



**European School of Molecular Medicine**  
**University of Naples Federico II**



**HARVARD**  
**MEDICAL SCHOOL**



**Harvard Medical School**  
**Beth Israel Deaconesses Medical Center**



**PhD in Molecular Medicine XXV Cycle**  
*Curriculum Molecular Oncology*

***New pseudogenes with the function of competing  
endogenous RNA (ceRNA)***

**Tutor**

**Prof. Francesco Salvatore**

**External Supervisor**

**Prof. Pier Paolo Pandolfi**

**PhD student**

**Anna Ruocco**

**Internal Supervisor**

**Prof. Luigi Del Vecchio**

*“Ad maiora semper”*

## **Acknowledgements**

I would like to express my deepest gratitude to my tutor, Prof. Francesco Salvatore, for his precious guidance and advice he has constantly given to me during the PhD. Once more, and particularly “thanks”, for caring about me and for providing me with the best environment to do research.

My sincerest appreciation goes to Prof. Pier Paolo Pandolfi, my external tutor, for all I have learned from him and for always pushing me to improve. Also, I would like to thank him for encouraging me to express my ideas and share them with the others.

My deep gratitude and respect are addressed to Prof. Luigi Del Vecchio as well, who has let me experience the innovative and charming field of flow cytometry, with which I wasn't acquainted yet.

I would like to express my deepest appreciation to Florian A. Karreth, my bench guidance. With the everyday bench work he has let my passion for research grow. Furthermore, his suggestions and advice have been really invaluable to me.

My appreciation also goes to friends and co-workers, in Italy and in USA, who have always offered their support and advice in my decisions.

Finally, I would like to thank my family and beloved, for their love, for understanding, and for always believing in me.

## TABLE OF CONTENTS

<b>LIST OF ABBREVIATIONS</b>	<b>p. 7</b>
<b>FIGURES INDEX</b>	<b>p. 10</b>
<b>TABLES INDEX</b>	<b>p. 12</b>
<b>ABSTRACT</b>	<b>p. 13</b>

### 1. INTRODUCTION

#### 1.1 Competing endogenous RNA world **p. 15**

- 1.1.1 The dark matter of genome: non coding RNA (ncRNA)
- 1.1.2 Pseudogenes
- 1.1.3 microRNA
- 1.1.4 The ceRNA network (ceRNET)
- 1.1.5 Pseudogene as ceRNAs

#### 1.2 B-RAF **p. 25**

- 1.2.1 Gene and protein
- 1.2.2 B-RAF function: the RAF–MEK–ERK MAPK cascade
- 1.2.3 B-RAF in cancer
- 1.2.4 B-RAF pseudogene
- 1.2.5 B-RAF gene and pseudogene homology in mice and human
- 1.2.6 TRE-B-Rafps mice develop diffuse large B-cell lymphoma

#### 1.3 The Membrane Cofactor Protein: CD46 **p. 33**

- 1.3.1 Gene and protein
- 1.3.2 Expression pattern and its regulation
- 1.3.3 CD46 functions
  - 1.3.3.1 Complement
  - 1.3.3.2 Pathogen interactions

1.3.3.3	Fertilization	
1.3.4	CD46 role in cancer	
1.3.5	The CD46P1 pseudogene	
<b>2.</b>	<b>AIM OF THE STUDY</b>	<b>p. 40</b>
<b>3.</b>	<b>MATERIALS AND METHODS</b>	
3.1	Cell lines	p. 41
3.2	RNA	p. 42
3.2.1	RNA isolation and DNase treatment	
3.2.2	RNA quantification	
3.2.3	cDNA synthesis	
3.2.4	Reverse Transcriptase PCR	
3.2.5	Data Analysis	
3.3	DNA	p. 47
3.3.1	DNA isolation	
3.3.2	DNA quantification	
3.4	Cloning	p. 48
3.4.1	General condition	
3.4.2	Topo cloning	
3.4.3	Human and mouse BRAFps cloning	
3.5	Transfection	p. 50
3.6	Protein extraction and western blot analysis	p. 50
3.7	Proliferation Assay	p. 51
3.8	microRNA prediction	p. 52
3.9	Flow cytometry	p. 52
3.9.1	BD Lyoplate	
3.9.2	Cell line flow cytometry	

<b>4. RESULTS</b>	<b>p. 54</b>
<b>4.1</b> High-throughput screening of the prostate cancer cell line V-CaP by flow cytometry	
<b>4.2</b> Membrane complement regulatory proteins expression in prostate cancer cell lines	
<b>4.3</b> CD46 pseudogene expression in prostate cancer cell lines	
<b>4.4</b> CD46 pseudogene expression in other cancer cell lines	
<b>4.5</b> B-RAFps overexpression in mouse and human cell lines	
<b>4.6</b> B-RAFps sponging effect is Dicer/miRNA dependent	
<b>4.7</b> B-RAFps sponging effect is B-RAF dependent	
<b>5. DISCUSSION</b>	<b>p. 68</b>
<b>5.1</b> CD46 and its pseudogene CD46P1 in prostate cancer	
<b>5.2</b> B-RAF as competing endogenous RNA	
<b>6. FINAL OBSERVATIONS</b>	<b>p. 75</b>
<b>7. BIBLIOGRAPHY</b>	<b>p. 76</b>

## LIST OF ABBREVIATIONS

<b>22Rv1</b>	Human prostate carcinoma epithelial cell line
<b>AA</b>	Amino acid
<b>Ad-Cre</b>	Adenovirus- Cre recombinase
<b>Ad-Mock</b>	Control adenovirus for virus infection
<b>BD</b>	Becton Dickinson
<b>B-RAF</b>	human v-raf murine sarcoma viral oncogene homolog B
<b>B-Raf</b>	mouse v-raf murine sarcoma viral oncogene homolog B
<b>B-RAFps</b>	human v-raf murine sarcoma viral oncogene homolog B pseudogene
<b>B-Rafps</b>	mouse v-raf murine sarcoma viral oncogene homolog B pseudogene
<b>Bp</b>	Base pairs
<b>CD</b>	Cluster differentiation
<b>cDNA</b>	Complementary DNA
<b>CDS</b>	Protein coding sequence
<b>CR</b>	Conserved region
<b>CR1</b>	Complement receptor 1
<b>CR2</b>	Complement receptor 2
<b>CRD</b>	Cysteine-rich domain
<b>CYT</b>	Cytoplasmic domain
<b>DAF</b>	Decay-acceleration factor - CD59
<b>dNTP</b>	Deoxyribonucleotide
<b>DMEM</b>	Dulbecco's Modified Eagle Medium
<b>DNA</b>	Deoxyribonucleic acid
<b>Du-145</b>	Human androgen-independent prostate cancer line
<b>E2</b>	Mouse fibroblast B-Raf <sup>f1/f1</sup>
<b>ERK</b>	Extracellular signal-regulated proteins

<b>EtOH</b>	Ethanol
<b>FAM</b>	Fluorescein amidite
<b>FBS</b>	Fetal Bovine Serum
<b>Kb</b>	Kilobases
<b>KPCD</b>	Mouse Kras <sup>G12D</sup> ; p53 <sup>-/-</sup> ; MSCV-Cre; Dicer1 <sup>fl/-</sup> sarcoma cell line
<b>K-SFM</b>	Keratinocyte serum free medium
<b>LincRNA</b>	Long intergenic non coding RNA
<b>LINE</b>	Long interspersed elements
<b>LNCaP</b>	Human prostate adenocarcinoma cell line lymph node metastasis
<b>LncRNA</b>	Long non-coding RNA
<b>Non-LTR</b>	Non-long terminal repeat
<b>ncRNAs</b>	Non-coding RNAs
<b>MAC</b>	Membrane attack complex
<b>MAPK</b>	Mitogen-activated protein kinase
<b>Mdm2</b>	Murine double minute oncogene
<b>MOPS</b>	3-(N-morpholino) propanesulfonic acid
<b>MRE</b>	MicroRNA response elements
<b>Myc</b>	Myelocytomatosis viral oncogene homolog
<b>NC</b>	Negative Control
<b>NFQ</b>	Non-fluorescent quencher
<b>OD</b>	Optical density
<b>PC-3</b>	Human prostatic adenocarcinoma cell line
<b>qRT-PCR</b>	Quantitative Real Time-Polymerase chain Reaction
<b>RAS</b>	Ras sarcoma oncoproteins
<b>RBD</b>	RAS-binding domain
<b>RNA</b>	Ribonucleic acid
<b>Rpm</b>	Revolutions per minute

**Rt** Room temperature

**RT** Retro transcriptase

**SDS-PAGE** Sodium dodecyl sulfate -polyacrylamide gel electrophoresis

**V-CaP** Human vertebral-Cancer of the Prostate cell line

**VEGF** Vascular endothelial growth factor

## FIGURES INDEX

- Figure 1: Unitary pseudogene graphical representation** p.17  
Adapted from Sen K. and Ghosh T.C. “Pseudogenes and their composers: delving in the ‘debris’ of human genome” Briefing in functional genomics, 2013.
- Figure 2: Processed pseudogene graphical representation** p.18  
Adapted from Sen K. and Ghosh T.C. “Pseudogenes and their composers: delving in the ‘debris’ of human genome” Briefing in functional genomics, 2013.
- Figure 3: Duplicated pseudogene graphical representation** p.19  
Adapted from Sen K. and Ghosh T.C. “Pseudogenes and their composers: delving in the ‘debris’ of human genome” Briefing in functional genomics, 2013.
- Figure 4: Maturation process of miRNAs** p.21
- Figure 5: Competitive endogenous RNA network suggested dynamics** p.23  
Adapted from Karreth FA et Pandolfi PP, ceRNA cross-talk in cancer: when ce-bling rivalries go awry. Cancer Discov. 2013.
- Figure 6: Human B-RAF gene graphic structure** p.25
- Figure 7: Graphical comparison of RAF proteins structure** p.26
- Figure 8: RAS/RAF/MEK/ERK/MAPK pathway** p.27  
Hong Joo Kim & Dafna Bar-Sagi, “Modulation of signalling by Sprouty: a developing story” Nature Reviews Molecular Cell Biology 5, June 2004.
- Figure 9: Frequency B-RAF mutations in human cancers** p.29
- Figure 10: Total comparing MREs in B-RAF and B-RAFps** p.31
- Figure 11: H&E staining of kidney and lung (\*)** p.32
- Figure 12: Weight of control and B-Rafps mice (\*)** p.32
- Figure 13: Schematic structure of human CD46 gene** p.33  
Adapted from Yamamoto H. et al., CD46: The ‘multitasker’ of complement proteins”, The International Journal of Biochemistry & Cell Biology, 2009.
- Figure 14: Human CD46 isoforms** p.34  
Adapted from Ni Choileain S, Astier AL, CD46 processing: a means of expression, Immunobiology. 2012.

<b>Figure 15:</b> <i>CD46 protein mechanism of action</i> Adapted from Riley-Vargas RC et al., “CD46: expanding beyond complement regulation”, Trends in Immunology, 2004.	<b>p.36</b>
<b>Figure 16:</b> <i>CD46P1 pseudogene genomic locus</i>	<b>p.39</b>
<b>Figure 17:</b> <i>RT-PCR samples preparation</i>	<b>p.43</b>
<b>Figure 18:</b> <i>Syber Green RT-PCR program</i>	<b>p.45</b>
<b>Figure 19:</b> <i>TaqMan RT-PCR program</i>	<b>p.46</b>
<b>Figure 20:</b> <i>Gating strategy BD Lyoplate .fcs files</i>	<b>p.55</b>
<b>Figure 21:</b> <i>Graphical representation of cluster differentiation markers on V-CaP cell line after the BD Lyoplate</i>	<b>p.56</b>
<b>Figure 22:</b> <i>Flow cytometry dot plots mCRPs molecules</i>	<b>p.58</b>
<b>Figure 23:</b> <i>qRT-PCR CD46 and CD46P1 in prostate cancer cell lines</i>	<b>p.59</b>
<b>Figure 24:</b> <i>qRT-PCR CD46 and CD46P1 in lymphoma cell lines</i>	<b>p.60</b>
<b>Figure 25:</b> <i>Murine B-Rafps construct expression vector</i>	<b>p.61</b>
<b>Figure 26:</b> A) <i>Western Blot</i> B) <i>proliferation assay of NIH3T3 fibroblasts overexpressing B-Rafps (*)</i>	<b>p.62</b>
<b>Figure 27:</b> <i>Human B-RAFps cloning construct</i>	<b>p.62</b>
<b>Figure 28:</b> A) <i>Western Blot</i> B) <i>proliferation assay of PC-9 cell line overexpressing B-RAFps (*)</i>	<b>p. 63</b>
<b>Figure 29:</b> <i>Western blot A) HCT-116 B) KPCD overexpressing B-RAFps (*)</i>	<b>p.64</b>
<b>Figure 30:</b> <i>Proliferation assay A) HCT-116 B) KPCD overexpressing B-RAFps (*)</i>	<b>p.65</b>
<b>Figure 31:</b> <i>Conditional B-Raf knock-out fibroblasts (*)</i>	<b>p.66</b>
<b>Figure 32:</b> <i>Western blot B-Raf<sup>fl/fl</sup> fibroblasts overexpressing B-Rafps in presence or absence of Adeno-Cre infection (*)</i>	<b>p.67</b>

(\*) Karreth A. F., Reschke M., Ruocco A., Ng C., Ala U., Lèopold Valentine, Seitzer N., Langelotto F., Rodig S.J., Chiarle R., Pandolfi PP.; The BRAF Pseudogene is a proto-oncogenic competitive endogenous RNA (**under revision**).

## **TABLES INDEX**

<b>Table 1:</b> <i>Master Mix for cDNA synthesis</i>	<b>p.43</b>
<b>Table 2:</b> <i>Reverse transcription cycling program</i>	<b>p.44</b>
<b>Table 3:</b> <i>Primers sequences for RT-PCR</i>	<b>p.44</b>
<b>Table 4:</b> <i>PCR mix and Thermal Cycler program for CD46 and CD46P1 PCR</i>	<b>p.49</b>
<b>Table 5:</b> <i>Multicolor flow cytometry panel for mCRPs</i>	<b>p.57</b>

## ABSTRACT

Pseudogenes have been defined as non-functional sequences of genomic DNA originally derived from functional genes. Over the past decade pseudogenes have been suitably investigated and often exhibit functional roles, such as gene expression, gene regulation, and generation of genetic (antibody, antigenic, and other) diversity. In particular, recent studies showed that pseudogene can also act as microRNA-decoying competitive endogenous RNA and when deregulated can promote oncogenesis.

In this study, I investigated the role of B-RAFps as microRNA-decoying competitive endogenous RNA that is able to regulate BRAF expression and MAPK pathway activation. Furthermore, I participated in demonstrating that mice engineered to overexpress either the full-length B-Rafps or only its “CDS” or “3’ UTR” develop aggressive diffuse large B-cell lymphoma. I demonstrated *in vitro* that this effect is due to the microRNA-decoying competitive endogenous RNA potential of B-RAFps toward its coding gene BRAF. In this way B-RAFps is able to regulate BRAF expression and MAPK activation. These results were submitted for publication in an article by Karreth et al. **(103)** “The BRAF Pseudogene is a competitive endogenous RNA”.

In addition to the study of BRAFps I started to investigate the role of the CD46P1 pseudogene as ceRNA. Preliminary data show that CD46P1 is expressed in some prostate cancer and lymphoma cell lines. In particular, it seems expressed in cancer cell lines established from primary tumours and not from metastasis. It could be possible that the primary tumour microenvironment could influence the transcription activation of genes in the RCA cluster where CD46P1 and CD46 are located.

These preliminary data on CD46P1 might help to design and execute experiments that will elucidate the reasons for differential CD46P1 expression in cancer cell lines and its possible role as ceRNA.

## INTRODUCTION

### 1.1 The competing endogenous RNA world

#### 1.1.1 *The dark matter of the genome: non-coding RNA (ncRNA)*

In the past few years, thanks to new technologies such as the whole-genome and whole-transcriptome sequencing, it has become even more clear that only a small proportion of the mammalian genome (i.e. human, mouse), is transcribed into mRNAs which is subsequently translated into protein. In fact, less than 2% of our genome encodes protein-coding transcripts, even though over three quarters of the genome is transcribed into what was previously regarded as “dark matter”- non-coding RNA transcripts (ncRNAs) **(1)**.

It is increasingly evident that ncRNAs can and do have a very wide repertoire of biological functions in cellular physiology, development, metabolism, and that they are widely employed as a means of gene regulation especially in higher eukaryotes **(2)**.

The ncRNAs are a heterogeneous class of RNAs which are divided into two major categories on the basis of their average sizes: the small non-coding RNAs (sncRNAs) with a length less than 200 base pairs (bp) and the long non-coding RNAs (lncRNAs) with a length within a range of 200 bp to ~100 kilobases (kb).

The small ncRNAs include many different RNAs that are highly conserved in evolution and have their own peculiar functions. Some well-studied sncRNAs are: microRNAs (miRNAs), small interfering RNAs (siRNAs), PIWI interacting RNAs (piRNAs), transfer RNA (tRNA), small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) **(3)**.

The lncRNAs, instead, are poorly conserved among different species when compared to sncRNAs. They are usually low expressed making them look like

transcriptional noise. Despite this, a lot of evidence has accumulated showing that lncRNAs play a significant role in a wide variety of important biological processes (4, 5), including transcription, splicing, translation, protein localization, regulation of nuclear import (Nrn1) (6), cellular structure integrity, imprinting (H19) (7), cell cycle, apoptosis and X-chromosome inactivation (Xist) (8). In addition to this, it has been suggested that lncRNAs may regulate cancer progression (9) and development of many other human diseases (10).

The lncRNAs are sub-classified according to their genome localization (11). The large intergenic non-coding RNAs (lincRNAs) are distinct transcript units located in sequence space that do not overlap protein-coding gene. They have an average length of 1Kb, they are mainly transcribed by RNA Pol II from both strands, they are spliced and polyadenylated (12). A second group is made by pseudogenes which are considered “relics” of genes that have lost their coding potential due to mutation in their sequence. They can be actively transcribed but not translated into proteins (see paragraph). A third group of lncRNAs is the long intronic ncRNAs. They are encoded within the introns of protein-coding genes and for a long time they were recognized only as precursors of shorter RNAs (miRNA, snoRNAs etc.) (13). Recent evidences suggest that they have differential expression patterns, respond to stimuli, and can be misregulated in cancer (14). A final class of lncRNAs are the sense and antisense lncRNAs. The sense lncRNAs are transcribed from the sense strand of protein-coding genes, containing exons from protein-coding genes. They may overlap with part of protein-coding genes, or cover the entire sequence of a protein-coding gene. Antisense lncRNAs, instead, are transcribed from the antisense strand of protein-coding genes, overlapping with exonic or intronic regions or covering the entire protein-coding sequence through an intron (15).

### 1.1.2 Pseudogenes

Pseudogenes are sequences in the genome that have close similarities to one or more paralogous functional genes, but in general are unable to be transcribed (16). The non-functionality of the pseudogenes is often caused by the lack of functional promoters or other regulatory elements. As a result, these sequences are released from selection pressure and are free to accumulate non-gene-like features such as frame disruptions (frameshifts, in-frame stop codons, or disrupting interspersed repeats) in the original protein-coding sequence (CDS) (17). Currently, approximately twenty thousand pseudogenes are estimated which is comparable to the number of protein-coding genes in human (18).

On the basis of origin and characteristic features, the pseudogenes can be categorized as unitary pseudogenes, processed or retrotransposed pseudogenes and duplicated or non-processed pseudogenes.

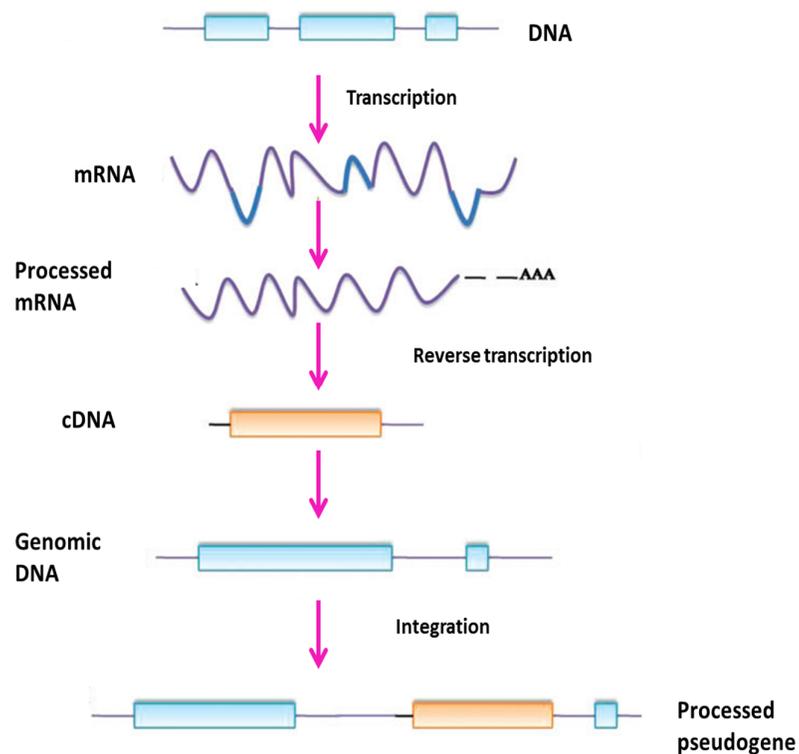
Unitary pseudogenes are like genetic “relics” of genes whose functions were important in ancestral species but became unnecessary in modern species. These pseudogenes are a natural consequence of mutations that failed to be eliminated by negative natural selection, because the functions of their products become unnecessary in present species (Figure 1).



**Figure 1:** Unitary pseudogene graphical representation.

More than 70 different unitary pseudogenes have been identified in the human genome in the last years (19). A classic example of a unitary pseudogene is the gene that coded for the enzyme L-gulono- $\gamma$ -lactone oxidase (GULO) in primates. In all mammals studied besides primates, GULO catalyses the terminal step of L-ascorbic acid biosynthesis: in human and other primates it exists only as a unitary pseudogene (GULOP) (20).

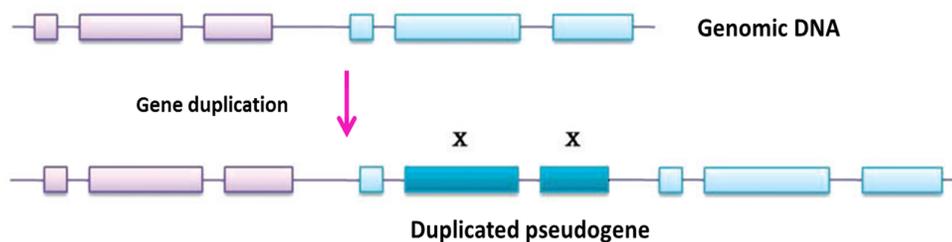
The second class of pseudogenes is the processed or retrotransposed pseudogenes. They originate through germline retroposition, where the processed transcript of a functional gene is reverse transcribed and incorporated into a staggered chromosome break, followed by DNA synthesis and repair (21) (Figure 2). Once these pseudogenes are inserted back into the genome, they usually contain a poly-A tail and lack introns and 5' promoter sequence (22).



**Figure 2:** *Processed pseudogene graphical representation.*

Processed pseudogenes expression is usually regulated by promoter and enhancer elements that differ from their parental genes. Several studies suggest that the enzymatic machinery that guides the process of reverse transcription and insertion of cDNA sequences are LINE1 non-LTR retrotransposons (23).

The third pseudogene category is the duplicated or non-processed pseudogenes. They usually originate from a gene duplication mechanism producing an extra copy of the gene which, being unnecessary, can accumulate mutations without damaging the organism, but they can also be generated by unequal crossing-over mechanisms (24) (Figure 3). Duplicated pseudogenes usually have all the same characteristics of genes, including an intact exon-intron structure and promoter sequences. They are often located next to their paralogous parent gene in clusters of similar functional sequences and even can be inserted into a different chromosome. Depending on the order in which mutations accumulate over evolutionary time, a duplicated pseudogene may still be transcribed and, even if rarely, could become a functional unit acquiring a novel function or mode of expression and become fixed within a population (25).

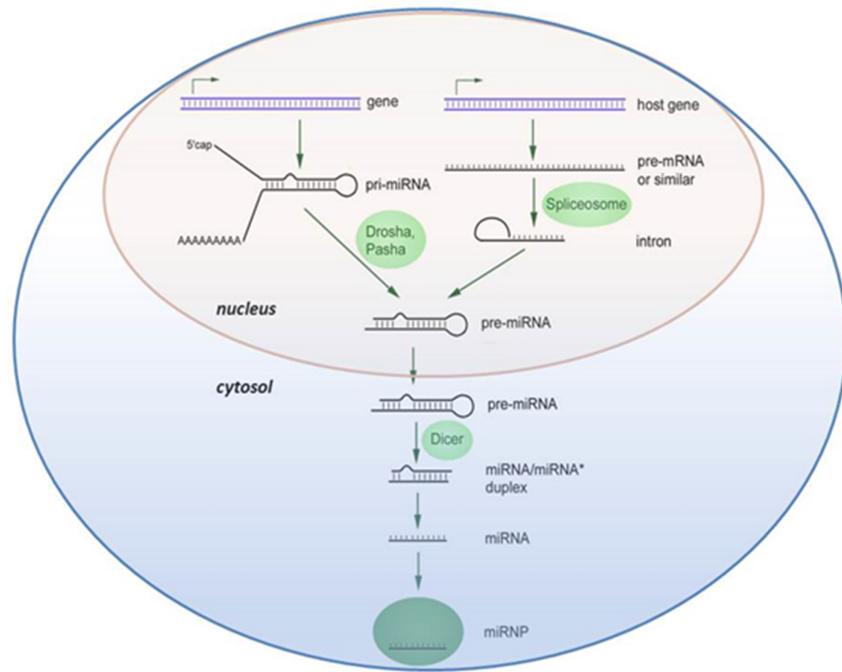


**Figure 3:** *Duplicated pseudogene graphical representation.*

### 1.1.3 *MicroRNAs*

MicroRNAs (miRNAs) comprise a large family of ~21-nucleotide-long ncRNAs that are highly conserved in the genome of viruses, plants, animals, and humans. Biogenesis of miRNAs takes place through a multi-step process that starts into the nucleus where the RNA polymerase II transcribes a hairpin loop of RNA (pri-miRNA) (26). The pri-miRNA is several kilobases long, polyadenylated at its 3' end and capped with a 7-methylguanosine cap at its 5' end. After the pri-miRNA transcript has been generated in the nucleus, it is cleaved into a precursor-miRNAs (pre-miRNAs) by Drosha, a RNA polymerase III. The pre-miRNA is then exported through the nuclear membrane protein exportin 5 into the cytosol (27). In the cytosol, the RNA polymerase III Dicer processes the pre-miRNAs into a ~20-bp miRNA/miRNA duplex containing a mature miRNA strand (leading strand) and its opposite complementary miRNA strand.

Only the leading strand is incorporated into the RNA-induced silencing complex (RISC), which directs the regulation of mRNA by recognizing the miRNA-response elements (MRE) (Figure 4). MREs are localized on the 3' untranslated region (UTR), coding sequence (CDS), and the 5' UTR of protein coding RNA, and, in addition, can also be found on non-protein coding transcripts such as long noncoding RNAs (28,29). The result of the binding of the miRNAs-RISC complex to a specific mRNAs usually results in the repression of target gene expression by either inhibiting translation or destabilizing the mRNA (30).



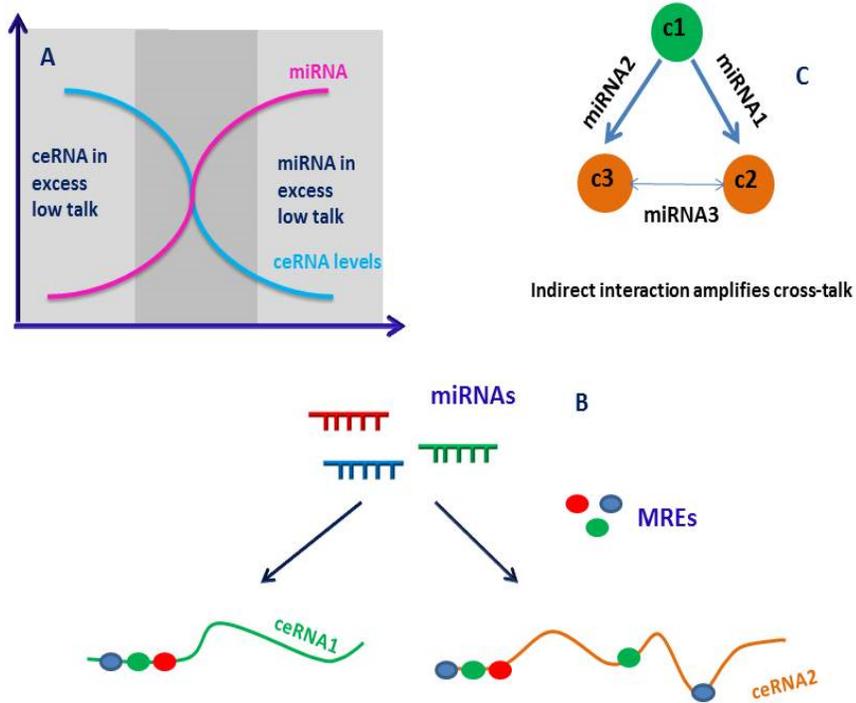
**Figure 4:** *Maturation process of miRNAs.*

The importance of miRNAs in normal cellular development and metabolism is now being realized. MiRNAs are involved in regulating many processes such as proliferation, differentiation, apoptosis and development (31, 32). Deficiencies or excesses have been correlated with a number of clinically important diseases ranging from myocardial infarction to cancers (33).

#### 1.1.4 *The ceRNA network (ceRNET)*

The competing endogenous RNA (ceRNA) hypothesis stems from the observation that RNA transcripts often share same microRNA response elements (MREs) for multiple distinct microRNAs and conversely, individual microRNAs often target at the same time multiple distinct transcripts (34). This target multiplicity has led to the hypothesis that different RNA transcripts, both protein-coding and non-coding, have the ability to compete for microRNA binding and co-regulate each other in complex ceRNA networks (ceRNETs) (35). In this “cross talk” MREs can be considered as the building blocks of a new language, the “RNA language”. Importantly, knowing the MRE overlap between multiple RNAs allows bioinformatic prediction of the interaction between RNAs, which are not yet correlated to each other (36).

In a recent paper, Ala et al. (37) define a precise set of rules to explain the “communication” within the ceRNA networks. Firstly they show that the optimal ceRNA-mediated cross-regulation occurs at a near equimolar equilibrium of all ceRNAs and miRNAs within a network (Figure 5-A). Secondly, they state that the relative abundance of miRNA and ceRNAs influences the cross-talk as does the number of shared MREs. In fact, the ceRNA cross-interaction increases with the numbers of shared miRNA and it is weakened when a ceRNA pair is targeted by too many non-shared miRNAs (38) (Figure 5-B). Thirdly the cross-talk is not limited to two transcripts which share MREs but can be amplified by the interaction with a third transcript (Figure 5-C).



**Figure 5:** Competitive endogenous RNA network suggested dynamics.

Changes in this cross-talk equilibrium may promote diseases like cancers. Misregulation of miRNAs, in fact, appears to play a fundamental role in the occurrence, growth and dissemination of many cancers (39).

### 1.1.5 Pseudogenes as ceRNAs

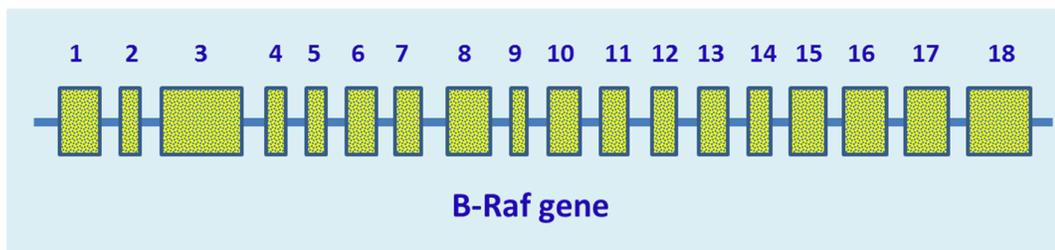
In 2010, the group of Pier Paolo Pandolfi hypothesized and demonstrated experimentally that pseudogenes, due to their high-sequence homology with their ancestor genes, can act as legitimate bona fide microRNA competitors (ceRNAs), thereby actively competing with their ancestral protein-coding genes for the same pool of microRNAs through sets of conserved MREs (35). These findings were particularly relevant because they attributed a new function to transcribed pseudogenes, which have been considered to be biologically useless owing to their inability to be translated into functional proteins. In particular, they showed this by studying the role of the PTEN pseudogene (PTENP1) as competing endogenous RNA for its protein-coding gene PTEN. PTEN is a haploinsufficient tumor suppressor that is commonly mutated or lost in tumours (35, 40). The 3' UTR regions of either the PTENP1 pseudogene and PTEN gene are highly homologous and contains numerous seed matches for miRNAs that were previously validated to repress PTEN expression. Overexpression and RNA interference silencing experiments confirmed that PTENP1 regulates PTEN thanks to the shared MREs (35, 40, 41).

A similar relationship was also observed between the pseudogene KRAS1P and its parental oncogene KRAS (35).

## 1.2 B-RAF

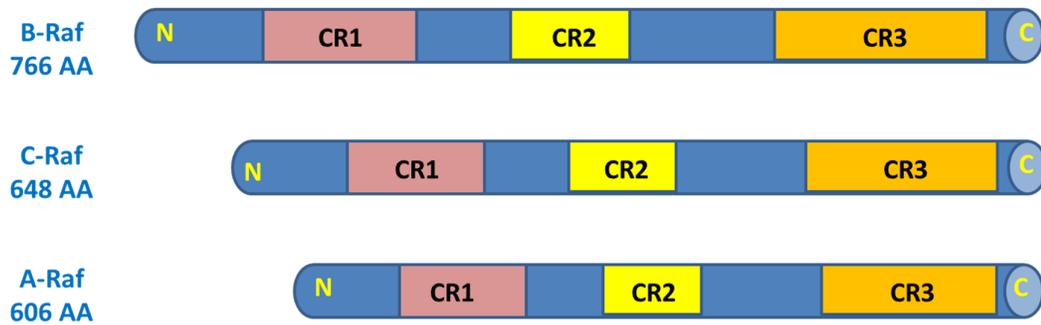
### 1.2.1 Gene and protein

The v-raf murine sarcoma viral oncogene homolog B (B-RAF) is located on the long (q) arm of chromosome 7 at position 34. The B-RAF gene is composed of 18 exons spanning a region of almost 200 kb (42) (Figure 6). The gene is transcribed into a transcript of 2478 bp, which is translated into a protein of 94-kDa.



**Figure 6:** Human B-Raf gene graphic structure.

B-RAF protein belongs to a family of serine-threonine protein kinases that includes A-RAF and C-RAF (Raf1). The three kinases have homology in their structure, in particular they share three conserved regions (CR): CR1, CR2, and CR3. CR1 is composed of a RAS-binding domain (RBD) and a cysteine-rich domain (CRD), which can bind two zinc ions. It is located near the N-terminus of the protein and it allows the interaction of RAF proteins with RAS. CR2 is a serine/threonine rich domain that contains a serine which when phosphorylated can bind to 14-3-3, an inhibitory regulatory protein. Finally, the CR3 region, located near the C-terminus, is the catalytic domain of the protein (43, 44) (Figure 7).



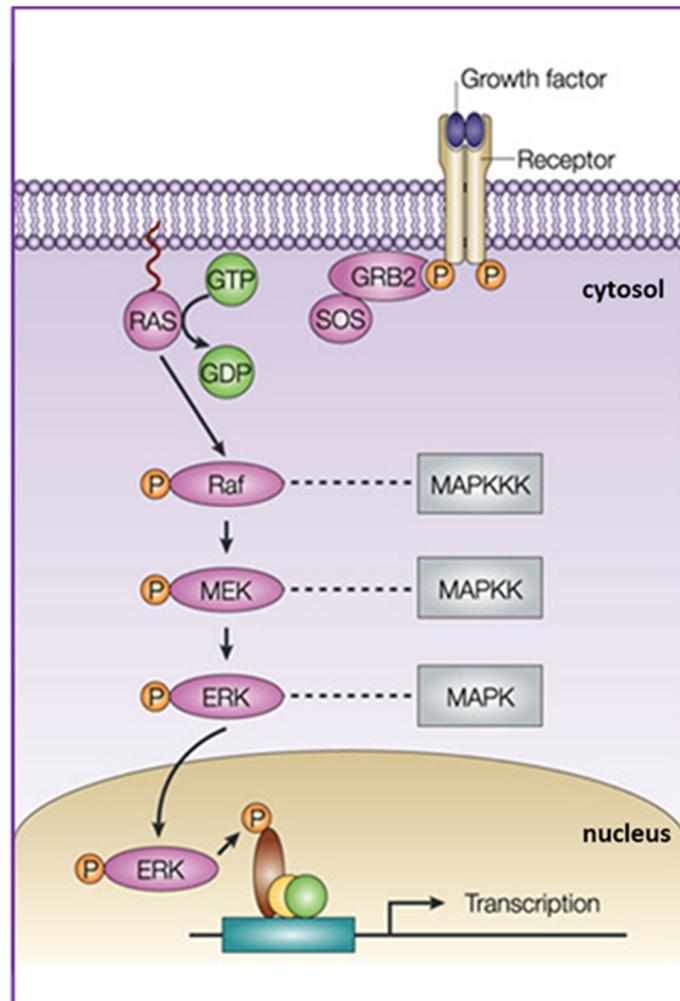
**Figure 7:** Graphical comparison of Raf proteins structure.

### 1.3.2 B-RAF function: the RAF–MEK–ERK MAPK cascade

B-RAF is a serine/threonine kinase that belongs to the RAS/RAF/MEK/ERK/MAPK pathway, which is involved in the transduction of mitogenic signals from the cell membrane to the nucleus (45). Binding of extracellular ligands such as growth factors, cytokines and hormones to their cell-surface receptors activates RAS-GTP protein, which resides at the inner side of the cell membrane.

When RAS proteins are activated, RAF proteins are able to recognize the phosphorylated sites and bind them with their RAS binding domain (RBD) in the CR1 region (46). The activation of RAF proteins lead to activation of the dual-specificity protein kinases MEK1 and MEK2 (mitogen-activated protein kinase -MAPK) through phosphorylation of two serine residues at positions 217 and 221 found in the activation loop (47). While RAF isoforms are enzymes of relatively low abundance, the high concentration of MEKs allows for amplification of signaling. The downstream targets of MEK proteins are the extracellular signal-regulated proteins ERK1 and ERK2.

Activated ERKs regulate growth factor-responsive targets in the cytosol and also translocate to the nucleus where they phosphorylate a number of transcription factors regulating gene expression. More of 160 proteins are described as downstream targets of the RAS/RAF/MEK/ERK/MAPK pathway (48).



**Figure 8:** *RAS/RAF/MEK/ERK/MAPK pathway*

Depending on the cellular context, this pathway mediates diverse biological functions such as cell growth, survival and differentiation predominantly through the regulation of transcription, metabolism and cytoskeletal rearrangements (43, 49).

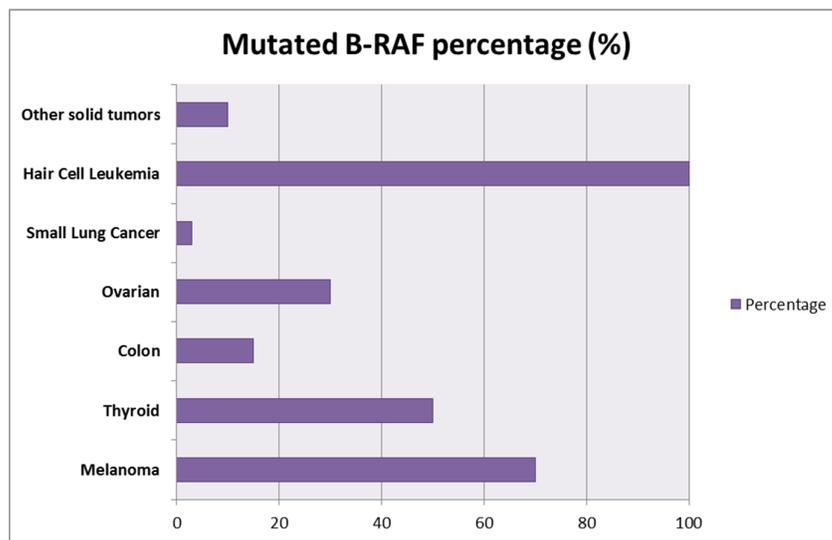
Alteration of one or more effectors of this pathway can deeply influence the cells fate. In fact, the RAS/RAF/MEK/ERK/MAPK pathway is one of the most frequently deregulated signal transduction pathways in human cancers, often through gain-of-function mutations of RAS and RAF family members **(50)**.

Some examples of genes activated by RAS/RAF/MEK/ERK/MAPK pathway that can promote cell transformation and metastasis are: cyclin D1, cyclin D2 and cyclin D3 (self-sufficiency in growth), VEGF (angiogenesis), c-myc (insensitivity to antigrowth signals), b3-integrin (tissue invasion and metastasis) and mdm2 (apoptosis evasion, limitless replicative potential and angiogenesis) **(51)**.

### *1.2.3 B-RAF in cancer*

B-RAF protein was the second RAF protein isoform to be identified and in contrast to the ubiquitous C-RAF expression was initially considered to be mainly a brain specific isoform **(52)**. For this reason the attention was focused for a long time on the study of the C-RAF isoform. This focus shifted however to some degree with the identification of B-RAF as a mutational target in various human cancers. In fact, in June 2002, Davies H. and colleagues **(53)** performed a sequencing screen of 923 cancer samples and identified missense mutations of the B-RAF gene in approximately 70% of human malignant melanomas **(54)**. Gradually, other studies revealed the B-RAF involvement in 50% of thyroid cancers **(55)** 15% of colorectal cancers **(56, 57)** and **35% of ovarian carcinomas (58)**.

Mutations were also detected at a low frequency in gliomas, lung, sarcomas, breast and liver cancers (59) (Figure 9).



**Figure 9:** Frequency B-RAF mutations in human cancers.

Most of the B-RAF mutations were identified within the kinase domain. The most frequent mutation, about 90% of all B-RAF mutations found in cancer, is a single amino acid substitution, a valine (V) being substituted for by glutamate (E) at codon 600 (V600E) (60). B-RAF<sup>V600E</sup> exhibits a 500-fold increase in kinase activity, constitutively stimulating the activation of MEK/ERK signalling in the absence of extracellular stimuli (61). Recently, Tiacci et al. using massively parallel sequencing observed that the V600E mutation is present in 100% of cases of hairy cell leukaemia (62).

Interestingly, a corresponding mutation to B-RAF<sup>V600E</sup> was not detected in C-RAF or A-RAF in human cancers. This is most likely due to the fact that C-RAF and A-RAF hold a tighter regulation of their kinase domain than B-RAF, thus, a single mutation is insufficient for substantial kinase activation (63).

Mutations in B-RAF commonly occur in the same cancer types that harbour RAS mutations in a mutually exclusive pattern with RAS mutations (< 1%), suggesting that these genetic alterations activate common downstream effectors of transformation **(69)**.

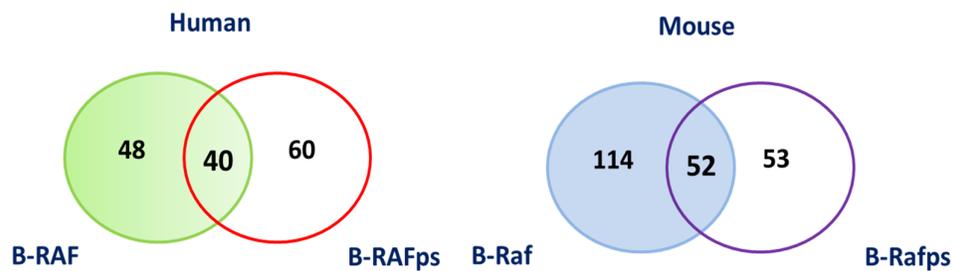
#### *1.2.4 B-RAF pseudogene*

Two human B-RAF loci (B-RAF-1 and B-RAF-2) have been identified in human genome. B-Raf-1 is located on chromosome 7q34 and encodes a functional B-Raf protein. B-RAF-2, instead, is located on chromosome Xq13 and has been predicted to be B-RAF-1 inactive unprocessed pseudogene (BRAFP1) **(42)**. The homology between the two transcripts is over 90%. The pseudogene contains 2 non-coding exons and one intron and it is transcribed in a transcript of 2,703 bps **(65)**.

In mice, B-RAF protein coding gene is located on chromosome 6 and is transcribed into two different transcripts, which give rise to two B-RAF isoform respectively of 804 AA and 750 AA. Also in mouse genome it has been identified a B-RAF pseudogene locus located on chromosome 10. It also shares a very high homology with its corresponding coding gene **(66)**.

### 1.2.5 *B-Raf* gene and pseudogene homology in mice and human

When analyzing the sequence of the B-RAF gene and its pseudogene and comparing those in either mice or human the sequence homology is > 90%. Karreth et al. in the paper under revision “The BRAF pseudogene is a proto-oncogenic competitive endogenous RNA” (103) shows that human B-RAF gene and pseudogene share 40 different miRNA response elements (MREs) while mouse B-RAF gene and pseudogene around 52 (Figure 10).

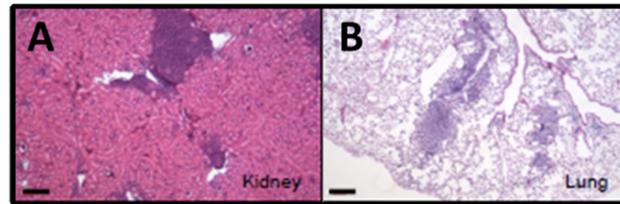


**Figure 10:** Total comparing MREs in *B-RAF* and *B-RAFps*.

The high number of MREs shared allows the transcripts to cross talk to each other following the ceRNA theory.

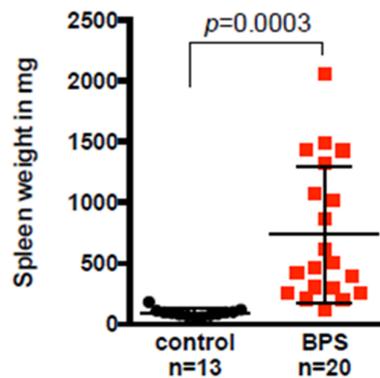
### 1.2.6 *TRE-B-Rafps* mice develop diffuse large B-cell lymphoma

Karreth et al. (103) report that mice engineered to overexpress either the full-length B-Rafps or only its “CDS” or “3’ UTR” developed aggressive diffuse large B-cell lymphoma. Lymphoma nodules were commonly observed in the kidneys, livers and lungs and histological analysis revealed microscopic organ infiltration (**Figure 11**)



**Figure 11:** *H&E staining of kidney and lung.*

In addition to these, all the *TRE-B-Rafps* mice showed splenomegaly and mild lymphadenopathy (**Figure 12**).



**Figure 12:** *Weight of spleen from control and B-Rafps mice.*

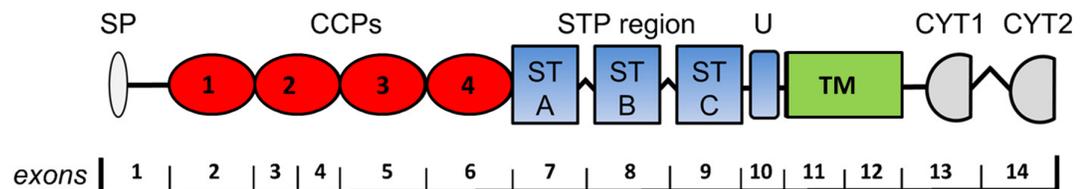
We demonstrated for the first time the oncogenic potential of a pseudogene in an engineered mouse model and indicate that ceRNA-mediated regulation is an important mechanism of gene expression *in vivo*.

### 1.3 The Membrane Cofactor Protein: CD46

#### 1.3.1 CD46: gene and protein

The Membrane Cofactor Protein (MCP)-CD46 gene is a member of the regulators of complement activation (RCA) family gene cluster located on the long (q) arm of chromosome 1 at position 32 (67). The RCA gene cluster contains more than 60 genes of which 15 are complement related genes. These genes code for both soluble and transmembrane proteins which play a pivotal role in regulating complement activity such as Factor H, the  $\alpha$ -chain of the C4b-binding protein (C4bp $\alpha$ ), the decay accelerating factor (DAF), the complement receptor type 1 (CR1) and the complement receptor type 2 (CR2) (68, 69).

The CD46 gene consists of 14 exons and 13 introns for a total length of ~43 kb (Figure 13).



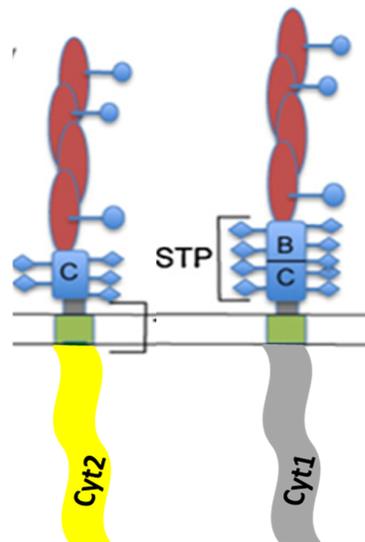
**Figure 13:** Schematic structure of human CD46 gene.

Exon 1 encodes the 5'UTR and signal peptide, exons 2-6 encode four short consensus repeat (SCRs), while exons 7, 8 and 9 form an area of high O-glycosylation rich in serine, proline and threonines (STP). Exon 10 encodes a 13-amino-acid stretch of unknown significance (UK) while exon 11 and 12 encode for the hydrophobic transmembrane domain (TM) and the basic amino acid anchor. Exon 13 is

alternatively spliced: if it is present a 93-bp stretch codes for the cytoplasmic domains of 16 amino acids (CYT-1), if it is absent the exon 14 encodes for a totally different cytoplasmic tail (CYT-2) of 23 amino acids. Due to the presence of an in-frame stop codon, the presence of exon 13 converts exon 14 into the 3'UTR. In the case exon 13 is spliced out exon 14 encodes for CYT-2 and the 3'UTR region of MCP (70).

### 1.3.2 Expression pattern and its regulation

CD46 is expressed on all human nucleated cells except on red blood cells. Four are the main isoforms generated by an alternative splicing. These isoforms differ in the quantity of juxtamembranous O-glycosylation regions and by expression of a 16 or 23 amino acid cytoplasmic tail (CYT-1 or CYT-2) (**Figure 14**).



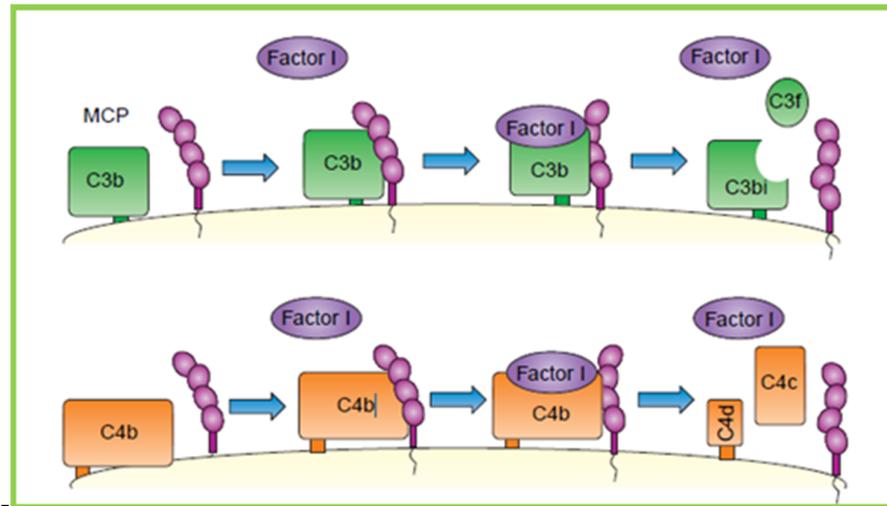
**Figure 14:** Human CD46 isoforms.

The relative ratio of the isoforms is inherited in an autosomal co-dominant way with three phenotypes in the population: 65% express predominantly the high O-glycosylated isoform, 6% express the less O-glycosylated and 29% express both forms with approximately equal ratio (71). In addition to this, the CD46 isoforms' expression pattern is tissue-specific: the less O-glycosylated form bearing the CYT-2 cytoplasmic tail is mostly expressed in male germ cells and brain while the higher O-glycosylated is expressed in kidney, the brain and salivary glands (72). While most cells express all four isoforms, some tissues are characterized by a distinct and specific CD46 expression pattern: brain, kidney and salivary glands, for instance, only express CYT-2-bearing isoforms (73).

#### *1.3.3.1 Complement*

The complement system can be activated via three different pathways: the classical pathway (activated by antibodies and immune complexes), the lectin pathway (triggered by pathogen-specific lectins) and the alternative pathway (initiated directly on microbial surfaces or triggered by properdin binding to targets). CD46 is most effective in controlling the amplification loop of the alternative pathway (74). In fact, CD46 is an intrinsic cofactor for the factor-I-mediated cleavage of C3b and C4b (75). The cleavage of C3b produces the fragment C3bi, which is unable to support further complement activation. In this way, C3b is not deposited onto cells surfaces and cannot induce the formation of the membrane attack complex (MAC) and subsequent lysis and/or the uptake of C3b-opsonized pathogens by macrophages and neutrophils.

The cleavage of C4b by factor I and CD46 produces the fragments C4c and C4d. C4c is liberated into the extracellular milieu whereas C4d, incapable of continuing the complement cascade, remains attached to the target (76) (Figure 15).



**Figure 15:** *CD46 protein mechanism of action.*

CD46 is not the only membrane-bound protein able to protect host cells from complement attack. Other proteins with similar function are CD35 (complement receptor type-1 (CR1)), CD55 (decay-accelerating factor (DAF)) and CD59. They differ from each other in their mechanism of action and in the way in which they attach to the cells surfaces. In particular, CD35 and CD55 contribute to the control of C3 activation while CD59 interferes with the assembly of the MAC at the terminal step of the complement activation cascade (77).

### *1.3.3.2 Pathogen interaction*

CD46 has been “labeled as a pathogen’s magnet” **(78)** due to the several pathogens for which it is a cell receptor. This includes the Edmonston strain of measles virus, human herpesvirus-6 (HHV-6), adenoviruses A and B, type IV pili of *Neisseria gonorrhoeae* and *Neisseria meningitidis* as well as group A streptococcus **(79-82)**.

Engagement of CD46 by pathogens has implications beyond the simple cell adhesion. Pathogens may use CD46 for activation of intracellular signaling cascades that give them a foothold in the host, by either facilitating the breach of protective host/environmental borders or creating an altered immune microenvironment, or both. For example the binding of measles virus (MV) to CD46 decreases IL-12 production in human macrophages **(83)** and alters internalization pathways of CD46 in non-lymphoid cell lines. In addition the MV attachment induces CD46 turnover resulting in enhanced susceptibility of the cell to complement-mediated lysis **(84)**.

### *1.3.3.4 Fertilization*

The research of a murine homologue of human CD46 revealed that murine CD46 gene is only expressed on spermatozoa **(85)**. This led to investigations into the role of CD46 in the human fertilization process. Many studies have revealed that CD46 plays a key role in egg-sperm fusion in human and mice **(86)**. In spermatozoa, CD46 is located on the inner acrosomal membrane where it is critical in the negative regulation of the acrosome reaction, a fusion of the plasma membranes of the sperm

and oocyte **(87)**. Lack or severely reduced expression of CD46 on spermatozoa has been connected with idiopathic male infertility **(88)**.

#### *1.3.4 CD46 role in cancer*

Trace of activated complement components (C3, C4) and the terminal MAC (C5b-C9) have been found deposited in tumor tissue such as breast **(89)** and papillary thyroid carcinoma **(90)**. In vitro studies of tumor cell lines have demonstrated that human tumor cells can activate to some degree human complement but are rather resistant to lysis **(91)**. Tumor cell resistance to complement-mediated lysis employs a collection of protective strategies and most of them are probably also utilized by normal tissues to resist accidental cell damage following local activation of complement **(92)**.

The resistance of tumor cells to complement-mediated cytotoxicity depends on basal or induced mechanisms. The basal mechanisms are spontaneously expressed in cells without a need for prior activation. An example is the membrane-associated complement regulatory proteins (CD46, CD55 and CD59) which are normally expressed on normal cells but appear to be over-expressed on tumors cells **(93)**. Numerous studies have also shown that the overexpression of these proteins plays an important role in cancer immunoresistance. In fact in order to succeed in tumor eradication with antibody and complement, novel methods and reagents that overcome the protective capacity of mCRPs are needed such as specific inhibition of mCRP activity achieved with monoclonal antibodies directed to CD46, CD55 or CD59 **(94)**.

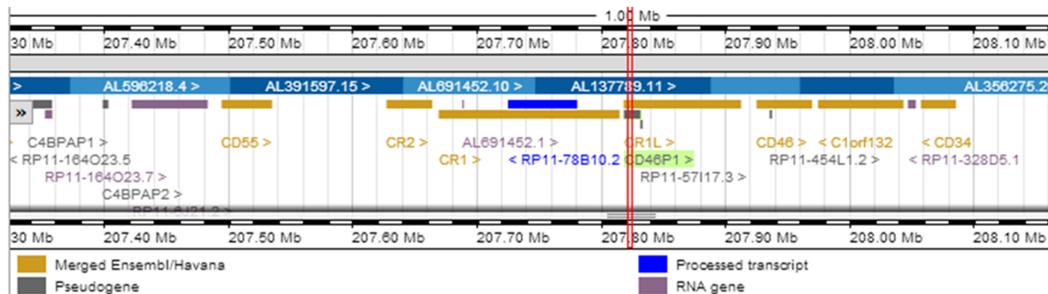
Furthermore, tumor cells can also secrete several soluble complement inhibitors in the microenvironment, such as C1 inhibitor, factor H and factor H-like protein (95) and may also express proteases that degrade complement proteins, such as C3. Besides this basal resistance, nucleated cells resist complement damage by removing the membrane attack complexes (MAC) from their surface. Calcium ion influx and activation of protein kinase C (PKC) (96) and of mitogen-activated protein kinase (MAPK) have been demonstrated to be associated with the complement-induced enhanced resistance to lysis (97).

### 1.3.5 The CD46P1 pseudogene

The CD46P1 pseudogene (Gene ID: 4182) was discovered in 1992 when the human regulators of complement activation gene cluster (RCA cluster) was partially characterized with yeast artificial chromosomes (YACs) (98).

It is an unprocessed pseudogene located on the long (q) arm of chromosome 1 at position 32 very close to its coding- CD46 gene locus (Figure 16).

The CD46P1 pseudogene has a total length of ~12.18 kb.



**Figure 16:** CD46P1 pseudogene genomic locus.

## **2. Aims of the study**

This study has been conducted in Italy at the CEINGE-Biotecnologie Avanzate of Naples, and in USA at Harvard Medical School-Beth Israel Deaconesses Medical Center in Boston.

The main aim of this study was to investigate the role of long non-coding RNAs (lncRNAs), in particular pseudogenes, as competing endogenous RNAs (ceRNAs) in cancer. Thanks to recent discoveries, it has become clear that two different transcripts can cross-talk to each other by sharing the same microRNA responsive elements (MREs).

In particular, I focused my attention on the two pseudogenes BRAFps and CD46P1. Their deregulation could affect cell physiology and could lead to oncogenesis.

### 3. Materials and Methods

#### 3.1 Cell lines

NIH3T3 and HCT116 cells lines were purchased from ATCC, Dicer mutant HCT116 cells were provided by B. Vogelstein and  $Dicer^{flox/delta}$  and  $Dicer^{delta/delta}$  mouse sarcoma cells were provided by P. Sharp (KPCD cell line). The prostate cancer cell lines Du-145, PC-3, LNCaP, V-CaP, 22Rv1 were obtained by CEINGE-Biotecnologie Avanzate cell bank (Naples, Italy). PC-9 cells were provided L. Cantley (Harvard Medical School, Boston, USA) and Ly8 cells were provided by R. Chiarle. NIH3T3, HCT116, Dicer mutant HCT116,  $Dicer^{flox/delta}$  and  $Dicer^{delta/delta}$  mouse sarcoma cells, E2 B-Raf<sup>fl/fl</sup> LNCaP, DU-145, PC-3, V-CaP cell lines were grown in Dulbecco's Modified Eagle Medium (DMEM, Sigma-Aldrich, Oakville, ON) supplemented with 10% Fetal Bovine Serum (FBS, Gibco, Carlsbad, CA), 2mM L-glutamine (Cambrex, East Rutherford, NJ) . PC-9 and Ly8 cells were grown in RPMI-1640 containing 10% DMEM and 2mM L-glutamine.

Cells were regularly tested with MycoAlert (Lonza) to ascertain that cells were not infected with mycoplasma and incubated at 37°C in a humidified atmosphere containing 5% CO<sub>2</sub>. The cells were fed fresh growth medium every 2-3 days and at confluence they were subcultured using trypsin-EDTA (0.05%, 0.02%).

## 3.2 RNA

### 3.2.1 RNA isolation and DNase treatment

Total RNA was extracted by using TRIzol Reagent (Ambion, USA). 500  $\mu$ L of TRIzol was added to cells pellets (around  $3 \times 10^6$  cells for sample) and incubated for 5 minutes at room temperature (Rt). 100  $\mu$ L of Chloroform was added to each sample, shaken vigorously and rested for 3 minutes. After centrifugation at 13,000 rpm for 15 minutes at 4°C the aqueous phase was transferred in a new Eppendorf tube and mixed with an equal amount of isopropanol. The samples were left to rest at Rt temperature for 10 minutes and then were centrifuged at 13,000 rpm at 4°C for 10 minutes. The RNA pellet was washed once with 250  $\mu$ L of ethanol 70% and then dried 15 minutes at Rt. The RNA pellet was re-suspended in 50  $\mu$ L ribonuclease-free water.

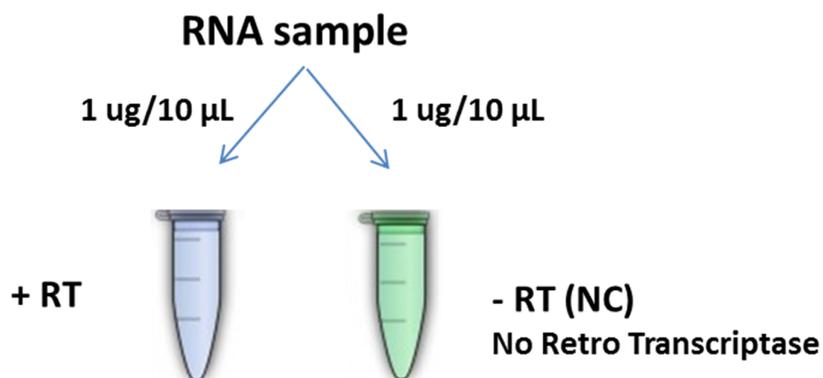
In order to eliminate any trace of DNA, RNA was treated with RQ1 RNase-Free DNase kit (Promega, USA). Briefly, the 50  $\mu$ L of RNA were incubated for 40 minutes at 37°C with a mix of: 6  $\mu$ L of 10X RQ1 DNase Buffer, 3  $\mu$ L of RQ1 RNase-Free DNase and 2  $\mu$ L RNasin® Plus RNase Inhibitor (Promega, USA). After the DNase treatment the RNA was cleaned up by using the Pure Link RNA mini kit (Ambion), according to manufacturer's recommendation. All RNA samples were eluted in a final volume of 30  $\mu$ L of RNase-Free water and stored at -20°C.

### 3.2.2 RNA quantification

RNA samples were quantified by using a NanoDrop 2000 spectrophotometer (Thermo Scientific) according to manufacturer's recommendations.

### 3.2.3 cDNA synthesis

Complementary DNA (cDNA) was synthesized using the High Capacity cDNA Reverse Transcription Kit (Applied BioSystems, California, USA). cDNA synthesis was performed in duplicate: +RT reaction with Retro Transcriptase enzyme (RT) and a -RT-negative control reaction with no enzyme (**Figure 17**).



**Figure 17:** RT-PCR samples preparation.

The cDNA reaction volume is of 20 µL and it is made up by 1ug RNA in 10 µL of Nuclease-free H<sub>2</sub>O plus 10 µL of master mix +RT or -RT:

Master mix +RT for 1 sample		Master mix -RT for 1 sample	
Components	Volume µL	Components	Volume µL
10X RT Buffer	2.0	10X RT Buffer	2.0
25X dNTP mix (100nM)	0.8	25X dNTP mix (100nM)	0.8
10X RT Random Primers	2.0	10X RT Random Primers	2.0
Reverse Transcriptase	1.0	-	-
Nuclease-free H <sub>2</sub> O	4.2	Nuclease-free H <sub>2</sub> O	5.2

**Table 1:** Master Mix for cDNA synthesis.

The cDNA synthesis reaction was performed using MJ Research PTC-200 PCR Thermal Cycler following these thermal conditions:

	<b>Step1</b>	<b>Step 2</b>	<b>Step 3</b>	<b>Step 4</b>
<b>Temperature</b>	25°C	37°C	85°C	4°C
<b>Time</b>	10 min	120 min	5 min	∞

**Table 2:** *Reverse transcription cycling program.*

After the cDNA synthesis reaction 80  $\mu$ L of Nuclease-free H<sub>2</sub>O was added to each samples. The cDNA products were stored at -20°C for future uses.

### 3.2.4 Reverse Transcriptase PCR

#### Syber Green

Real-time PCR was performed using the Applied Biosystems StepONE Plus PCR machine and SYBR® Select Master Mix (Applied BioSystems, California, USA). For one single reaction were used 5  $\mu$ L of 2X SYBR® Select Master Mix, 4  $\mu$ L of cDNA and 1  $\mu$ L of Primer Mix 1:1 ratio. The primers stock concentration was 25  $\mu$ M.

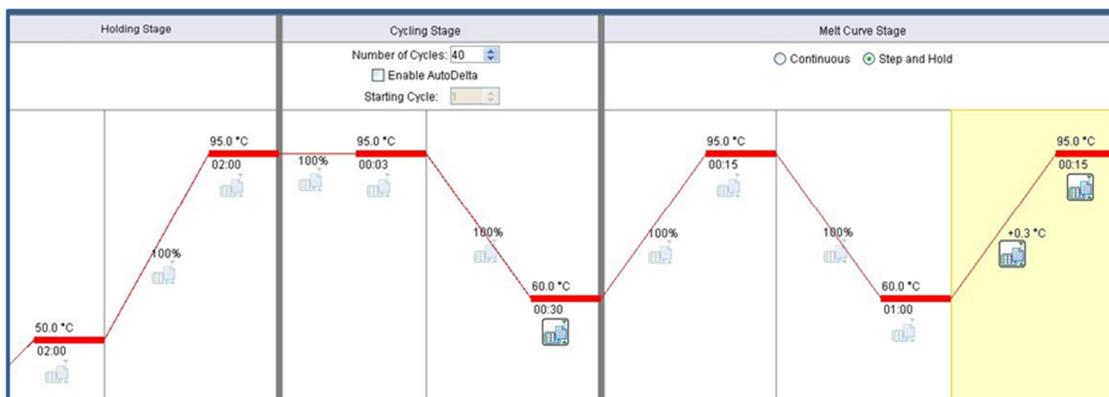
All primers were ordered from Invitrogen™ Custom DNA Oligos service. They were designed using the bioinformatics tool Primer3 (<http://frodo.wi.mit.edu>).

<b>Name</b>	<b>Orientation</b>	<b>Sequence 5'-3'</b>
<b>CD46</b>	F	GTCCAGTGCCTCAGGTCCTA
<b>CD46</b>	R	TGACCCAAACATCCAAACTG
<b>CD46P1</b>	F	GGACCAGATCCATTTTCACTT
<b>CD46P1</b>	R	ACATACACCTGCTTTATTTG
<b>Tubulin</b>	F	CAGTTTGTGGATTGGTGCCC

<b>Tubulin</b>	R	CAATGGCCGTGGTGTTC
<b>Murine BRAF</b>	F	GGCCTATGAAGAGTACAC
<b>Murine BRAF</b>	R	GGAGGATGTAACGGTGTCC
<b>Murine BRAFps</b>	F	TGGCCAACCAGACCTGTCCTCA
<b>Murine BRAFps</b>	R	GTTGACCCTCCATCACCACAAATT
<b>Murine <math>\beta</math>-Actin</b>	F	CGTCGACAACGGCTCCGGCA
<b>Murine <math>\beta</math>-Actin</b>	R	TGGGCCTCGTCACCCACATAGG

**Table 3:** Primers sequences for RT-PCR.

The following thermal conditions were used (**Figure 18**):



**Figure 18:** Syber Green RT-PCR program.

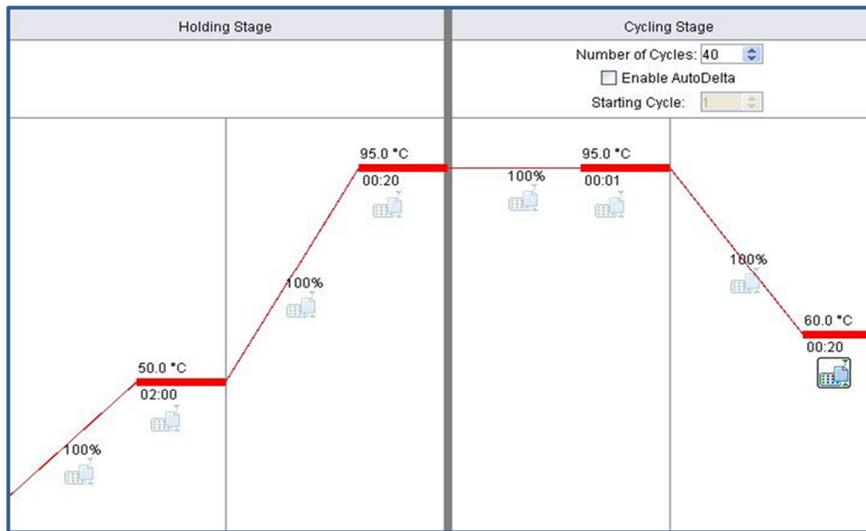
Each sample reaction was performed in triplicates.

### TaqMan Probe

Real-time PCR with TaqMan Probe was performed using the Applied Biosystems StepONE Plus PCR machine and TaqMan® Fast Universal PCR Master Mix (Applied BioSystems, California, USA). For one single reaction were used 5  $\mu$ L of 2 X TaqMan® Fast Universal PCR Master Mix, 4.5  $\mu$ L of cDNA and 0.5  $\mu$ L of

TaqMan Probe. The TaqMan probes were ordered from Invitrogen™ Custom DNA Oligos service. The TaqMan MGB probes were customized with FAM at 5' as reporter dye and NFQ at 3' as non-fluorescent quencher.

The following thermal conditions were used (**Figure 19**):



**Figure 19:** *TaqMan RT-PCR program.*

Each sample reaction was performed in triplicates.

### 3.2.5 Data Analysis

Excel and StepOne Software were used for graphing, correlation and statistical analyses. Data were analyzed using unpaired Student's t test and Mantel-Cox test. Values of  $p < 0.05$  were considered statistically significant. The mean  $\pm$  standard deviation of three or more independent experiments performed in triplicates is reported.

### 3.3 DNA

#### 3.3.1 DNA isolation

After processing of TRIzol preparations for the extraction of RNA (see 2.1) the remaining interphase and organic phase were stored at 4°C until DNA extraction. In order to extract DNA 300 µL of 100% EtOH was added to each tube and mixed for inversion several times at Rt. After 3 minutes the samples were spun at 5,000 x g for 5 min at + 4°C. The supernatants containing the proteins fractions were removed carefully and the DNA pellets were washed with 500 µL of sodium citrate/ ethanol solution (0.1 M sodium citrate in 10% ethanol, pH 8.5). The samples were left to rest for 30 min at Rt. The samples were centrifuged at 2,000 x g for 5 min a +4°C and the supernatants were removed gently. To each DNA pellet was added 1.5 mL of 75% EtOH. After 20 min of incubation the samples were centrifuged at 12,000 x g for 5 min. The supernatant of each sample was removed and the DNA pellets were dried at Rt for 10 min. The DNA pellets were resuspended in 500 µL of Nuclease-free water. The samples were stored at -20°C.

#### 3.3.2 DNA quantification

DNA samples were quantified by using a NanoDrop 2000 spectrophotometer (Thermo Scientific) according to manufacturer's recommendations.

### **3.4 Cloning**

#### *3.4.1 General Conditions*

The bacteria used for transformation reaction were One Shot® Chemically Competent *E. coli* (Invitrogen, CA, USA). Each bacteria vial was left to thaw on ice for 10 minutes. To each vial were added 2  $\mu\text{L}$  -3  $\mu\text{L}$  of ligation reaction and mixed gently. A negative control of ligation reaction was always set up. The bacteria were left to rest for 5 minutes on ice and then were heat-shocked for 30 seconds at 42°C without shaking. After the heat-shock the bacteria were immediately transferred to ice. 250  $\mu\text{L}$  of room temperature SOC medium was added to each bacteria vial and let shake (300 rpm) at 37°C for 1 hour. The bacteria were spun for 1 minute at 4000 rpm and the pellet was dissolved in 200  $\mu\text{L}$  of SOC medium. The bacteria were spread on pre-warmed LB- kanamycin or ampicillin selective plates and incubated overnight at 37°C. The day after, the positive colonies were picked and prepared for an overnight pre-culture with 2 ml of LB-Broth medium with antibiotics. The day after, 1 ml of each pre-culture was collected for DNA isolation using Plasmid Mini Purification kit (QIAGEN). The plasmid positivity was firstly controlled by restriction digestion to identify positive clones. Only the positive pre-cultures were prepared for an overnight culture with 250 ml of LB-Broth medium with antibiotics. The plasmid purifications were performed using the standard protocol of the Plasmid Midi Purification kit (QIAGEN). The plasmid sequences were controlled by Sanger sequencing using the GENEWIZ, Inc. - DNA Sequencing Services (Cambridge, MA, USA).

### 3.4.2 Cloning CD46 and CD46P1

In order to clone the exon 5 of CD46 gene and CD46P1 pseudogene into TOPO pcDNA 3.1 expression vector each sequence was amplified by using Phusion® High-Fidelity DNA Polymerase (ThermoScientific). The following protocol was used:

PCR Mix		Thermocycler program	
Components	Volume $\mu\text{L}$		
Phusion Polymerase	0.2	<b>Step 1</b>	98°C x 30 sec
Primer F	1	<b>Step 2</b>	98°C x 10 sec
Primer R	1		68°C x 30 sec
			72°C x 1 min
25X dNTP mix (100nM)	0.4		35 cycles
DMSO	0.6	<b>Step 3</b>	72°C x 10 min
5X HF buffer	4	<b>Step 4</b>	4°C
Nuclease-free H <sub>2</sub> O	12.8		

**Table 4:** PCR mix and Thermal Cycler program for CD46 and CD46P1 PCR.

According to the Invitrogen's TOPO TA Cloning System kit instructions 2  $\mu\text{L}$  of fresh PCR product was used for the ligation mix into TOPO-Cloning Vector. The reaction was performed at Rt for 5 minutes then the reaction was placed on ice.

2  $\mu\text{L}$  of the TOPO Cloning reaction was added into a vial of One Shot® Chemically Competent E. coli (Invitrogen) and mixed gently. The mix was incubated on ice for 10 minutes before the cells were heat-shocked for 30 seconds at 42°C without shaking. The cells were then immediately transferred to ice. 250  $\mu\text{L}$  of room temperature SOC medium was added to each bacteria vial and let shake (300 rpm) at

37°C for 1 hour. 10 µL from each transformation was spread on a pre-warmed kanamycin selective plate and incubate overnight at 37°C.

### *3.4.3 Human and mouse BRAFps cloning*

Mouse BRAFps was cloned into pCCL.sin.PPT.hPGK.GFP.Wpre kindly given by L. Naldini from San Raffaele-Telethon Institute for Gene Therapy, Milano, Italy. The same sequence has been cloned into pcDNA3-neo vector obtained by Invitrogen. Human BRAF pseudogene was cloned into pLenti-CMV-GFP-Puro (Addgene 25873).

## **3.5 Transfection**

Plasmid transfection, cells were seeded in 12-well dishes at a density of 120,000 cells per well. Transfection was performed 24 hours later with Lipofectamine 2000 (Life Technologies) according to the manufacturer's recommendations. Cell Medium was changed after 4 hours of transfection. Cells were trypsinized and seeded for the various assays 8 hours post-transfection.

## **3.6 Protein extraction and western blot analysis**

For protein extraction cells were washed twice with cold PBS then scraped into 500 µL of PBS and centrifuged at 13.000 rpm for 5 minutes at 4°C. The PBS was removed and the cell pellets were lysed with 100 µL of 1X RIPA lysis buffer

containing 1X HALT protease and phosphatase inhibitors on ice for 10 minutes. The lysates were centrifuged at 13,000 rpm for 5 minutes at 4°C. The supernatant was collected and stored at -20°C for future uses.

For western blot analysis, 5-20 µg of total protein was size-fractionated by SDS-PAGE on 4-12% Bis-Tris acrylamide NuPAGE gels in MOPS-SDS running buffer (Invitrogen). The proteins were transferred to nitrocellulose membranes (Whatman) in NuPage transfer buffer (Invitrogen) containing 10% methanol. The membranes were then probed with specific primary antibodies. HSP90 (610419, BD), BRAF (sc5284, Santa Cruz) pERK (9101, Cell Signaling) and tERK (9102, Cell Signaling). Secondary HRP-tagged antibodies and ECL detection reagent were from Amersham.

### **3.7 Proliferation Assay**

The proliferation assays were performed by seeding 3 separate 12-well plates at a final density of  $20 \times 10^4$  cells for well. Starting from the following day (d0), one plate per day was washed once with PBS, fixed in 4% paraformaldehyde for 10 min at room temperature. Cells were washed one time with PBS and then kept in PBS at 4°C. On the last day, all the well cells were stained with crystal violet solution (0.5% crystal violet in 20% methanol) for 10 min. Unbound crystal violet was removed by rinsing with distilled water, and cells were subsequently air dried. Next, crystal violet was extracted from cells with 500 µL of 10% acetic acid solution. 100 µL of crystal violet-acetic acid solution was used for absorbance measurement with the GloMax®-Multi Detection System - Promega at optical density (OD) of 595 nm.

### **3.8 microRNA prediction**

MicroRNA target prediction was performed using TargetScan release 6.2 (June 2012). The tools if freely available on the web at the following web-address:

[http://www.targetscan.org/vert\\_61/docs/help.html](http://www.targetscan.org/vert_61/docs/help.html)

### **3.9 Flow Cytometry**

#### *3.9.1 BD Lyoplate*

V-CaP cell line was characterized by using a panel of 242 purified monoclonal antibodies (BD Lyoplate™ Human Cell Surface Marker Screening Panel). According to the manufacture instructions the lyophilized antibodies of the kit were reconstituted by adding 110 µL of 1X sterile PBS to each well of 96well/plates. The reconstituted antibodies were stored at 4°C and used for screening within 10 days. Each antibody was incubated with  $1 \times 10^5$  V-CaP according the kit guidelines. Data acquisition (10,000 gated events) was performed using a Cytomics FC 500 flow cytometer (Beckman Coulter) after performing an instrument quality control procedure. Experimental panels and data analyses were performed using the Kaluza (Beckman Coulter) and BD FACSDiva (Becton Dickinson)

### 3.9.2 Cell line flow cytometry

Adherent cells lines were trypsinized and centrifuged at 1200 rpm for 4 minutes. Then the cell pellet was washed with 1 ml of FACS Flow solution. Then  $2 \times 10^5$  cells were incubated at 4°C for 30 min with the appropriate amount of MoAbs, following the manufacturer's instructions. The Becton Dickinson (BD) MoAbs used are: CD46-FITC (cat.n.555949), CD35-FITC (cat.n.555452), CD55-PE (cat.n.555694), CD59-PE (cat.n.555764), CD49f-PE (cat.n.555736). CD44-PE-Cy7 was purchased by eBioscience (cat.n. 25-0441-82).

The cell staining was analysed using an unmodified BD FACSCantoII flow cytometer (BD) and compensation was set in FACS-DiVa (BD) software.

## 4. RESULTS

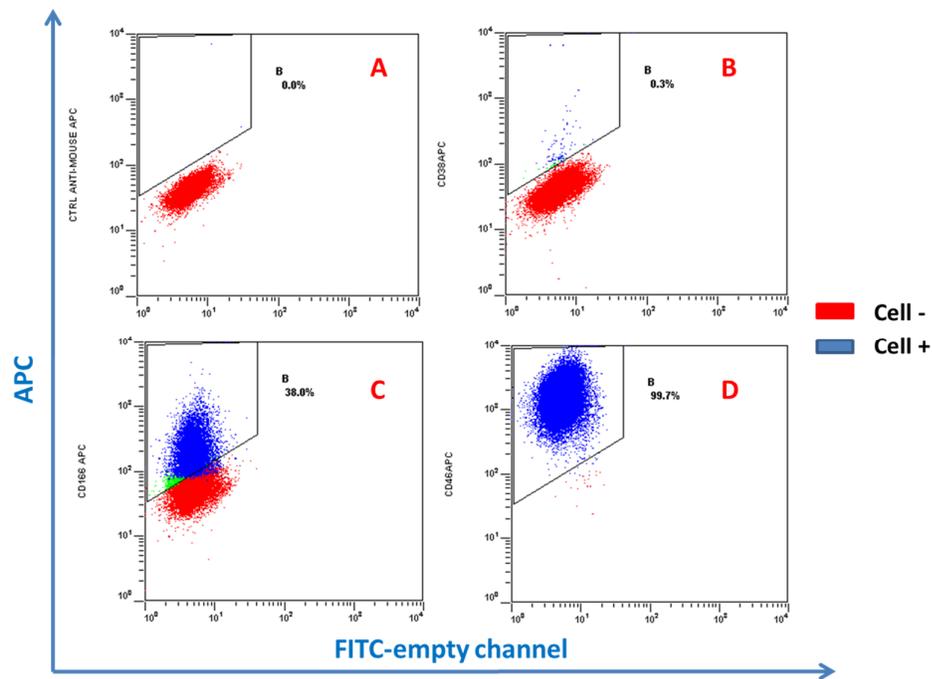
### *4.1 High-throughput screening of the prostate cancer cell line V-CaP by flow cytometry*

Prostate cancer (CaP) is the most frequently diagnosed cancer in men. Progression of CaP from primary to metastatic disease is associated with several molecular and genetic changes that can affect the expression of specific tumor-associated antigens (TAAs) or receptors on the cells surface.

In order to investigate the complete cell surface pattern of a prostate cancer cell line I used an innovative high-throughput screening flow cytometry approach. The aim of this experiment was to know exactly which antigens were expressed or not and identify new antigens subsets useful for a better characterization of prostate cancer cells. The cell line chose to be screened was the V-CaP cell line which is an example of prostate metastatic cell line.

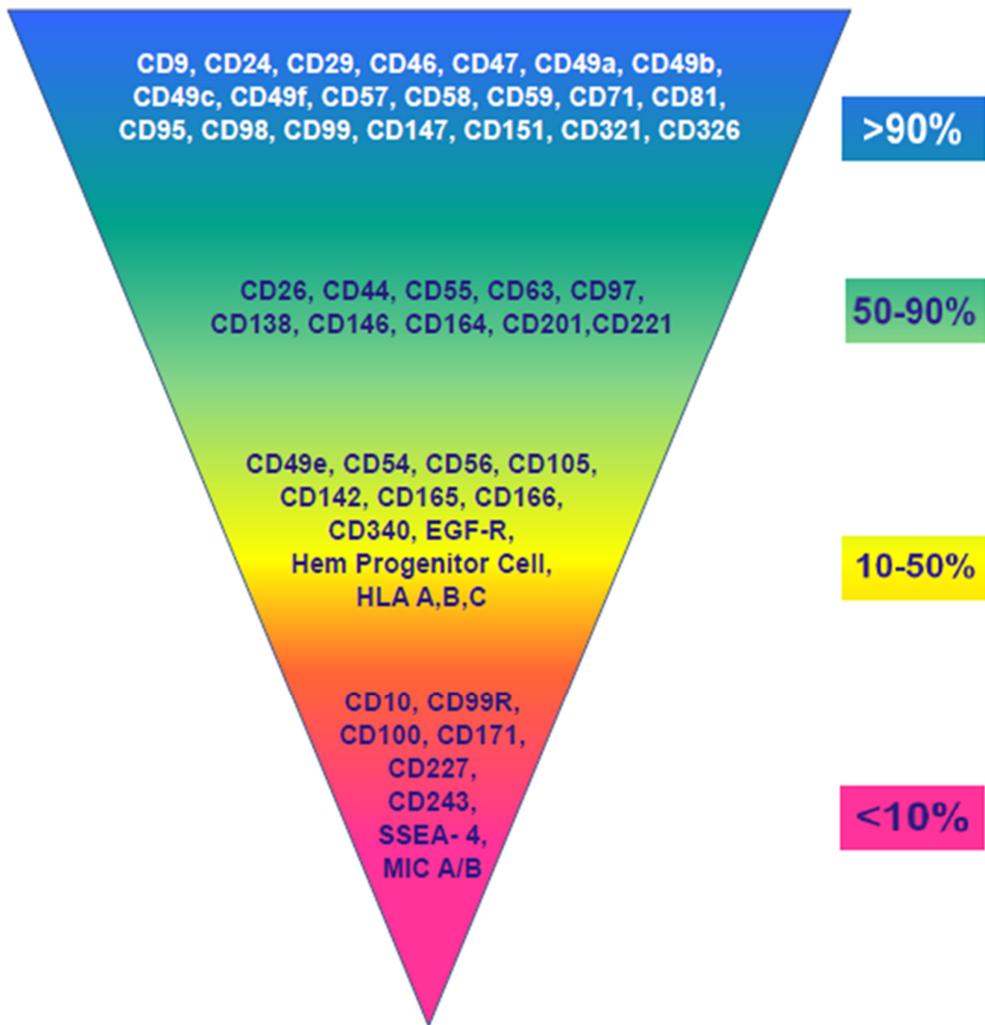
The V-CaP cell line was tested with a panel of 242 different purified monoclonal antibodies of the BD Lyoplate screening panel. As negative control unstained cell samples were used. In this way it has been possible to determine the level of background due to cellular autofluorescence.

The unstained cells were first gated on physical parameters (FSC and SSC) to exclude the majority of debris and dead cells. Afterwards, an APC vs FITC-empty channel dot plot was used to gate the positive cells. The last gate created was used to analyze all 242 .fcs files (**Figure 20**).



**Figure 20:** Gating strategy *BD Lyoplate .fcs* files: A) Gate to detect positive cells on APC vs FITC channel dot plot. B), C), D) Examples of gating strategy applied to stained cells samples.

After the complete analysis of all the data, V-CaP cell line shows a variable positivity for 51 of totally cluster differentiation molecules (CDs) tested. In particular, V-CaP cells are positive (> 90%) for CD9, CD24, CD29, CD44, CD46, CD47, CD49a, CD49b, CD49c, CD49f, CD57, CD58, CD59, CD71, CD81, CD95, CD98, CD99, CD147, CD151, CD321 and CD326. The CD26, CD44, CD55, CD63, CD97, CD138, CD146, CD164, CD201, CD221 positivity was in the 50-90% range and reactivity for CD49e, CD54, CD56, CD105, CD142, CD165, CD166, CD340, EGF-R (Epidermal Growth Factor Receptor), Hem Progenitor Cell and HLA A,B,C was comprised between 10-50%. Finally, it has found a small proportion of cells (< 10%) reactive for CD10, CD99R, CD100, CD171, CD227, CD243, SSEA-4 (**Figure 21**).



**Figure 21:** Graphical representation of cluster differentiation markers on V-CaP cell line after the BD Lyoplate screening.

#### 4.2 Membrane complement regulatory proteins expression on prostate cancer cell lines

Thanks to the results obtained by the BD Lyoplate screening, I focused my attention on the study of mCRPs (CD46, CD55 and CD59). I wanted to test, with use of flow cytometry, the mCRPs expression pattern of other prostate cancer cell lines: 22Rv1, PC-3, Du-145 and LNCaP.

For this experiment, I designed a specific multicolor antibody panel (**Table 5**).

	<b>Tube 1</b>	<b>Tube 2</b>	<b>Tube3</b>
<b>FITC</b>	<b>CD46</b>	<b>CD35</b>	<b>CD46</b>
<b>PE</b>	<b>CD55</b>	<b>CD59</b>	<b>CD49f</b>
<b>PE-Cy7</b>	<b>CD44</b>	<b>CD44</b>	<b>CD44</b>

**Table 5:** Multicolor flow cytometry panel for mCRPs.

$2 \times 10^5$  cells of each cell lines were incubated at 4°C for 30 min with the appropriate amount of MoAbs according the manufacture recommendation. For all staining the appropriate isotype controls and negative controls were set up. The .fcs files were analyzed by using BD FACSDiva Software.

The 22Rv1, PC-3, Du-145 cell lines showed positivity for CD46 > 90% as previously observed with the V-CaP cell line. The LNCaP, instead, has positivity for CD46 of only 50%. The 22Rv1, PC-3, Du-145 CD46+ cells are also positive for CD44, CD49f, CD55 and CD59. The intensity of CD59 is not as high as it is the

CD55. All four cell lines studied are negative for CD35 expression. The LNCaP cell lines has a very low expression of CD44 less than <30% (Figure 22).

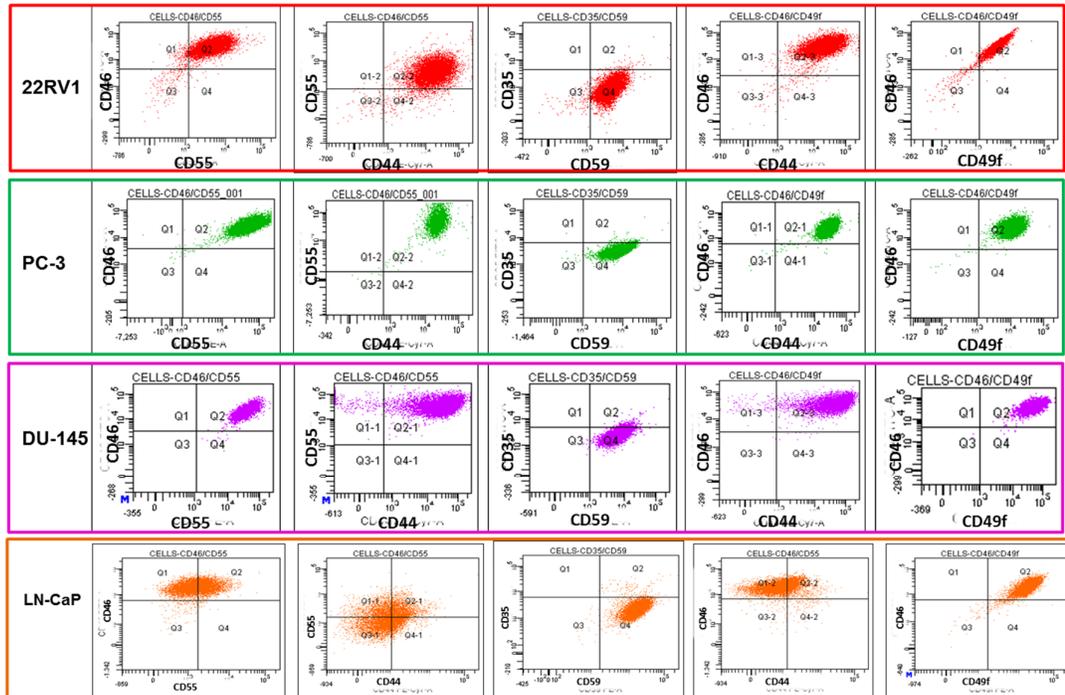


Figure 22: Flow cytometry dot plots mCRPs molecules.

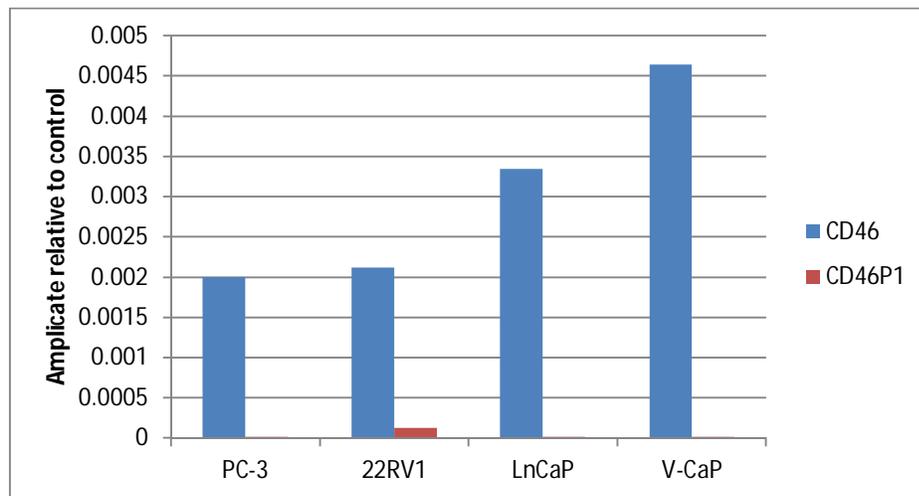
#### 4.3 CD46 pseudogene expression in prostate cancer cell lines

In order to determine the expression level of CD46P1 pseudogene in prostate cancer cells a quantitative RT-PCR (qRT-PCR) was performed.

RNA was isolated from 4 different prostate cancer cell lines: V-CaP, LNCaP, PC-3 and 22Rv1. Primers were specific designed to target exon 5 in CD46P1ps in a high

mismatch region when aligned with CD46 sequence. Primer specificity for CD46P1ps was tested by using two plasmid carrying the exon 5 region specific for CD46 and the pseudogene. For all qRT-PCRs tubulin was used as internal control.

As expected CD46 transcript is highly expressed in all cell lines analyzed. These results are in accordance with the results obtained by flow cytometry. Regarding the pseudogene levels, they are not detectable in PC-3, LNCaP and V-CaP. Basal expression of CD46P1 pseudogene can be detected in 22Rv1 cell line (Figure 23).



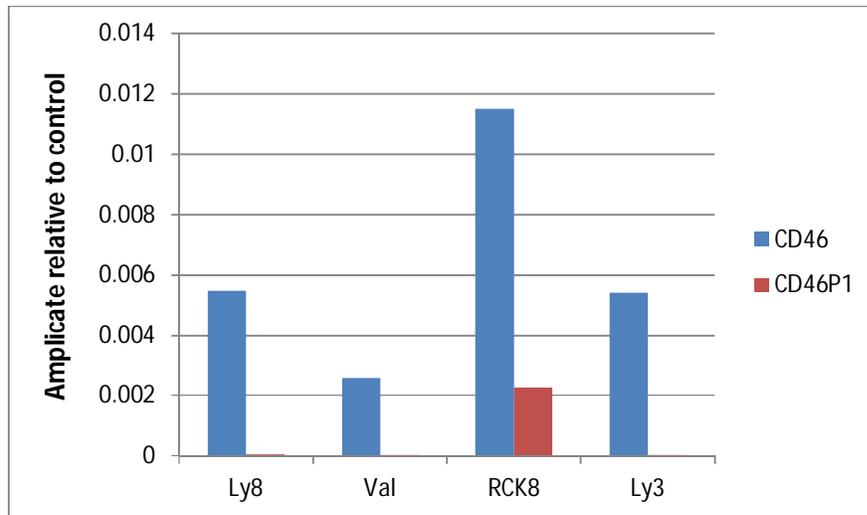
**Figure 23:** *qRT-PCR CD46 and CD46P1 in prostate cancer cell lines*

#### *4.4 CD46 psuedogene expression in lymphoma cell lines*

In order to investigate if CD46 pseudogene could be expressed in other tumor cell lines we tested the lymphoma cell lines: Ly8, Val, RCK-8, Ly3.

As for the prostate cancer cell lines qRT-PCR was performed using Syber Green detection method and tubulin as internal loading control.

In all four lymphoma cell lines CD46 gene is expressed. The levels of expression are similar between Ly8, Val and Ly3 with an increase in RCK-8 cell line. RCK-8 cell line is also the only one expressing the CD46ps. In the other three cell lines CD46ps is either very low or undetectable (**Figure 24**).



**Figure 24:** *qRT-PCR CD46 and CD46P1 in lymphoma cell lines.*

#### 4.5 B-RAFps overexpression in mouse and human cell lines

Due to the high homology between B-RAF gene and B-RAF pseudogene (B-RAFps) in mice and human we hypothesized that overexpression of the non-coding transcript B-RAFps (in mouse and human) is able to antagonize multiple shared miRNAs, thus relieving translational inhibition of the mRNAs of B-RAF coding gene.

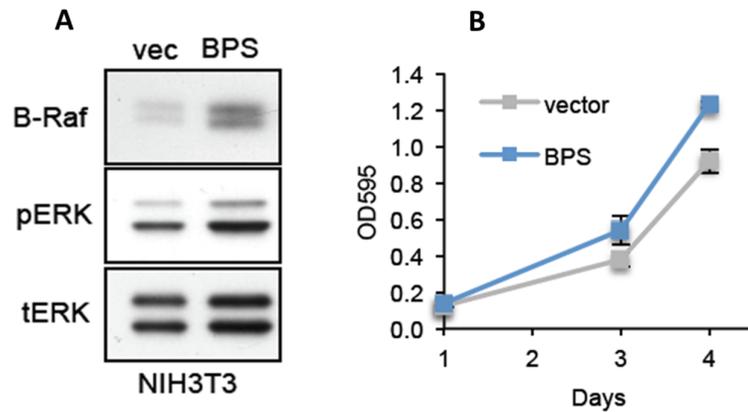
In order to study the sponging effect of B-Rafps versus B-Raf gene, we cloned the murine full length B-Rafps into pCCL.sin.PPT.hPGK.GFP.Wpre expression vector.



**Figure 25:** Mouse *B-Rafps* construct expression vector.

The mouse NIH3T3 cell line was stably transfected with the B-Rafps-pCCL.sin.PPT.hPGK.GFP.Wpre vector and a control NIH3T3 cell line was established with a stable transfection of an empty-pCCL.sin.PPT.hPGK.GFP.Wpre vector. Stable clones of both cell lines (B-Rafps and empty) were expanded and tested by western blotting and proliferation assay.

NIH3T3 cells showed up-regulated B-Raf and pERK expression compared with empty vector-transfected NIH3T3 cells (**Figure 26-A**) and also an increase in proliferation rate (**Figure 26-B**).



**Figure 26:** A) Western Blot B) proliferation assay of NIH3T3 fibroblasts overexpressing B-Rafps.

The same approaches have been used to study of B-RAFps sponging effect in human derived cell lines: PC-9 (lung). The human full length B-RAFps has been cloned into a pLenti-CMV-GFP-Puro expression plasmid (**Figure 27**).

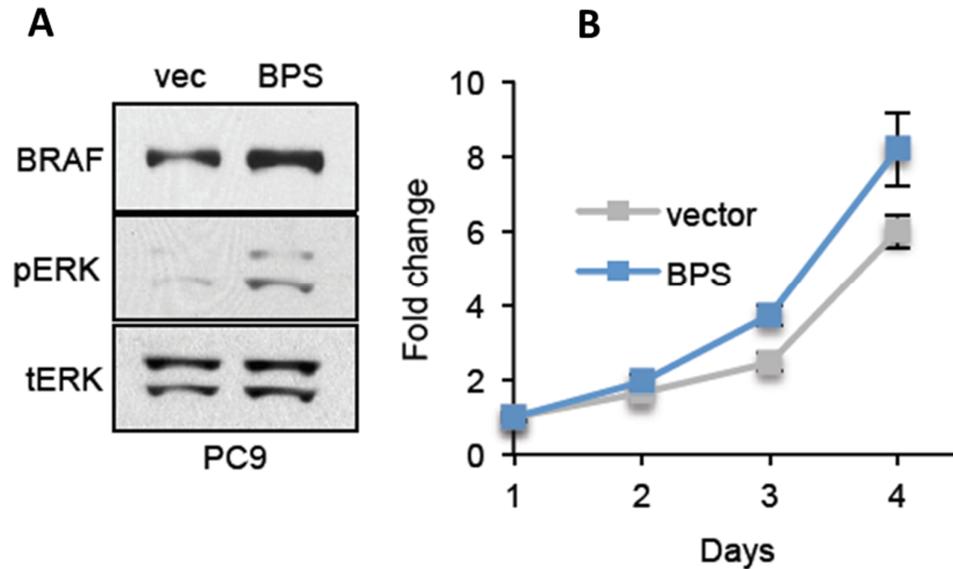


**Figure 27:** Human B-RAFps cloning construct.

Control PC-9 cells were also setup with empty- pLenti-CMV-GFP-Puro vector. All the clones have been selected with puromycin (1 $\mu$ g/ml) and the tested for western blot and proliferation assay.

As observed in the murine cell line system in presence of B-RAFps overexpression there is an increase in B-RAF protein levels, which is followed by the increase of pERK levels (**Figure 28-A**). In addition to these also the PC-9-B-RAFps-

vector cells show an increase in proliferation rate when compared to PC-9-empty-vector (**Figure 28-B**).



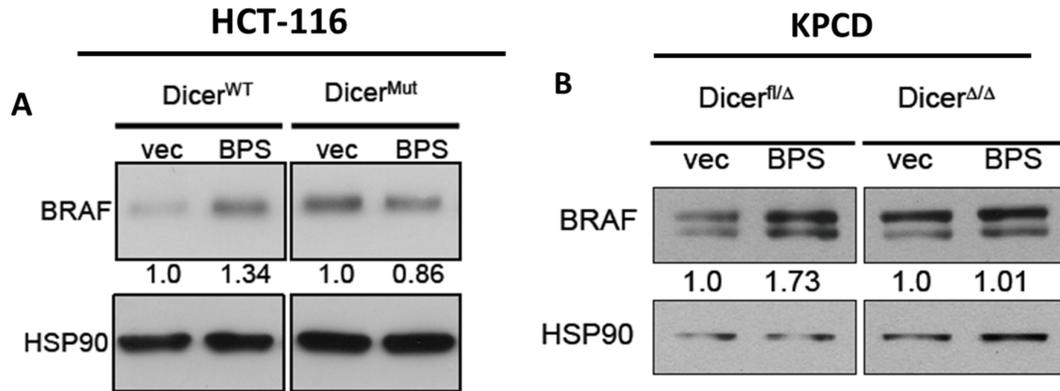
**Figure 28:** A) *Western Blot* B) *proliferation assay of PC-9 cell line overexpressing B-RAFps.*

#### 4.6 *B-RAFps sponging effect is Dicer/miRNA dependent*

To investigate the role for Dicer-controlled miRNAs mediated cross-talk between B-RAF transcript and its pseudogene we utilized cell lines engineered to lack functional Dicer1, a ribonuclease that mediates critical step during miRNA biogenesis. The lack of Dicer causes a drastic reduction of mature miRNA levels.

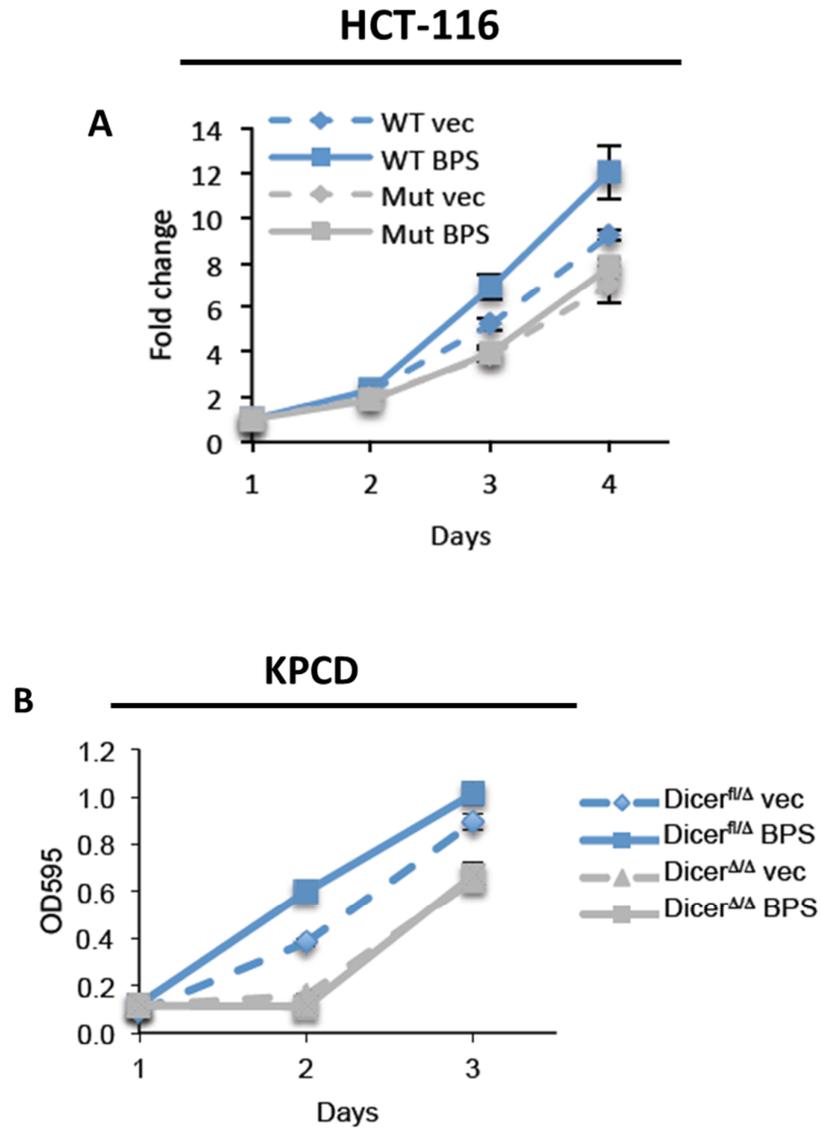
The cell lines used are murine sarcoma cell line KPCD  $Dicer^{fl/\Delta}$  and KPCD  $Dicer^{\Delta/\Delta}$  (**101**) and human HCT-116 colon cancer cell line  $Dicer^{wt}$  and HCT-116  $Dicer^{mut}$  (**102**). These cell lines have been transfected specifically with murine or human B-RAF pseudogene as in *paragraph 4.4*.

Ectopic overexpression B-RAFps in HCT-116  $Dicer^{wt}$  and KPCD  $Dicer^{fl/\Delta}$  induced an increase in B-RAF expression as shown by western blots. When B-RAFps is overexpressed in HCT-116  $Dicer^{mut}$  or KPCD  $Dicer^{\Delta/\Delta}$  I did not observe an increase in B-RAF levels (**Figure 29 A-B**). These data indicate that mature miRNAs are critical for the crosstalk.



**Figure 29:** Western blot A) HCT-116 B) KPCD overexpressing B-RAFps.

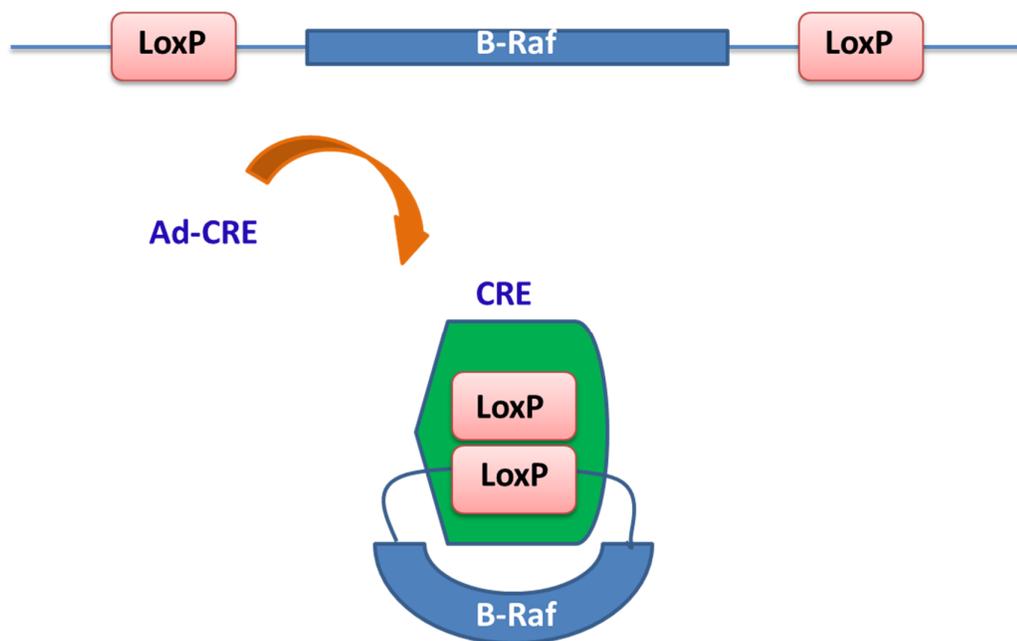
The increased level of B-RAF protein in HCT-116  $Dicer^{wt}$  and KPCD  $Dicer^{fl/\Delta}$  was confirmed by proliferation assays (**Figure 30**). The proliferation rate of the cell lines containing a functional Dicer in homozygosis or heterozygosis proliferate more than the ones without.



**Figure 30:** Proliferation assay A) HCT-116 B) KPCD overexpressing B-RAFps.

#### 4.7 B-RAFps sponging effect is B-RAF dependent

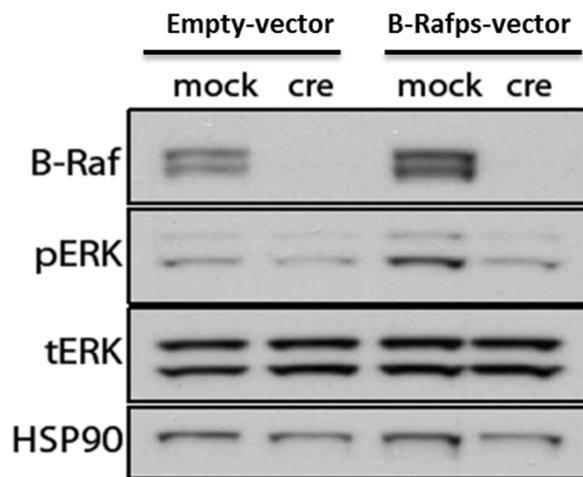
Mouse B-Raf<sup>fl/fl</sup> fibroblasts were used in order to determine if the B-Rafps sponging effect is B-Raf dependent. These B-Raf<sup>fl/fl</sup> fibroblasts were derived from a conditional knockout mouse for B-Raf gene. When the cells are infected with an adenovirus carrying the Cre-recombinase the B-Raf gene is genetically deleted because it is flanked by two loxP sites as shown in **Figure 31**.



**Figure 31:** *Conditional B-Raf knock-out fibroblasts.*

The B-Raf<sup>fl/fl</sup> fibroblasts were transfected with an expression vector to overexpress B-Rafps and an empty-vector as negative control. The same cells were then infected with Ad-Cre, to knock out B-Raf gene and a negative control infection virus (Ad-Mock).

The cells infected with Ad-Mock virus do not lose B-Raf gene. When the cells are infected with Ad-Cre virus the B-Raf gene is genetically deleted and no B-Raf protein is detectable with western blot. In addition to these, B-Raf<sup>fl/fl</sup> fibroblasts transfected with B-Rafps vector show an increase in B-Raf protein levels when compared to the fibroblasts transfected with the empty-vector. Increases of pERK levels follow the B-Raf protein increase (**Figure 32**). These dynamics are not observed in the same cells infected with Ad-Cre virus where B-Raf gene is deleted. In absence of B-Raf gene no mRNA transcript is produced. In this way, no cross-talk there is between the B-Raf gene transcript and its pseudogene.



**Figure 32:** Western blot B-Raf<sup>fl/fl</sup> fibroblasts overexpressing B-Rafps in presence or absence of Adeno-Cre infection.

## 5. DISCUSSION

### *5.1 CD46 and its pseudogene CD46P1 in prostate cancer*

Prostate cancer is an important disease in the modern era, based on its prevalence, the controversies surrounding diagnosis and treatment selection. Progression of prostate cancer from primary to metastatic disease is associated with several molecular and genetic changes that can affect the expression of specific tumor-associated antigens (TAAs) or receptors on the cells surface. Understanding how the cancer cells differ from the normal cells, might allow them to be targeted selectively and eliminated, thus improving therapeutic outcome.

The first aim of this thesis was to determinate the complete immunophenotype of a metastatic prostate cancer cell line (V-CaP) in order to identify specific sets of antigens that could better characterize the prostate cancer cells and they could be specifically associated with a prostate cancer stage. In order to this, the Vertebral-Cancer of the Prostate (V-CaP) cell line was screened for 242 cell surface molecules by using a high-throughput flow cytometry approach: the BD Lyoplate screening panel. This approach has been considered the most suitable technique for this screening because it provides the possibility to acquire data from multiple assays in a single experiment reducing the intra-lab variability and it is the most completed kit for human cell surface antigens screening available on market.

After the analysis of all 242 .fcs files, it emerged that the V-CaP cell line has a variable positivity for 51 of totally cluster differentiation molecules (CDs) tested (**Figure 21**). As expected the V-CaP cell line is negative for hematological cell-lineage specific antigens (i.e. CD3, CD4, CD8, CD34...) while it is positive for

molecules involved in cell adhesion such as integrins (i.e. CD49a,b,c,f).

Interestingly, the V-CaP cell line shows, also, a high positivity for membrane complement regulatory proteins (mCRPs) expression: CD46, CD55, and CD59.

The mCRPs are known for their role in the immune evasion strategy of tumors cells by inhibiting complement cascade activation. An overexpression of these molecules on cancer cells significantly limit the anti-tumor mAb therapeutic efficacy and increase the tumor cell metastatic potential.

Because of that, I wanted to test other prostate cancer cell lines (22Rv1, PC-3, Du-145 and LNCaP) for the mCRPs expression by using a specific multicolor antibody panel shown in **Table 5**. In addition to the mCRPs molecules, I investigated the expression of other membrane cell surface proteins such as CD35, CD44 and CD49f, due to their role in cell homeostasis. In particular, CD35 was interesting to study because it is an N-glycosylated member of the RCA (regulators of complement activation) family of proteins. As CD46 and CD46P1 genes, its genomic locus is in the RCA gene cluster on chromosome 1. CD35's main function is to protect cells from complement-mediated lysis by serving as a cofactor for Factor I and inhibiting the C3 and C5 convertases. The CD44 antigen is a cell-surface glycoprotein involved in cell-cell interactions, cell adhesion and migration. CD44 together with CD49f (commonly expressed on human prostate basal cells) are reported to be used as cell surface markers to isolate breast and prostate cancer stem cells (**99, 100**).

Three of the cell line tested (22Rv1, PC-3 and Du-145) showed positivity for CD46 > 90% as previously observed in the V-CaP cell line. The LNCaP cell line, instead, has positivity for CD46 of only 50%. The 22Rv1, PC-3 and Du-145 CD46+ cells are also positive for CD44, CD49f, CD55 and CD59. The intensity of CD59 is

not as high as it is the CD55 one. All four cell lines (22Rv1, PC-3, Du-145 and LNCaP) are negative for CD35 expression (**Figure 21**). The CD35 negativity is not surprising because it is described in literature that CD35 is expressed on monocytes, dendritic cells and lymphocytes.

The results obtained from the described experiments let arise other questions such as which mechanisms could lead to an upregulation of mCRPs expression in prostate cancer cell lines, in particular the upregulation of CD46 protein.

According to me, the CD46 molecule is so interesting to be studied because:

- 1- It is the highest representative molecule among the mCRPs;
- 2- It is the only mCRPs molecule which has a pseudogene in human genome which is located in the same genomic locus of CD46 gene (RCA gene cluster);
- 3- It is a multifunctional protein;
- 4- It is used as target for prostate specific immunotherapy.

At the time of writing this thesis, no studies have been conducted to investigate the role of CD46P1 as modulator molecule for CD46 gene expression through the ceRNA hypothesis. As previously described in *paragraph 1.1.4* a pseudogene may play a role as competing endogenous RNA if some conditions are present. First of all, it has to be transcribed.

For this reason, I decided to investigate if CD46P1 is transcribed in prostate cancer cell line by using qRT-PCR. As shown and discussed in *paragraph 4.3* CD46P1 pseudogene is expressed at low levels only in 22Rv1 cell line while in the other cell lines studied (V-CaP, PC-3 and LNCaP) it is undetectable. Because of that, I decided to extend the CD46P1 transcript research to lymphoma cancer cell lines, which are an example of non-solid tumor.

In three lymphoma cell lines (Ly3, Ly8, Val) studied no CD46P1 transcript expression is undetectable. Only in one cell line, the RCK-8 it is possible to detect some expression of CD46 pseudogene (**Figure 24**).

Why only the prostate cell line 22Rv1 and the lymphoma cell line Ly8 express the pseudogene CD46P1 has to be investigated. It could be possible that these cell lines, being derived from primary tumors, could have been more exposed to the complement activation. A direct consequence of that could be a transcriptional hyper-activation of RCA gene cluster and an increase in CD46P1 levels.

## 5.2 B-RAF as competing endogenous RNA

This PhD work focused its attention on the study of B-RAF pseudogene for several reasons. First of all, some scientific research report that B-RAF pseudogene is overexpressed in human thyroid cancer (**104**) and other cancer types (**65**), suggesting that its deregulation may promote cancer development. Secondly, the MAPK pathway is commonly hyperactivated in cancer and a potential ceRNA activity of B-RAFps toward B-RAF may impact this signaling pathway. Finally, it is well known that some human hematopoietic malignancies are associated with overdosage of the X chromosome, which is where the B-RAF pseudogene is located.

In this thesis it has been demonstrated that B-RAF pseudogene elicits its oncogenic function by decoying miRNAs from its parental gene B-RAF through a series of *in vitro* experiments.

Murine and human B-RAF corresponding pseudogenes have been cloned in two expression vectors and transfected in two cell lines: PC-9 for human B-RAFps and NIH3T3 cell line for murine B-Rafps. According to the ceRNA hypothesis, if a RNA transcript with ceRNA potential is overexpressed in a cell, it can easily sponge and subsequently compete for the common miRNA pool shared with another transcript.

When the human and murine B-RAF pseudogenes were overexpressed in PC-9 and NIH3T3 respectively, we observed, by western blot, an increase in B-RAF and pERK protein levels (**Figure 26-A and 28-A**). In addition to these, when the same cells were used for proliferation assays I observe an increase in cells proliferation rate (**Figure 26-B and 28-B**) correlated to an activation of MAPK pathway suggested also by the western blot results discussed above.

Because of that, I investigated how many miRNA response elements (MREs) the B-RAF pseudogene and its coding-gene share both in human and in mice. As shown in **Figure 10**, BRAFps and BRAF share a large number of common MREs: 40 different MREs in human while in mice around 52. These numbers suggest that the two transcripts can cross-talk using common miRNAs molecules.

The next question we wanted to answer was: if no more mature miRNAs are produced in the cell what happens to the cross-talk between the B-RAFps and its coding-gene transcript? According to the ceRNA hypothesis without miRNAs no cross-talk could be possible.

In order to investigate the role for Dicer-controlled miRNAs mediated cross-talk between B-RAF transcript and its pseudogene I utilized cell lines engineered to lack functional Dicer1, a ribonuclease that mediates critical step during miRNA biogenesis. The lack of Dicer causes a drastic reduction of mature miRNA levels. The cell lines used are murine sarcoma cell line KPCD Dicer<sup>fl/Δ</sup> and KPCD Dicer<sup>Δ/Δ</sup> and human HCT-116 colon cancer cell line Dicer<sup>wt</sup> and HCT-116 Dicer<sup>mut</sup>.

The results obtained clearly show that in absence of Dicer1, no mature miRNAs are generated and B-RAF gene transcript and its pseudogene cannot cross-talk. In fact, when we compare the B-RAF protein level in HCT-116 Dicer<sup>wt</sup> and HCT-116 Dicer<sup>mut</sup> cells overexpressing the pseudogene, we observed that B-RAF protein level increase only in the cell line with a functional Dicer (HCT-116 Dicer<sup>wt</sup>) (**Figure 29-A**).

The next point I wanted to investigate was to understand if the sponging effect of B-RAF pseudogene is B-RAF dependent. In order to do that, I used a mouse B-Raf<sup>fl/fl</sup> fibroblast cell line where it was possible to delete both allele copies of B-Raf gene after Ad-Cre virus infection (**Figure 31**). As shown in **Figure 32**

When the B-Rafps is overexpressed in the B-Raf<sup>f1/f1</sup> fibroblast cell line infected with the control virus (Ad-Mock), unable to induce the genetic B-Raf gene deletion, the B-Raf protein levels increase followed by pERK increase too. When the same cell line is infected with Ad-Cre virus, the two copies of B-Raf gene are deleted and no B-Raf is detectable. What is important is that without B-Raf no increase in pERK protein is detectable too in response to B-Rafps overexpression.

This experiment supports the notion that the effect of BRAFps on MAPK activation is indeed mediated by B-RAF.

## 6. FINAL OBSERVATIONS

This thesis has had two protagonists: the B-RAF pseudogene and the CD46P1 pseudogene. The *in vitro* experiments conducted on B-RAF pseudogene, clearly show that it is a pseudogene that functions as ceRNA and promotes oncogenesis when deregulated. The *in vitro* results also give confirmation to what observed *in vivo* by Karreth et al. (103).

At this point it is not possible to define CD46P1 as a pseudogene with ceRNA ability. The data obtained are preliminary and many other *in vitro* experiments need to be done. Nevertheless, CD46 and its CD46P1 revealed to be interesting molecules to be studied.

## 7. BIBLIOGRAPHY

1. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. *Nature* 2012;489:101–8
2. Gesteland RF, Cech TR, Atkins JF (eds) *The RNA world*, 3rd edn. Cold Spring Harbor Laboratory Press, New York, pp 257–285.
3. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006 Apr 15;15 Spec No 1:R17-29.
4. Clark MB, Mattick JS. Long noncoding RNAs in cell biology. *Semin Cell Dev Biol* 2011.
5. Ponting CP, Oliver PL, Reik W. Evolution and functions of long noncoding RNAs. *Cell.* 2009 Feb 20; 136(4):629-41.
6. Willingham, A. T., Orth, A. P., Batalov, S., Peters, E. C., Wen, B. G., Azabanc, P., Hogenesch, J. B. & Schultz, P. G. A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. *Science* 309:1570-3, 2005.
7. Brannan, C. I., Dees, E. C., Ingram, R. S. & Tilghman, S. M. The product of the H19 gene may function as an RNA. *Mol. Cell. Biol.* 10, 28–36 (1990).
8. Brown CJ, Ballabio A, Rupert JL, Lafreniere RG, Grompe M, Tonlorenzi R, Willard HF. A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. *Nature.* 1991 Jan 3; 349(6304):38–44.

9. Tano K, Akimitsu N. Long non-coding RNAs in cancer progression. *Front Genet.* 2012 Oct 24;3:219.
10. Wapinski O, Chang HY. Long noncoding RNAs and human disease. *Trends Cell Biol* 2011; 21:354-61.
11. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013 Mar; 193(3):651-69.
12. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al., Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* 2011.
13. Rearick, D., A. Prakash, A. McSweeney, S. S. Shepard, L. Fedorova et al., 2011 Critical association of ncRNA with introns. *Nucleic Acids Res.* 39: 2357–2366.
14. Huarte M, Rinn JL: Large non-coding RNAs: missing links in cancer? *Hum Mol Genet* 2010, 19(R2):R152-161.
15. Katayama S, Tomaru Y, Kasukawa T, Waki K, Nakanishi M, Nakamura M, et al.; RIKEN Genome Exploration Research Group; Genome Science Group (Genome Network Project Core Group); FANTOM Consortium. Antisense transcription in the mammalian transcriptome. *Science* 2005.
16. Proudfoot N. Pseudogenes. *Nature* 1980; 286:840–1.
17. Zhang ZL, Harrison PM, Liu Y, Gerstein M. Millions of years of evolution preserved: A comprehensive catalog of the processed pseudogenes in the human genome. *Genome Research.* 2003; 13:2541–2558.

18. Y. J. Han, S. F. Ma, G. Yourek, Y.-D. Park, and J. G.N. Garcia, "A transcribed pseudogene of MYLK promotes cell proliferation," *The FASEB Journal*, vol. 25, no. 7, pp. 2305–2312, 2011.
19. Zhang ZD, Frankish A, Hunt T, et al. Identification and analysis of unitary pseudogenes: historic and contemporary gene losses in humans and other primates. *GenomeBiol* 2010; 11:R26.
20. Nishikimi M, Fukuyama R, Minoshima S, et al. Cloning and chromosomal mapping of the human nonfunctional gene for L-gulono-gamma-lactone oxidase, the enzyme for l-ascorbic-acid biosynthesis missing in man. *J Biol Chem* 1994; 269:13685–8.
21. H. Kaessmann et al. RNA-based gene duplication: mechanistic and evolutionary insights *Nat. Rev. Genet.*, 10 (2009), pp. 19–31.
22. Vanin EF. Processed pseudogenes - Characteristics and evolution. *Annu Rev Genet* 1985; 19:253–72.
23. Esnault C, Maestre J, Heidmann T. Human LINE retrotransposons generate processed pseudogenes. *Nat Genet* 2000; 24:363–7.
24. Eric Christian Rouchka, and I. Elizabeth Cha. Current Trends in Pseudogene Detection and Characterization. *Current Bioinformatics*, 2009, 4, 112-119.
25. Zhang Z, Gerstein MB. 2004. Large-scale analysis of pseudogenes in the human genome. *Curr Opin Genet Dev* 14:328–35.
26. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, et al. MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* 2004; 23:4051–4060.

27. Yan Zeng and Bryan R. Cullen. Structural requirements for pre-microRNA binding and nuclear export by Exportin 5. *Nucleic Acids Res.* 2004.
28. Paraskevopoulou MD , Georgakilas G , Kostoulas N , Reczko M ,Maragkakis M , Dalamagas TM , et al. DIANA-LncBase: experimentally verified and computationally predicted microRNA targets on long non-coding RNAs . *Nucleic Acids Res* 2013; 41 : D239 – 45.
29. Thomas M, Lieberman J, Lal A. Desperately seeking microRNA targets. *Nat Struct Mol Biol.* 2010 Oct; 17(10):1169-74.
30. Pratt, A.J. and MacRae, I.J. (2009) The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.*, 284, 17897–17901.
31. Ambros, V. microRNAs: tiny regulators with great potential. *Cell* 107, 823-826 (2001)
32. Alvarez-Garcia I. and Miska, EA. MicroRNA Functions in Animal Development and Human Disease. *Development.* 2005; 132 (21): .4653-4662.
33. Calin, G.A. and Croce, C.M. MicroRNA Signatures in Human Cancers. *Nature Reviews Cancer.* 2006; 6(11):857-66
34. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009
35. Poliseno L., Salmena L., Zhang J., Carver B., Haveman W.J., Pandolfi P.P.; Salmena; Zhang; Carver; Haveman; Pandolfi (2010). "A coding-independent function of gene and pseudogene mRNAs regulates tumour biology". *Nature* 465 (7301): 1033–1038.

36. L. Salmena, L. Poliseno, Y. Tay, L. Kats, and P. P. Pandolfi, “A ceRNA hypothesis: the rosetta stone of a hidden RNA language?” *Cell*, vol. 146, no. 3, pp. 353–358, 2011.
37. Ala U., Karreth F. A., Bosia C., Pagnani A., Taulli R., Leopold V., et al. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc. Natl. Acad. Sci. U.S.A* (2013).
38. Karreth FA, Pandolfi PP. ceRNA Cross-Talk in Cancer: When ce-bling Rivalries Go Awry. *Cancer Discov.* 2013 Oct; 3(10):1113-21.
39. Xi JJ. MicroRNAs in Cancer. *Cancer Treat Res.* 2013; 158:119-37.
40. Tay Y , Kats L , Salmena L , Weiss D , Tan SM , Ala U , et al. Coding independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs . *Cell* 2011; 147 : 344 – 57
41. de Giorgio A, Krell J, Harding V, Stebbing J, Castellano L. Emerging roles of competing endogenous RNAs in cancer: insights from the regulation of PTEN. *Mol Cell Biol.* 2013 Oct; 33(20):3976-82.
42. Eychène A, Barnier JV, Apiou F, Dutrillaux B, Calothy G. Chromosomal assignment of two human B-raf(Rmil) proto-oncogene loci: B-raf-1 encoding the p94Braf/Rmil and B-raf-2, a processed pseudogene. *Oncogene.* 1992 Aug; 7(8):1657-60.
43. Wellbrock, M. Karasarides, R. Marais, The RAF proteins take centre stage, *Nat. Rev. Mol. Cell Biol.* 5 (2004) 875–885.

44. Roskoski R J. RAF protein-serine/threonine kinases: structure and regulation. *Biochem Biophys Res Commun.* 2010 Aug 27; 399(3):313-7.
45. Avruch J, Khokhlatchev A, Kyriakis JM, Luo Z, Tzivion G, Vavvas D, Zhang XF. Ras Activation of the Raf Kinase: Tyrosine Kinase Recruitment of the MAP Kinase Cascade. *Recent Prog Horm Res.* 2001; 56:127-55.
46. Zebisch A, Troppmair J. Back to the roots: the remarkable RAF oncogene story. *Cell Mol Life Sci.* 2006 Jun; 63(11):1314-30.
47. Alessi DR, Saito Y, Campbell DG, Cohen P, Sithanandam G, Rapp U, Ashworth A, Marshall CJ, Cowley S. Identification of the sites in MAP kinase kinase-1 phosphorylated by p74raf-1. *EMBO J.* 1994 Apr 1;13(7):1610-9.
48. Yoon S, Seger R. (2006). The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions. *Growth Factors* 24: 21–44.
49. Junttila MR, Li SP, Westermarck J. Phosphatase-mediated crosstalk between MAPK signaling pathways in the regulation of cell survival. *FASEB J.* 2008 Apr;22(4):954-65. Epub 2007 Nov 26.
50. McCubrey JA, Steelman LS, Chappell WH, et al. Roles of the Raf/MEK/ERK pathway in cell growth, malignant transformation and drug resistance. *Biochim Biophys Acta.* 2007;1773:1263–84.
51. Domingo E, Schwartz S Jr . BRAF (v-raf murine sarcoma viral oncogene homolog B1). *Atlas Genet Cytogenet Oncol Haematol.* September 2004Zhang Y, Dong C. Regulatory mechanisms of mitogen-activated kinase signaling. *Cell. Mol. Life Sci.* (2007). 64(21), 2771–89.

52. Barnier, J. V., Papin, C., Eychene, A., Lecoq, O. and Calothy, G. The mouse B-raf gene encodes multiple protein isoforms with tissue-specific expression. *J Biol Chem.* 1995 Oct 6;270(40):23381-9.
53. Davies H., Bignell G. R., Cox C., Stephens P., Edkins S., Clegg S. et al. Mutations of the BRAF gene in human cancer. *Nature* (2002) 417: 949–95443
54. Flaherty KT, McArthur G. BRAF, a target in melanoma: implications for solid tumor drug development. *Cancer.* 2010;116:4902–4913.
55. Ricarte-Filho JC, Ryder M, Chitale DA, Rivera M, Heguy A, Ladanyi M, Janakiraman M, Solit D, Knauf JA, Tuttle RM, Ghossein RA, Fagin. Mutational profile of advanced primary and metastatic radioactive iodine-refractory thyroid cancers reveals distinct pathogenetic roles for BRAF, PIK3CA, and AKT1. *JACancer Res.* 2009 Jun 1.
56. Forbes SA, Tang G, Bindal N, Bamford S, Dawson E, Cole C, Kok CY, Jia M, Ewing R, Menzies A, Teague JW, Stratton MR, Futreal PA. COSMIC (the Catalogue of Somatic Mutations in Cancer): a resource to investigate acquired mutations in human cancer. *Nucleic Acids Res.* 2010 Jan;38
57. Li WQ, Kawakami K, Ruzkiewicz A, Bennett G, Moore J, Iacopetta B. BRAF mutations are associated with distinctive clinical, pathological and molecular features of colorectal cancer independently of microsatellite instability status. *Mol Cancer.* 2006 Jan 10;5:2.
58. Wong KK, Tsang YT, Deavers MT, Mok SC, Zu Z, Sun C, Malpica A, Wolf JK, Lu KH, Gershenson DM. BRAF mutation is rare in advanced-stage low-grade ovarian serous carcinomas. *Am J Pathol.* 2010 Oct;177(4):1611-7.

59. Garnett MJ, Marais R. Guilty as charged: B-RAF is a human oncogene. *Cancer Cell*. 2004 Oct;6(4):313-9.
60. YH, Liu Y, Eu KW, Ang PW, Li WQ, Salto-Tellez M, Iacopetta B, Soong R. Detection of BRAF V600E mutation by pyrosequencing. *Pathology* April 2008. 40 (3): 295–8.
61. Cantwell-Dorris ER, O'Leary JJ, Sheils OM. BRAFV600E: implications for carcinogenesis and molecular therapy. *Mol Cancer Ther*. 2011 Mar;10(3):385-94.
62. Tiacci E, Trifonov V, Schiavoni G, Holmes A, Kern W, Martelli MP, Pucciarini A, Bigerna B, Pacini R, Wells VA, Sportoletti P et al. BRAF mutations in hairy-cell leukemia. *N Engl J Med*. 2011 Jun 16;364(24):2305-15. doi: 10.1056/NEJMoa1014209. Epub 2011 Jun 11.
63. J.W. Lee, Y.H. Soung, S.Y. Kim, W.S. Park, S.W. Nam, W.S. Min, S.H. Kim, J.Y. Lee, N.J. Yoo, S.H. Lee, Mutational analysis of the ARAF gene in human cancers, *APMIS* 113 (2005) 54–57.
64. Brose MS, Volpe P, Feldman M, Kumar M, Rishi I, Gerrero R, Einhorn E, Herlyn M, Minna J, Nicholson A, Roth JA et al. BRAF and RAS mutations in human lung cancer and melanoma. *Cancer Res*. 2002 Dec 1;62(23):6997-7000.
65. Kalyana-Sundaram S, Kumar-Sinha C, Shankar S, Robinson DR, Wu YM, Cao X, Asangani IA, Kothari V, Prensner JR, Lonigro RJ et al. Expressed pseudogenes in the transcriptional landscape of human cancers. *Cell* (2012) 149, 1622-1634.

66. Vacher J, Bernard H., Genetic localization and transmission of the mouse osteopetrotic grey-lethal mutation. *Mamm Genome*. 1999 Mar;10(3):239-43.
67. Lublin, D.M., M.K. Liszewski, T.W. Post, M.A. Arce, M.M. Le Beau, M.B. Rebentisch, L.S. Lemons, T. Seya, and J.P. Atkinson. Molecular cloning and chromosomal localization of human membrane cofactor protein (MCP). Evidence for inclusion in the multigene family of complement-regulatory proteins. *J. Exp. Med*. 1988. 168:181–194.
68. Krushkal J, Bat O, Gigli I. Evolutionary relationships among proteins encoded by the regulator of complement activation gene cluster. *Mol Biol Evol*. 2000 Nov;17(11):1718-30.
69. Heinen S, Sanchez-Corral P, Jackson MS, Strain L, Goodship JA, Kemp EJ, Skerka C, Jokiranta TS et al. De novo gene conversion in the RCA gene cluster (1q32) causes mutations in complement factor H associated with atypical hemolytic uremic syndrome. *Hum Mutat*. 2006 Mar;27(3):292-3.
70. Liszewski, M.K., T.W. Post, and J.P. Atkinson. Membrane cofactor protein (MCP or CD46): newest member of the regulators of complement activation gene cluster. *Annu. Rev. Immunol*. 1991 9:431–455.
71. Wilton AN, Johnstone RW, McKenzie IFC, et al. Strong associations between RFLP and protein polymorphism for CD46. *Immunogenetics*. 1992;36:79–85.
72. Seya T, Hirano A, Matsumoto M, et al. Human membrane cofactor protein (MCP, CD46): multiple isoforms and functions. *Int J Biochem Cell Biol*. 1999;31:1255–60.

73. Johnstone RW, Russell SM, Loveland BE, Jayasuriya N, McKenzie IF. Tissue-specific expression of CD46 isoforms. *Transplant Proc* 1992;24:2331–2.
74. Liszewski, M.K. et al. Control of the complement system. *Adv. Immunol.* (1996) 61, 201–283
75. Le Friec G, Kemper C. Complement: coming full circle. *Arch Immunol Ther Exp (Warsz)* 2009; 57:393–407
76. Pickering MC, Cook HT. Translational mini-review series on complement factor H: renal diseases associated with complement factor H: novel insights from humans and animals. *Clin Exp Immunol* 2008; 151:210–30.
77. Kim DD, Song WC. Membrane complement regulatory proteins. *Clin Immunol.* 2006 Feb-Mar; 118(2-3):127-36.
78. Cattaneo R. Four viruses, two bacteria, and one receptor: membrane cofactor protein (CD46) as pathogens' magnet. *J Virol* 2004; 78:4385–8.
79. Riley-Vargas, R.C., Gill, D.B., Kemper, C., Liszewski, M.K., Atkinson, J.P., CD46: expanding beyond complement regulation. *Trends Immunol.* 2004. 25, 496–503.
80. Naniche D, Varior-Krishnan G, Cervoni F et al. Human membrane cofactor protein (CD46) acts as a cellular receptor for measles virus. *J Virol* 1993; 67: 6025.
81. Santoro F, Kennedy PE, Locatelli G, Malnati MS, Berger EA, Lusso P. CD46 is a cellular receptor for human herpesvirus 6. *Cell* 1999; 99: 817.

82. Giannakis E, Jokiranta TS, Ormsby RJ et al. Identification of the streptococcal M protein binding site on membrane cofactor protein (CD46). *J Immunol* 2002; 168: 4585.
83. Katayama Y, Hirano A, Wong TC. Human receptor for measles virus (CD46) enhances nitric oxide production and restricts virus replication in mouse macrophages by modulating production of alpha/beta interferon. *J Virol* (2000) 74:1252–1257.
84. Schnorr JJ, Dunster LM, Nanan R et al., Measles virus-induced down-regulation of CD46 is associated with enhanced sensitivity to complement-mediated lysis of infected cells. *Eur J Immunol* (1995) 25:976–984.
85. Mizuno M, Harris CL, Suzuki N, Matsuo S, Morgan BP. Expression of CD46 in developing rat spermatozoa: ultrastructural localization and utility as a marker of the various stages of the seminiferous tubuli. *Biol reprod* 2005, 72: 908.
86. Riley RC, Kemper C, Leung M, Atkinson JP. Characterization of human membrane cofactor protein (MCP; CD46) on spermatozoa. *Mol Reprod Dev* 2002a;62:534–46.
87. Cummerson JA, Flanagan BF, Spiller DG, Johnson PM. The complement regulatory proteins CD55 (decay accelerating factor) and CD59 are expressed on the inner acrosomal membrane of human spermatozoa as well as CD46 (membrane cofactor protein). *Immunology* 2006;118:333–42.
88. Nomura M, Kitamura M, Matsumiya K, Tsujimura A, Okuyama A, Matsumoto M, et al. Genomic analysis of idiopathic infertile patients with sperm-specific depletion of CD46. *Exp Clin Immunogenet* 2001;18:42–50.

89. Nishioka, K., Kawamura, K., Hirayama, T., Kawashima, T., Shimada, K., 1976.  
The complement system in tumor immunity: significance of elevated levels of complement in tumor bearing hosts. *Ann. N. Y. Acad. Sci.* 276, 303–315.
90. Yamakawa, M., Yamada, K., Tsuge, T., Ohrui, H., Ogata, T., Dobashi, M., Imai, Y., Protection of thyroid cancer cells by complement-regulatory factors. *Cancer*, 1994, 73, 2808–2817.
91. Niehans GA, Cherwitz DL, Staley NA, Knapp DJ, Dalmaso AP. Human carcinomas variably express the complement inhibitory proteins CD46 (membrane cofactor protein), CD55 (decay-accelerating factor), and CD59 (protectin) *Am J Pathol.* 1996;149:129–42.
92. Jurianz, K., Ziegler, S., Garcia-Schuler, H., Kraus, S., Bohana-Kashtan, O., Fishelson, Z., Kirschfink, M., Complement resistance of tumor cells: basal and induced mechanisms. *Mol. Immunol.* 1999b, 36, 929–939.
93. Jurianz, K., Ziegler, S., Donin, N., Reiter, Y., Fishelson, Z., Kirschfink, M., 2001. K562 erythroleukemic cells are equipped with multiple mechanisms of resistance to lysis by complement. *Int. J. Cancer* 93, 848–854.
94. Fishelson Z, Donin N, Zell S, Schultz S, Kirschfink M. Obstacles to cancer immunotherapy: expression of membrane complement regulatory proteins (mCRPs) in tumors. *Mol Immunol* (2003) 40:109–123
95. Junnikkala S, Hakulinen J, Jarva H, Manuelian T, Bjørge L, Bützow R, Zipfel PF, Meri S. Secretion of soluble complement inhibitors factor H and factor H-like protein (FHL-1) by ovarian tumour cells. *Br J Cancer.* 2002 Nov 4;87(10):1119-27.

96. Kraus, S., Fishelson, Z., Cell desensitization by sublytic C5b-9 complexes and calcium ionophores depends on activation of protein kinase C. *Eur. J. Immunol.* 2000 30, 1272–1280.
97. Kraus, S., Seger, R., Fishelson, Z., Involvement of the ERK mitogen activated protein kinase in cell resistance to complement-mediated lysis. *Clin. Exp. Immunol.* 2001 123, 366–374.
98. Hourcade D, Garcia AD, Post TW, Taillon-Miller P, Holers VM, Wagner LM, Bora NS, Atkinson JP. Analysis of the human regulators of complement activation (RCA) gene cluster with yeast artificial chromosomes (YACs). *Genomics.* 1992 Feb;12(2):289-300.
99. (104) Yamamoto H, Masters JR, Dasgupta P, Chandra A, Popert R, Freeman A, Ahmed A. CD49f is an efficient marker of monolayer- and spheroid colony-forming cells of the benign and malignant human prostate. *PLoS One.* 2012;7(10):e46979. doi: 10.1371.
100. Changyong Guo, Haibo Liu, Bao-Hui Zhang, Radu M. Cadaneanu, Aqila M. Mayle, Isla P. Garraway; Epcam, CD44, and CD49f Distinguish Sphere-Forming Human Prostate Basal Cells from a Subpopulation with Predominant Tubule Initiation Capability, April 13, 2012 DOI: 10.1371
101. Ravi A, Gurtan AM, Kumar MS, Bhutkar A, Chin C, Lu V, Lees JA, Jacks T, Sharp PA; Proliferation and tumorigenesis of a murine sarcoma cell line in the absence of DICER1. *Cancer Cell.* 2012 Jun 12;21(6):848-55

102. Jordan M. Cummins, Yiping He, Rebecca J. Leary, Ray Pagliarini, Luis A. Diaz, Jr., Tobias Sjoblom, et al.; The colorectal microRNAome, Proc Natl Acad Sci U S A. 2006 March 7; 103(10): 3687–3692.
103. Karreth A. F., Reschke M., Ruocco A., Ng C., Ala U., Lèopold Valentine, Seitzer N., Langellotto F., Rodig S.J., Chiarle R., Pandolfi PP.; The BRAF Pseudogene is a proto-oncogenic competitive endogenous RNA. (**under revision**).
104. Zou M, Baitei EY, Alzahrani AS, et al. Oncogenic activation of MAP kinase by BRAF pseudogene in thyroid tumors. Neoplasia 2009;11:57–65