers and emergence of latent integrants." Carcinogenesis **27**(11): 2341-2353.

Hopman, A. H., F. Smedts, W. Dignef, M. Ummelen, G. Sonke, M. Mravunac, G. P. Vooijs, E. J. Speel and F. C. Ramaekers (2004). "Transition of high-grade cervical intraepithelial neoplasia to micro-invasive carcinoma is characterized by integration of HPV 16/18 and numerical chromosome abnormalities." J Pathol **202**(1): 23-33.

Horikawa, I. and J. C. Barrett (2001). "cis-Activation of the human telomerase gene (hTERT) by the hepatitis B virus genome." J Natl Cancer Inst **93**(15): 1171-1173.

Houben, R., M. Shuda, R. Weinkam, D. Schrama, H. C. Feng, Y. A. Chang, P. S. Moore and J. C. Becker (2010). "Merkel Cell Polyomavirus-Infected Merkel Cell Carcinoma Cells Require Expression of Viral T Antigens." Journal of Virology **84**(14): 7064-7072.

Hruban, R. H., A. D. M. Vanmansfeld, G. J. A. Offerhaus, D. H. J. Vanweering, D. C. Allison, S. N. Goodman, T. W. Kensler, K. K. Bose, J. L. Cameron and J. L. Bos (1993). "K-Ras Oncogene Activation in Adenocarcinoma of the Human Pancreas - a Study of 82 Carcinomas Using a Combination of Mutant-Enriched Polymerase Chain-Reaction Analysis and Allele-Specific Oligonucleotide Hybridization." American Journal of Pathology **143**(2): 545-554.

Huang, J., Q. Deng, Q. Wang, K. Y. Li, J. H. Dai, N. Li, Z. D. Zhu, B. Zhou, X. Y. Liu, R. F. Liu, Q. L. Fei, H. Chen, B. Cai, B. P. Zhou, H. S. Xiao, L. X. Qin and Z. G. Han (2012). "Exome sequencing of hepatitis B virus-associated hepatocellular carcinoma." Nature Genetics **44**(10): 1117-+.

Hubank, M. and D. G. Schatz (1994). "Identifying differences in mRNA expression by representational difference analysis of cDNA." Nucleic Acids Res **22**(25): 5640-5648.

Hustinx, S. R., L. M. Leoni, C. J. Yeo, P. N. Brown, M. Goggins, S. E. Kern, R. H. Hruban and A. Maitra (2005). "Concordant loss of MTAP and p16/CDKN2A expression in pancreatic intraepithelial neoplasia: evidence of homozygous deletion in a noninvasive precursor lesion." Mod Pathol **18**(7): 959-963.

Hwang, S., D. Lee, Y. Gwack, H. Min and J. Choe (2003). "Kaposi's sarcoma-associated herpesvirus K8 protein interacts with hSNF5." Journal of General Virology **84**: 665-676.

Iacobuzio-Donahue, C. A., B. J. Fu, S. Yachida, M. D. Luo, H. Abe, C. M. Henderson, F. Vilardell, Z. Wang, J. W. Keller, P. Banerjee, J. M. Herman, J. L. Cameron, C. J. Yeo, M. K. Halushka, J. R. Eshleman, M. Raben, A. P. Klein, R. H. Hruban, M. Hidalgo and D. Laheru (2009). "DPC4 Gene Status of the Primary Carcinoma Correlates With Patterns of Failure in Patients With Pancreatic Cancer." Journal of Clinical Oncology **27**(11): 1806-1813.

Javier, R. T. and J. S. Butel (2008). "The history of tumor virology." Cancer Res **68**(19): 7693-7706.

Jiang, Z. S., S. Jhunjhunwala, J. F. Liu, P. M. Haverty, M. I. Kennemer, Y. H. Guan, W. Lee, P. Carnevali, J. Stinson, S. Johnson, J. Y. Diao, S. Yeung, A. Jubb, W. L. Ye, T. D. Wu, S. B. Kapadia, F. J. de Sauvage, R. C. Gentleman, H. M. Stern, S. Seshagiri, K. P. Pant, Z. Modrusan, D. G. Ballinger and Z. M. Zhang (2012). "The effects of hepatitis B virus integration into the genomes of hepatocellular carcinoma patients." Genome Research **22**(4): 593-601.

Jiao, Y., C. Shi, B. H. Edil, R. F. de Wilde, D. S. Klimstra, A. Maitra, R. D. Schulick, L. H. Tang, C. L. Wolfgang, M. A. Choti, V. E. Velculescu, L. A. Diaz, Jr., B. Vogelstein, K. W. Kinzler, R. H. Hruban and N. Papadopoulos (2011). "DAXX/ATRX, MEN1, and mTOR pathway genes are frequently altered in pancreatic neuroendocrine tumors." Science **331**(6021): 1199-1203.

Jones, S., X. Zhang, D. W. Parsons, J. C. Lin, R. J. Leary, P. Angenendt, P. Mankoo, H. Carter, H. Kamiyama, A. Jimeno, S. M. Hong, B. Fu, M. T. Lin, E. S. Calhoun, M. Kamiyama, K. Walter, T. Nikolskaya, Y. Nikolsky, J. Hartigan, D. R. Smith, M. Hidalgo, S. D. Leach, A. P. Klein, E. M. Jaffee, M. Goggins, A. Maitra, C. Iacobuzio-Donahue, J. R. Eshleman, S. E. Kern, R. H. Hruban, R. Karchin, N. Papadopoulos, G. Parmigiani, B. Vogelstein, V. E. Velculescu and K. W. Kinzler (2008). "Core signaling pathways in human pancreatic cancers revealed by global genomic analyses." Science **321**(5897): 1801-1806.

Jox, A., C. Rohen, G. Belge, S. Bartnitzke, M. Pawlita, V. Diehl, J. Bullerdiek and J. Wolf (1997). "Integration of Epstein-Barr virus in Burkitt's lymphoma cells leads to a region of enhanced chromosome instability." Ann Oncol **8 Suppl 2**: 131-135.

Kanda, M., H. Matthaei, J. Wu, S. M. Hong, J. Yu, M. Borges, R. H. Hruban, A. Maitra, K. Kinzler, B. Vogelstein and M. Goggins (2012). "Presence of Somatic Mutations in Most Early-Stage Pancreatic Intraepithelial Neoplasia." Gastroenterology **142**(4): 730-U129.

Kieff, E. and A. B. Rickinson (2007). Epstein-barr virus and its replication. Fields Virology. D. M. Knipe and P. M. Howley. Philadelphia, PA, USA, Lippincott Williams and Wilkins**:** 2603–2654.

Kincaid, R. P. and C. S. Sullivan (2012). "Virus-encoded microRNAs: an overview and a look to the future." PLoS Pathog **8**(12): e1003018.

Koch, R. (1884). Die Aetiologie der Tuberkulose. Mittheilungen aus dem Kaiserlichen Gesundheitsamt. **2:** 1-88.

Koch, R. (1942). SourceBookofMedicalHistory. New York, Dover Publications, Inc.

Kovalchuk, A. L., C. F. Qi, T. A. Torrey, L. Taddesse-Heath, L. Feigenbaum, S. S. Park, A. Gerbitz, G. Klobeck, K. Hoertnagel, A. Polack, G. W. Bornkamm, S. Janz and H. C. Morse, 3rd (2000). "Burkitt lymphoma in the mouse." J Exp Med **192**(8): 1183-1190.

Kremsdorf, D., P. Soussan, P. Paterlini-Brechot and C. Brechot (2006). "Hepatitis B virus-related hepatocellular carcinoma: paradigms for viral-related human carcinogenesis." Oncogene **25**(27): 3823-3833.

Kumar, P., A. Saha and E. S. Robertson (2010). "Epstein-Barr virus hijacks cell cycle machinery." Microbe **5**.

Lambert, P. J., A. Z. Shahrier, A. G. Whitman, O. F. Dyson, A. J. Reber, J. A. McCubrey and S. M. Akula (2007). "Targeting the PI3K and MAPK pathways to treat Kaposi's-sarcoma-associated herpes virus infection and pathogenesis." Expert Opin Ther Targets **11**(5): 589-599.

Lee, C. P., J. Y. Chen, J. T. Wang, K. Kimura, A. Takemoto, C. C. Lu and M. R. Chen (2007). "Epstein-Barr virus BGLF4 kinase induces premature chromosome condensation through activation of condensin and topoisomerase II." Journal of Virology **81**(10): 5166-5180.

Lee, J., D. H. Kim, S. Lee, Q. H. Yang, D. K. Lee, S. K. Lee, R. G. Roeder and J. W. Lee (2009). "A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4." Proc Natl Acad Sci U S A **106**(21): 8513-8518.

Lee, K. (1980). "Long term effects of Marek's disease vaccination with cell-free herpesvirus of turkey and age at debeaking on performance and mortality of white leghorns." Poult Sci **59**(9): 2002-2007.

Lee, S. G. and H. M. Rho (2000). "Transcriptional repression of the human p53 gene by hepatitis B viral X protein." Oncogene **19**(3): 468-471.

Levy, G. N. (1997). "Prostaglandin H synthases, nonsteroidal anti-inflammatory drugs, and colon cancer." Faseb Journal **11**(4): 234-247.

Li, J., C. Yen, D. Liaw, K. Podsypanina, S. Bose, S. I. Wang, J. Puc, C. Miliaresis, L. Rodgers, R. McCombie, S. H. Bigner, B. C. Giovanella, M. Ittmann, B. Tycko, H.

Hibshoosh, M. H. Wigler and R. Parsons (1997). "PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer." Science **275**(5308): 1943-1947.

Li, M., H. Zhao, X. S. Zhang, L. D. Wood, R. A. Anders, M. A. Choti, T. M. Pawlik, H. D. Daniel, R. Kannangai, G. J. A. Offerhaus, V. E. Velculescu, L. F. Wang, S. B. Zhou, B. Vogelstein, R. H. Hruban, N. Papadopoulos, J. Q. Cai, M. S. Torbenson and K. W. Kinzler (2011). "Inactivating mutations of the chromatin remodeling gene ARID2 in hepatocellular carcinoma." Nature Genetics **43**(9): 828-829.

Li, S. and M. Mao (2013). "Next generation sequencing reveals genetic landscape of hepatocellular carcinomas." Cancer Lett **340**(2): 247-253.

Li, Z., S. Van Calcar, C. Qu, W. K. Cavenee, M. Q. Zhang and B. Ren (2003). "A global transcriptional regulatory role for c-Myc in Burkitt's lymphoma cells." Proc Natl Acad Sci U S A **100**(14): 8164-8169.

Lisitsyn, N., N. Lisitsyn and M. Wigler (1993). "Cloning the Differences between two Complex Genomes." Science **259**(5097): 946-951.

Lisitsyn, N. A., N. M. Lisitsina, G. Dalbagni, P. Barker, C. A. Sanchez, J. Gnarra, W. M. Linehan, B. J. Reid and M. H. Wigler (1995). "Comparative Genomic Analysis of Tumors - Detection of DNA Losses and Amplification." Proceedings of the National Academy of Sciences of the United States of America **92**(1): 151-155.

Loeffler, F. (1884). "Utersuchung über die Bedeutung der Mikroorganismen für die Entstehung der Diphtherie beim Menschen, bei der Taube und beim Kalb." Mittheilungen aus dem Kaiserlichen Gesundheitsamt **2**: 421-499.

Lovat, F., N. Valeri and C. M. Croce (2011). "MicroRNAs in the pathogenesis of cancer." Semin Oncol **38**(6): 724-733.

Luber, B., N. Arnold, M. Sturzl, M. Hohne, P. Schirmacher, U. Lauer, J. Wienberg, P. H. Hofschneider and A. S. Kekule (1996). "Hepatoma-derived integrated HBV DNA causes multi-stage transformation in vitro." Oncogene **12**(8): 1597-1608.

Maitra, A., N. Fukushima, K. Takaori and R. H. Hruban (2005). "Precursors to invasive pancreatic cancer." Advances in Anatomic Pathology **12**(2): 81-91.

Marks, F., G. Furstenberger and K. Muller-Decker (2007). "Tumor promotion as a target of cancer prevention." Recent Results Cancer Res **174**: 37-47.

Mason, W. S., C. Liu, C. E. Aldrich, S. Litwin and M. M. Yeh (2010). "Clonal expansion of normal-appearing human hepatocytes during chronic hepatitis B virus infection." J Virol **84**(16): 8308-8315.

Matsuo, T., M. Heller, L. Petti, E. O'Shiro and E. Kieff (1984). "Persistence of the entire Epstein-Barr virus genome integrated into human lymphocyte DNA." Science **226**(4680): 1322-1325.

Matthaei, H., D. Wylie, M. B. Lloyd, M. Dal Molin, J. Kemppainen, S. C. Mayo, C. L. Wolfgang, R. D. Schulick, L. Langfield, B. F. Andruss, A. T. Adai, R. H. Hruban, A. E. Szafranska-Schwarzbach and A. Maitra (2012). "miRNA Biomarkers in Cyst Fluid Augment the Diagnosis and Management of Pancreatic Cysts." Clinical Cancer Research **18**(17): 4713-4724.

Meiring, T. L., A. T. Salimo, B. Coetzee, H. J. Maree, J. Moodley, I. I. Hitzeroth, M. J. Freeborough, E. P. Rybicki and A. L. Williamson (2012). "Next-generation sequencing of cervical DNA detects human papillomavirus types not detected by commercial kits." Virology Journal **9**.

Munger, K. and W. C. Phelps (1993). "The Human Papillomavirus-E7 Protein as a Transforming and Transactivating Factor." Biochimica Et Biophysica Acta **1155**(1): 111-123.

Murakami, Y., K. Saigo, H. Takashima, M. Minami, T. Okanoue, C. Brechot and P. Paterlini-Brechot (2005). "Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas." Gut **54**(8): 1162-1168.

Nakamura, T., Y. Tomita, R. Hirai, K. Yamaoka, K. Kaji and A. Ichihara (1985). "Inhibitory Effect of Transforming Growth-Factor-Beta on DNA-Synthesis of Adult-Rat Hepatocytes in Primary Culture." Biochemical and Biophysical Research Communications **133**(3): 1042-1050.

Nakatsukasa, H., P. Nagy, R. P. Evarts, C. C. Hsia, E. Marsden and S. S. Thorgeirsson (1990). "Cellular distribution of transforming growth factor-beta 1 and procollagen types I, III, and IV transcripts in carbon tetrachloride-induced rat liver fibrosis." J Clin Invest **85**(6): 1833-1843.

Needleman, P., J. Turk, B. A. Jakschik, A. R. Morrison and J. B. Lefkowith (1986). "Arachidonic-Acid Metabolism." Annual Review of Biochemistry **55**: 69-102.

O'Nions, F. and M. J. Allday (2004). "Deregulation of the cell cycle by the Epstein-Barr virus." Advances in Cancer Research, Vol 92 **92**: 119-186.

Olivares, O. and S. Vasseur (2015). "Metabolic rewiring of pancreatic ductal adenocarcinoma: New routes to follow within the maze." Int J Cancer.

Palermo, R. D., H. M. Webb, A. Gunnell and M. J. West (2008). "Regulation of transcription by the Epstein-Barr virus nuclear antigen EBNA 2." Biochemical Society Transactions **36**: 625-628.

Pantry, S. N. and P. G. Medveczky (2009). "Epigenetic regulation of Kaposi's sarcoma-associated herpesvirus replication." Seminars in Cancer Biology **19**(3): 153-157.

Parkin, D. M. and F. Bray (2006). "The burden of HPV-related cancers." Vaccine **24**: 11-25.

Parsonnet, J. (1999). Microbes and Malignancy. New York, Oxford Univ. Press.

Perz, J. F., G. L. Armstrong, L. A. Farrington, Y. J. Hutin and B. P. Bell (2006). "The contributions of hepatitis B virus and hepatitis C virus infections to cirrhosis and primary liver cancer worldwide." J Hepatol **45**(4): 529-538.

Peter, M., N. Stransky, J. Couturier, P. Hupe, E. Barillot, P. de Cremoux, P. Cottu, F. Radvanyi and X. Sastre-Garau (2010). "Frequent genomic structural alterations at HPV insertion sites in cervical carcinoma." Journal of Pathology **221**(3): 320-330.

Pett, M. and N. Coleman (2007). "Integration of high-risk human papillomavirus: a key event in cervical carcinogenesis?" J Pathol **212**(4): 356-367.

Pett, M. R., M. T. Herdman, R. D. Palmer, G. S. Yeo, M. K. Shivji, M. A. Stanley and N. Coleman (2006). "Selection of cervical keratinocytes containing integrated HPV16 associates with episome loss and an endogenous antiviral response." Proc Natl Acad Sci U S A **103**(10): 3822-3827.

Polack, A., K. Hortnagel, A. Pajic, B. Christoph, B. Baier, M. Falk, J. Mautner, C. Geltinger, G. W. Bornkamm and B. Kempkes (1996). "c-myc activation renders proliferation of Epstein-Barr virus (EBV)-transformed cells independent of EBV nuclear antigen 2 and latent membrane protein 1." Proc Natl Acad Sci U S A **93**(19): 10411-10416.

Popescu, N. C., M. C. Chen, S. Simpson, S. Solinas and J. A. Dipaolo (1993). "A Burkitt-Lymphoma Cell-Line with Integrated Epstein-Barr-Virus at a Stable Chromosome Modification Site." Virology **195**(1): 248-251.

Purchase, H. G., W. Okazaki and B. R. Burmester (1972). "Long-term field trials with the herpesvirus of turkeys vaccine against Marek's disease." Avian Dis **16**(1): 57-71.

Qin, D., N. H. Feng, W. F. Fan, X. T. Ma, Q. Yan, Z. G. Lv, Y. Zeng, J. Z. Zhu and C. Lu (2011). "Activation of PI3K/AKT and ERK MAPK signal pathways is required for the induction of lytic cycle replication of Kaposi's Sarcoma-associated herpesvirus by herpes simplex virus type 1." Bmc Microbiology **11**.

Reddy, B. S., Y. Hirose, R. Lubet, V. Steele, G. Kelloff, S. Paulson, K. Seibert and C. V. Rao (2000). "Chemoprevention of colon cancer by specific cyclooxygenase-2 inhibitor, celecoxib, administered during different stages of carcinogenesis." Cancer Res **60**(2): 293-297.

Roberts, N. J., Y. C. Jiao, J. Yu, L. Kopelovich, G. M. Petersen, M. L. Bondy, S. Gallinger, A. G. Schwartz, S. Syngal, M. L. Cote, J. Axilbund, R. Schulick, S. Z. Ali, J. R. Eshleman, V. E. Velculescu, M. Goggins, B. Vogelstein, N. Papadopoulos, R. H. Hruban, K. W. Kinzler and A. P. Klein (2012). "ATM Mutations in Patients with Hereditary Pancreatic Cancer." Cancer Discovery **2**(1): 41-46.

Ronco, L. V., A. Y. Karpova, M. Vidal and P. M. Howley (1998). "Human papillomavirus 16 E6 oncoprotein binds to interferon regulatory factor-3 and inhibits its transcriptional activity." Genes & Development **12**(13): 2061-2072.

Rozenblatt-Rosen, O., R. C. Deo, M. Padi, G. Adelmant, M. A. Calderwood, T. Rolland, M. Grace, A. Dricot, M. Askenazi, M. Tavares, S. J. Pevzner, F. Abderazzaq, D. Byrdsong, A. R. Carvunis, A. A. Chen, J. W. Cheng, M. Correll, M. Duarte, C. Y. Fan, M. C. Feltkamp, S. B. Ficarro, R. Franchi, B. K. Garg, N. Gulbahce, T. Hao, A. M. Holthaus, R. James, A. Korkhin, L. Litovchick, J. C. Mar, T. R. Pak, S. Rabello, R. Rubio, Y. Shen, S. Singh, J. M. Spangle, M. Tasan, S. Wanamaker, J. T. Webber, J. Roecklein-Canfield, E. Johannsen, A. L. Barabasi, R. Beroukhim, E. Kieff, M. E. Cusick, D. E. Hill, K. Munger, J. A. Marto, J. Quackenbush, F. P. Roth, J. A. DeCaprio and M. Vidal (2012). "Interpreting cancer genomes using systematic host network perturbations by tumour virus proteins." Nature **487**(7408): 491-495.

Ryan, D. P., T. S. Hong and N. Bardeesy (2014). "Pancreatic adenocarcinoma." N Engl J Med **371**(11): 1039-1049.

Saemundsen, A. K., D. T. Purtilo, K. Sakamoto, J. L. Sullivan, A. C. Synnerholm, D. Hanto, R. Simmons, M. Anvret, R. Collins and G. Klein (1981). "Documentation of Epstein-Barr Virus-Infection in Immunodeficient Patients with Life-Threatening Lymphoproliferative Diseases by Epstein-Barr Virus Complementary Rna-DNA and Viral DNA-DNA Hybridization." Cancer Research **41**(11): 4237-4242.

Scheurer, M. E., G. Tortolero-Luna and K. Adler-Storthz (2005). "Human papillomavirus infection: biology, epidemiology, and prevention." International Journal of Gynecological Cancer **15**(5): 727-746.

Schluter, V., M. Meyer, P. H. Hofschneider, R. Koshy and W. H. Caselmann (1994). "Integrated hepatitis B virus X and 3' truncated preS/S sequences derived from human hepatomas encode functionally active transactivators." Oncogene **9**(11): 3335-3344.

Schutte, M., L. T. Dacosta, S. A. Hahn, C. Moskaluk, A. T. M. S. Hoque, E. Rozenblum, C. L. Weinstein, M. Bittner, P. S. Meltzer, J. M. Trent, C. J. Yeo, R. H. Hruban and S. E. Kern (1995). "Identification by Representational Difference Analysis of a Homozygous Deletion in Pancreatic-Carcinoma That Lies within the Brca2 Region." Proceedings of the National Academy of Sciences of the United States of America **92**(13): 5950-5954.

Seeger, C. and W. S. Mason (2000). "Hepatitis B virus biology." Microbiology and Molecular Biology Reviews **64**(1): 51-+.

Shafritz, D. A., D. Shouval, H. I. Sherman, S. J. Hadziyannis and M. C. Kew (1981). "Integration of Hepatitis-B Virus-DNA into the Genome of Liver-Cells in Chronic Liver-Disease and Hepatocellular-Carcinoma - Studies in Percutaneous Liver Biopsies and Postmortem Tissue Specimens." New England Journal of Medicine **305**(18): 1067-1073.

Sherr, C. J. (1996). "Cancer cell cycles." Science **274**(5293): 1672-1677.

Shi, C. J., N. Fukushima, T. Abe, Y. S. Bian, L. Hua, B. J. Wendelburg, C. J. Yeo, R. H. Hruban, M. G. Goggins and J. R. Eshleman (2008). "Sensitive and quantitative detection of KRAS2 gene mutations in pancreatic duct juice differentiates patients with pancreatic cancer from chronic pancreatitis, potential for early detection." Cancer Biology & Therapy **7**(3): 353-360.

Slater, E. P., P. Langer, E. Niemczyk, K. Strauch, J. Butler, N. Habbe, J. P. Neoptolemos, W. Greenhalf and D. K. Bartsch (2010). "PALB2 mutations in European familial pancreatic cancer families." Clinical Genetics **78**(5): 490-494.

Smith, W. L., D. L. DeWitt and R. M. Garavito (2000). "Cyclooxygenases: Structural, cellular, and molecular biology." Annual Review of Biochemistry **69**: 145-182.

Steele, C., L. M. Cowsert and E. J. Shillitoe (1993). "Effects of Human Papillomavirus Type-18-Specific Antisense Oligonucleotides on the Transformed Phenotype of Human Carcinoma Cell-Lines." Cancer Research **53**(10): 2330-2337.

Stevenson, P. G., J. S. May, V. Connor and S. Efstathiou (2010). "Vaccination against a hit-and-run viral cancer." J Gen Virol **91**(Pt 9): 2176-2185.

Stewart and Wild (2014). World Cancer Report 2014. Lyon, France, International Agency for Research on Cancer.

Stimson, L., V. Wood, O. Khan, S. Fotheringham and N. B. La Thangue (2009). "HDAC inhibitor-based therapies and haematological malignancy." Annals of Oncology **20**(8): 1293-1302.

Straus, D. and F. M. Ausubel (1990). "Genomic subtraction for cloning DNA corresponding to deletion mutations." Proc Natl Acad Sci U S A **87**(5): 1889-1893.

Subramanian, C., J. S. Knight and E. S. Robertson (2002). "The Epstein Barr nuclear antigen EBNA3C regulates transcription, cell transformation and cell migration." Front Biosci **7**: d704-716.

Sung, W. K., H. C. Zheng, S. Y. Li, R. H. Chen, X. Liu, Y. R. Li, N. P. Lee, W. H. Lee, P. N. Ariyaratne, C. Tennakoon, F. H. Mulawadi, K. F. Wong, A. M. Liu, R. T. Poon, S. T. Fan, K. L. Chan, Z. L. Gong, Y. J. Hu, Z. Lin, G. Wang, Q. H. Zhang, T. D. Barber, W. C. Chou, A. Aggarwal, K. Hao, W. Zhou, C. S. Zhang, J. Hardwick, C. Buser, J. C. Xu, Z. Y. Kan, H. Y. Dai, M. Mao, C. Reinhard, J. Wang and J. M. Luk (2012). "Genome-

wide survey of recurrent HBV integration in hepatocellular carcinoma." Nature Genetics **44**(7): 765-U188.

Tai, P. C., F. M. Suk, W. H. Gerlich, A. R. Neurath and C. Shih (2002). "Hypermodification and immune escape of an internally deleted middle-envelope (M) protein of frequent and predominant hepatitis B virus variants." Virology **292**(1): 44-58.

Takakuwa, T., W. J. Luo, M. F. Ham, F. Sakane-Ishikawa, N. Wada and K. Aozasa (2004). "Integration of Epstein-Barr virus into chromosome 6q15 of Burkitt lymphoma cell line (Raji) induces loss of BACH2 expression." American Journal of Pathology **164**(3): 967-974.

Tan, T. M. and R. C. Ting (1995). "In vitro and in vivo inhibition of human papillomavirus type 16 E6 and E7 genes." Cancer Res **55**(20): 4599-4605.

Thorland, E. C., S. L. Myers, B. S. Gostout and D. I. Smith (2003). "Common fragile sites are preferential targets for HPV16 integrations in cervical tumors." Oncogene **22**(8): 1225-1237.

Thurau, M., G. Marquardt, N. Gonin-Laurent, K. Weinlander, E. Naschberger, R. Jochmann, K. R. Alkharsah, T. F. Schulz, M. Thome, F. Neipel and M. Sturzl (2009). "Viral Inhibitor of Apoptosis vFLIP/K13 Protects Endothelial Cells against Superoxide-Induced Cell Death." Journal of Virology **83**(2): 598-611.

Toyota, M., C. Ho, N. Ahuja, K. W. Jair, Q. Li, M. Ohe-Toyota, S. B. Baylin and J. P. J. Issa (1999). "Identification of differentially methylated sequences in colorectal cancer by methylated CpG island amplification." Cancer Research **59**(10): 2307-2312.

Tsai, W. L. and R. T. Chung (2010). "Viral hepatocarcinogenesis." Oncogene **29**(16): 2309-2324.

Tsurimoto, T., A. Fujiyama and K. Matsubara (1987). "Stable Expression and Replication of Hepatitis-B Virus Genome in an Integrated State in a Human Hepatoma-Cell Line Transfected with the Cloned Viral-DNA." Proceedings of the National Academy of Sciences of the United States of America **84**(2): 444-448.

Tzahar, E., J. D. Moyer, H. Waterman, E. G. Barbacci, J. Bao, G. Levkowitz, M. Shelly, S. Strano, R. Pinkas-Kramarski, J. H. Pierce, G. C. Andrews and Y. Yarden (1998). "Pathogenic poxviruses reveal viral strategies to exploit the ErbB signaling network." Embo Journal **17**(20): 5948-5963.

Vara, J. A. F., E. Casado, J. de Castro, P. Cejas, C. Belda-Iniesta and M. Gonzalez-Baron (2004). "PI3K/Akt signalling pathway and cancer." Cancer Treatment Reviews **30**(2): 193-204.

Varmus, H. and R. Swanstrom (1985). Replication of retroviruses. RNA tumor viruses. R. e. a. Weiss. New York, Cold Spring Harbor Laboratory**:** 75-134.

Vermeulen, K., D. R. Van Bockstaele and Z. N. Berneman (2003). "The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer." Cell Proliferation **36**(3): 131-149.

von Knebel Doeberitz, M. (2002). "New markers for cervical dysplasia to visualise the genomic chaos created by aberrant oncogenic papillomavirus infections." Eur J Cancer **38**(17): 2229-2242.

Wan, C., Y. C. Shen, T. Yang, T. Wang, L. Chen and F. Q. Wen (2012). "Diagnostic value of microRNA for pancreatic cancer: a meta-analysis." Archives of Medical Science **8**(5): 749-755.

Wang, H. C., W. Huang, M. D. Lai and I. J. Su (2006). "Hepatitis B virus pre-S mutants, endoplasmic reticulum stress and hepatocarcinogenesis." Cancer Sci **97**(8): 683-688.

Wang, J., F. Zindy, X. Chenivesse, E. Lamas, B. Henglein and C. Brechot (1992). "Modification of cyclin A expression by hepatitis B virus DNA integration in a hepatocellular carcinoma." Oncogene **7**(8): 1653-1656.

Welford, S. M., J. Gregg, E. Chen, D. Garrison, P. H. Sorensen, C. T. Denny and S. F. Nelson (1998). "Detection of differentially expressed genes in primary tumor tissues using representational differences analysis coupled to microarray hybridization." Nucleic Acids Research **26**(12): 3059-3065.

Wentzensen, N., R. Ridder, R. Klaes, S. Vinokurova, U. Schaefer and M. Doeberitz (2002). "Characterization of viral-cellular fusion transcripts in a large series of HPV16 and 18 positive anogenital lesions." Oncogene **21**(3): 419-426.

Wentzensen, N., S. Vinokurova and M. von Knebel Doeberitz (2004). "Systematic review of genomic integration sites of human papillomavirus genomes in epithelial dysplasia and invasive cancer of the female lower genital tract." Cancer Res **64**(11): 3878-3884.

Wies, E., Y. Mori, A. Hahn, E. Kremmer, M. Sturzl, B. Fleckenstein and F. Neipel (2008). "The viral interferon-regulatory factor-3 is required for the survival of KSHV-infected primary effusion lymphoma cells." Blood **111**(1): 320-327.

Williams, V. M., M. Filippova, U. Soto and P. J. Duerksen-Hughes (2011). "HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress." Future Virology **6**(1): 45-57.

Williamson, C. T., E. Kubota, J. D. Hamill, A. Klimowicz, R. Q. Ye, H. Muzik, M. Dean, L. R. Tu, D. Gilley, A. M. Magliocco, B. C. McKay, D. G. Bebb and S. P. Lees-Miller (2012). "Enhanced cytotoxicity of PARP inhibition in mantle cell lymphoma harbouring mutations in both ATM and p53." Embo Molecular Medicine **4**(6): 515-527.

Witter, R. L. and J. J. Solomon (1971). "Epidemiology of a herpesvirus of turkeys: possible sources and spread of infection in turkey flocks." Infect Immun **4**(4): 356-361.

Witter, R. L. and J. J. Solomon (1972). "Experimental infection of turkeys and chickens with a herpesvirus of turkeys (HVT)." Avian Dis **16**(1): 34-44.

Wolf, J., M. Pawlita, A. Jox, S. Kohls, S. Bartnitzke, V. Diehl and J. Bullerdiek (1993). "Integration of Epstein Barr virus near the breakpoint of a translocation 11;19 in a Burkitt's lymphoma cell line." Cancer Genet Cytogenet **67**(2): 90-94.

Wu, J., Y. Jiao, M. Dal Molin, A. Maitra, R. F. de Wilde, L. D. Wood, J. R. Eshleman, M. G. Goggins, C. L. Wolfgang, M. I. Canto, R. D. Schulick, B. H. Edil, M. A. Choti, V. Adsay, D. S. Klimstra, G. J. Offerhaus, A. P. Klein, L. Kopelovich, H. Carter, R. Karchin, P. J. Allen, C. M. Schmidt, Y. Naito, L. A. Diaz, Jr., K. W. Kinzler, N. Papadopoulos, R. H. Hruban and B. Vogelstein (2011). "Whole-exome sequencing of neoplastic cysts of the pancreas reveals recurrent mutations in components of ubiquitin-dependent pathways." Proc Natl Acad Sci U S A **108**(52): 21188-21193.

Wu, J., H. Matthaei, A. Maitra, M. Dal Molin, L. D. Wood, J. R. Eshleman, M. Goggins, M. I. Canto, R. D. Schulick, B. H. Edil, C. L. Wolfgang, A. P. Klein, L. A. Diaz, Jr., P. J. Allen, C. M. Schmidt, K. W. Kinzler, N. Papadopoulos, R. H. Hruban and B. Vogelstein (2011). "Recurrent GNAS mutations define an unexpected pathway for pancreatic cyst development." Sci Transl Med **3**(92): 92ra66.

Xing, L. and E. Kieff (2007). "Epstein-Barr virus BHRF1 micro- and stable RNAs during latency III and after induction of replication." J Virol **81**(18): 9967-9975.

Xu, B., S. Chotewutmontri, S. Wolf, U. Klos, M. Schmitz, M. Durst and E. Schwarz (2013). "Multiplex Identification of Human Papillomavirus 16 DNA Integration Sites in Cervical Carcinomas." Plos One **8**(6).

Yaginuma, K., H. Kobayashi, M. Kobayashi, T. Morishima, K. Matsuyama and K. Koike (1987). "Multiple integration site of hepatitis B virus DNA in hepatocellular carcinoma and chronic active hepatitis tissues from children." J Virol **61**(6): 1808-1813.

Yan, Z. J., K. R. Cui, D. M. Murray, C. Ling, Y. T. Xue, A. Gerstein, R. Parsons, K. J. Zhao and W. D. Wang (2005). "PBAF chromatin-remodeling complex requires a novel specificity subunit, BAF200, to regulate expression of selective interferon-responsive genes." Genes & Development **19**(14): 1662-1667.

Yoo, Y. D., H. Ueda, K. Park, K. C. Flanders, Y. I. Lee, G. Jay and S. J. Kim (1996). "Regulation of transforming growth factor-beta 1 expression by the hepatitis B virus (HBV) X transactivator. Role in HBV pathogenesis." J Clin Invest **97**(2): 388-395.

Young, L. S. and A. B. Rickinson (2004). "Epstein-Barr virus: 40 years on." Nat Rev Cancer **4**(10): 757-768.

Zerfass, K., A. Schulze, D. Spitkovsky, V. Friedman, B. Henglein and P. Jansendurr (1995). "Sequential Activation of Cyclin-E and Cyclin-a Gene-Expression by Human

Papillomavirus Type-16 E7 through Sequences Necessary for Transformation." Journal of Virology **69**(10): 6389-6399.

Zeschnigk, M., B. Horsthemke and D. R. Lohmann (1999). "Detection of homozygous deletions in tumors by hybridization of representational difference analysis (RDA) products to chromosome specific YAC clone arrays." American Journal of Human Genetics **65**(4): A331-A331.

zur Hausen, H. (1991). "Human papillomaviruses in the pathogenesis of anogenital cancer." Virology **184**(1): 9-13.

zur Hausen, H. (1999). "Viruses in human cancers." Eur J Cancer **35**(8): 1174-1181.

zur Hausen, H. (2000). "Papillomaviruses causing cancer: Evasion from host-cell control in early events in carcinogenesis." Journal of the National Cancer Institute **92**(9): 690-698.

Zur Hausen, H. (2001). "Oncogenic DNA viruses." Oncogene **20**(54): 7820-7823.

zur Hausen, H., H. Schulte-Holthausen, G. Klein, W. Henle, G. Henle, P. Clifford and L. Santesson (1970). "EBV DNA in biopsies of Burkitt tumours and anaplastic carcinomas of the nasopharynx." Nature **228**(5276): 1056-1058.