ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ ΥΓΕΙΑΣ
ΤΜΗΜΑ ΒΙΟΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΕΘΝΙΚΟ ΙΔΡΥΜΑ ΕΡΕΥΝΩΝ
ΙΝΣΤΙΤΟΥΤΟ ΒΙΟΛΟΓΙΑΣ, ΦΑΡΜΑΚΕΥΤΙΚΗΣ ΧΗΜΕΙΑΣ ΚΑΙ ΒΙΟΤΕΧΝΟΛΟΓΙΑΣ

ΔΙΙΔΡΥΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΒΙΟΕΠΙΧΕΙΡΕΙΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

‘’Έγκαιρη διάγνωση καρκίνου στο μηχανισμό δράσης της ασθένειας’’

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ: Καθηγητής, Λεωνίδας Δημήτριος

Γεωργουσάκης Χρήστος
Α.Μ:0003
Αθήνα, 2017

UNIVERSITY OF THESSALY
SCHOOL OF HEALTH SCIENCES
DEPARTMENT OF BIOCHEMISTRY AND BIOTECHNOLOGY

NATIONAL HELLENIC RESEARCH FOUNDATION
INSTITUTE OF BIOLOGY, MEDICINAL CHEMISTRY & BIOTECHNOLOC .

INTERISTITUTIONAL PROGRAM OF POSTGRADUATE STUDIES IN BIOENTREPRENEURSHIP

MASTER THESIS

''Disease mechanism based diagnosis of cancer''

SUPERVISOR: Professor Demetres D. Leonidas

Christos Georgousakis
A.M.0003
Athens, 2017

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο σπουδών για την απόκτηση του Μεταπτυχιακού Διπλώματος Ειδίκευσης στο

ΒΙΟΕΠΙΧΕΙΡΕΙΝ

που απονέμει το Τμήμα Βιοχημείας και Βιοτεχνολογίας του Πανεπιστημίου Θεσσαλίας, σε συνεργασία με την εταιρεία iCellate.

Εγκρίθηκε την ..... από την τριμελή εξεταστική επιτροπή:

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ

ΟΝΟΜΑΤΕΠΩΝΥΜΟ	**ΒΑΘΜΙΔΑ**	**ΥΠΟΓΡΑΦΗ**
(ΛΕΩΝΙΔΑΣ ΔΗΜΗΤΡΙΟΣ)	ΚΑΘΗΓΗΤΗΣ	
(ΣΚΑΜΝΑΚΗ ΒΑΣΙΛΙΚΗ)	ΛΕΚΤΟΡΑΣ	
(ΨΑΡΡΑ ΑΝΝΑ-ΜΑΡΙΑ)	ΕΠΙΚ.ΚΑΘΗΓΗΤΡΙΑ	

Table of contents.

Abstract

One of the main common cancer therapy problems is late detection. Circulating tumor cells (CTCs) invade the lymphatic system, causing metastases. Isolation of CTC's is a difficult process while current methodology is not always productive. However, genetic sequencing of isolated CTC's, can give us the panorama of the gene mutations, which are similar to those from the solid tumor they originated. Therefore, cancer treatment can be designed based on each patient's molecular profile of the disease. In the framework of this Thesis, collections of identified cancer driver genes were screened to identify drugs targeting specific cancer drivers and the number of anticancer agents that are currently on clinical trials. Successful isolation of CTC's from blood samples can be tedious. A new CTC detection method based on the telomerase activity in cancer cells has been used. Firstly, MCF7 and MDA-MB-231 cell lines were infected with the virus vector OBP-401, to test the efficiency of the detection method through the expression of a Green Fluorescent Protein (GFP). The procedure was further optimized by testing patient's blood samples. Although the detection was successful, the ratio of living cells to GFP positive cells was low while the number of GFP positive cells was much lower than expected. Nevertheless CTC'S were isolated successfully and through single cell genetic sequencing, the number of mutations for each patient has been obtained. By using the aforementioned cancer driver genes database, patients were able to find out whether their oncogenic mutations could be treated or whether they are on a clinical trial level, paving the way for individualized cancer treatment.

Keywords: Cancer driver genes, CTC's, telomerase activity, OBP-401.

Aim

The aim of thesis was to produce a database of drugs that target specific mutations on cancer driver genes which may be employed for personalized cancer therapy. This database may also serve as an additional diagnostic test when combined with mutations on circulating tumor cells for any clinically new treatment target. Thus, we provided a census of human cancer genes, including those that are currently considered druggable. For the experimental part, we provided an orthogonal verification that isolated single cells are viable, potentially immortal (conforming to cancer diagnosis), and providing triaging for genetic sequencing. Cancer cells for single cells project, were verified by the use of hTERT promoter in adenovirus vector with EGFP reporter (Oncolys Biopharma,Telomescan),combined with cancer cell detection and verification.

1. Introduction.

Cancer is a group of more than 100 distinctive diseases. Tumorigenesis is a process with many steps including alterations in cell genotypes (Diaz-Cano, 2012). Essential alterations in cell physiology include self-sufficiency in growth signals, insensitivity to anti-growth signals, evading apoptosis, limitless replicative potential, sustained angiogenesis tissue invasion and metastases (Futreal et al. 2004), deregulated metabolism, evading immune system, genome instability and inflammation (Futreal et al. 2010). Most human cancers are caused by two to eight alterations that develop over the course of 20 to 30 years (Vogelstein et al. 2013). Cancer originates at the single cell level but also spreads by way of one or more single cells from its tissue of origin to neighboring lymph nodes or to far distant organs through blood stream forming metastases. Furthermore, cancer cells can invade existing tumors with new variants of the original cancer clone, including new sub clones that are treatment resistant (Martin et al. 2013).

Cancer starts with a single cell; therefore cancer therapy needs to be based on properties of single tumor cells. Results of genetic sequencing showcase the limited number of cancer driver mutations. A driver mutation is causally implicated in oncogenesis. It has conferred growth advantage on the cancer cell and has been positively selected in the microenvironment of the tissue in which the cancer arises. A passenger mutation has not been selected, has not conferred clonal growth advantage and has therefore not contributed to cancer development (Stratton et al. 2009). A central goal of the cancer genome analysis is the identification of cancer driver genes. The main problem is to distinguish driver from passenger mutations. Driver mutations cluster in the subset of genes that are cancer genes whereas passenger mutations are more or less randomly distributed (Campbell et al. 2009).

This is important information in fighting cancer. Targeting cancer in the level of individual cells increases the chances of successful treatment. In order to do that, we need to know the specific genes that are mutated in each individual cancer. Close examination of individual genes, has identified a relatively small number of genes responsible for cancer (out of the 20,000 genes in the human genome). The cancer gene census is still ongoing. An individual cancer is caused by various combinations of 2 to 8 out of the total of hundred cancer genes. With single cell genetic sequencing, we can discover what gene mutations need to be targeted for treatment in each individual patient. As we complete the census of potential cancer genes and mutations and as the companies develop new drugs targeting them, we can treat the root causes of cancer, based on individual cell information (Ericsson et al. 2016).

2. Cancer driver genes.

The identification of new cancer driver genes constitutes a scientific breakthrough due to the challenges which occur from the heterogeneity of somatic mutations and the number of genes mutated at a low frequency. Cancer driver genes can be categorized into three groups: 1) gain of function (oncogenes), 2) loss of function (tumor suppressor genes), 3) switch of function, which promotes malignant transformation.

A central aim of cancer research has been to identify the mutated genes that are implicated in oncogenesis. COSMIC, the Catalogue of Somatic Mutations in Cancer and IntOGen-mutations platform, are the world's largest and most comprehensive resources for exploring the impact of somatic mutations in cancer.

2.1 Cosmic cancer genes census.

COSMIC (http://cancer.sanger.ac.uk/cancergenome/projects/cencus/) is an online database of somatically acquired mutations found in human cancer. It is the largest catalogue of exploring the impact of somatic mutations in cancer. COSMIC can accommodate information on base substitutions, insertions and deletions, translocations and changes in copy number. Data sources are identified from the published literature and online data portals. Over 300 cancer genome publications have now been curated, and COSMIC includes substantial data sets from The Cancer Genome Atlas and International Cancer Genome Consortium project (Pleasance et al. 2010).

Approximately half of COSMIC's cancer genomes are curated from these consortium data portals, the other half from curations of published literature. The details of samples and disease descriptions are curated into COSMIC manually, and the mutations, usually supplied as genomic co-ordinates, are annotated automatically via a software pipeline. Upon selection of a gene from the Census for full expert curation, all papers mentioning its mutation in human cancer are collected and exhaustively curated before it is released into a new version of COSMIC (Forbes et al. 2014).

Once this initial curation is released, the gene is updated as significant new information is published. Each curator is responsible for a defined set of 60 or more genes, developing substantial expertise. In parallel, cancer genomes are curated via a more bioinformatic

approach. Genomic data is obtained half from published supplementary information tables, and the other half from genome consortia such as TCGA, ICGC. Such molecular profiling includes point mutations, gene fusions, copy number annotations, structural breakpoints, gene expression and CpG island methylation variants (Forbes et al. 2014).

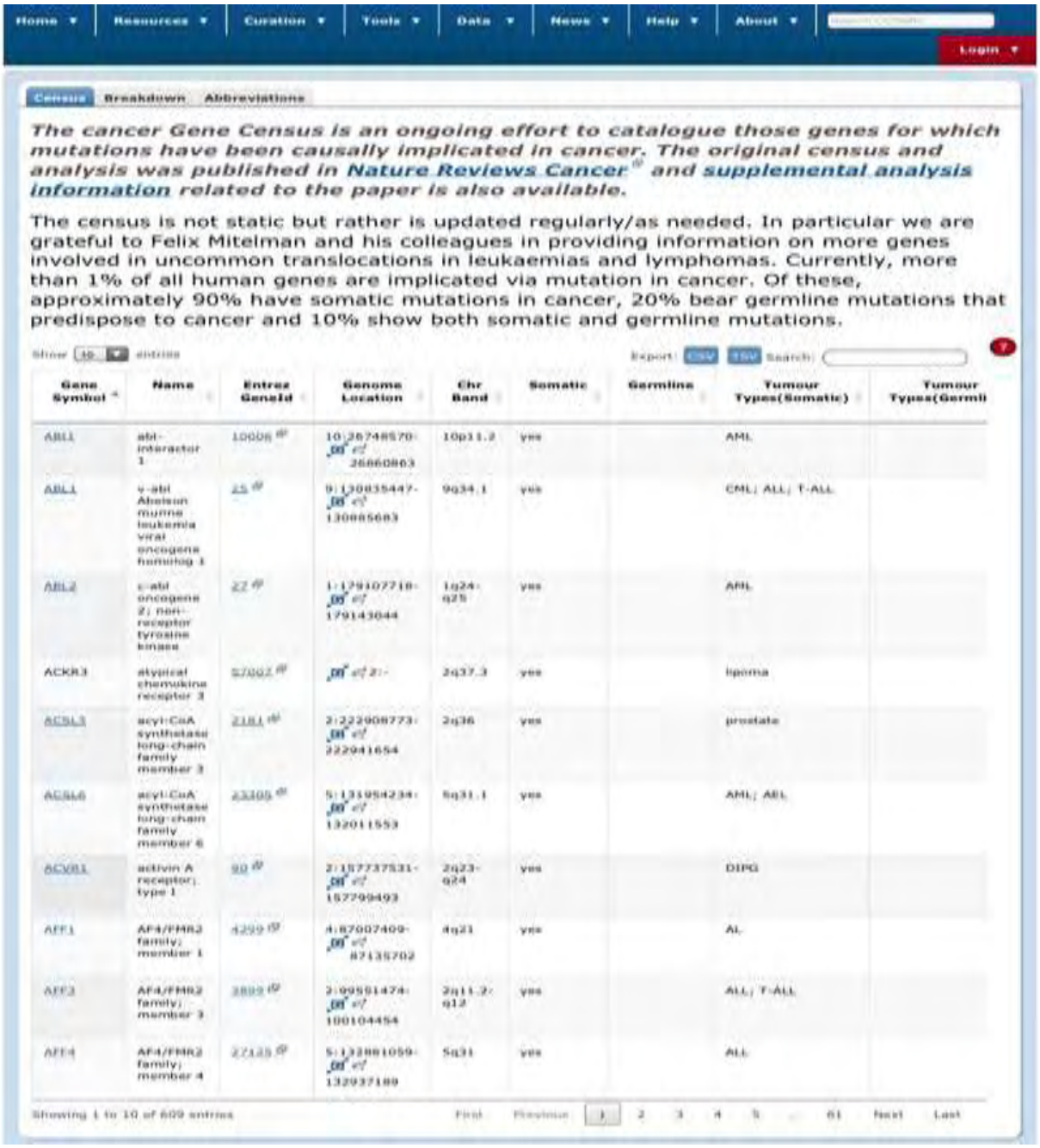

***Figure 1**: COSMIC list of driver genes as it is shown in the front page of the website.*

2.2 IntOGen.

The IntOGen-mutations platform (http://www.intogen.org) summarizes somatic mutations, genes and pathways involved in tumorigenesis. It identifies and visualizes cancer drivers, analyzing 4,623 exomes from 13 cancer sites. It provides support to cancer researchers, aids the identification of drivers across tumor cohorts and helps rank mutations for better clinical decision-making. The IntOGen-mutations pipeline integrates the results of tumor genomes analyzed with different mutation-calling workflows and is scalable to hundreds of thousands of tumor genomes (Gonzalez-Perez et al. 2013).

It currently includes OncodriveFM7, a tool that detects genes that are significantly biased toward the accumulation of mutations with high functional impact (FM bias) without the need to estimate background mutation rate, and OncodriveCLUST9, which picks up genes whose mutations tend to cluster in particular regions of the protein sequence with respect to synonymous mutations (CLUST bias) (Online Methods). Both tools detect signals of positive selection, which appear in genes whose mutations are selected during tumor development and are therefore likely drivers (Araya et al. 2016).


Table

Show 10 entries

Search:

Symbol	Mutated samples (CS)	Mutated samples (PAM)	% Mutated samples (PAM)
TP53	2227	2187	32.20
PIK3CA	789	747	11.00
KRAS	544	533	7.85
BRAF	498	472	6.95
PTEN	452	440	6.48
MLL3	451	391	5.76
APC	367	336	4.95
MLL2	362	300	4.42
ARID1A	321	290	4.27
NF1	354	286	4.21

***Figure 2**: IntOGen list of driver genes as it is shown in the front page of the website*

2.3 Candidate cancer gene database.

The CCGD (http://ccgd-starrlab) is a manually curated database containing a unified description of all identified candidate driver genes and the genomic location of transposon common insertion sites (CISs) from all currently published transposon-based screens. The current version includes data and results from 28 publications covering 40 individual screens. All data have been manually curated and genomic loci have been updated to the current genome build. Searches can be by gene, study or cancer type. This allows users to determine if a gene of interest is a putative cancer driver gene and quickly generate a list of driver genes that have been identified in a particular tumor type (Abott et al. 2015).

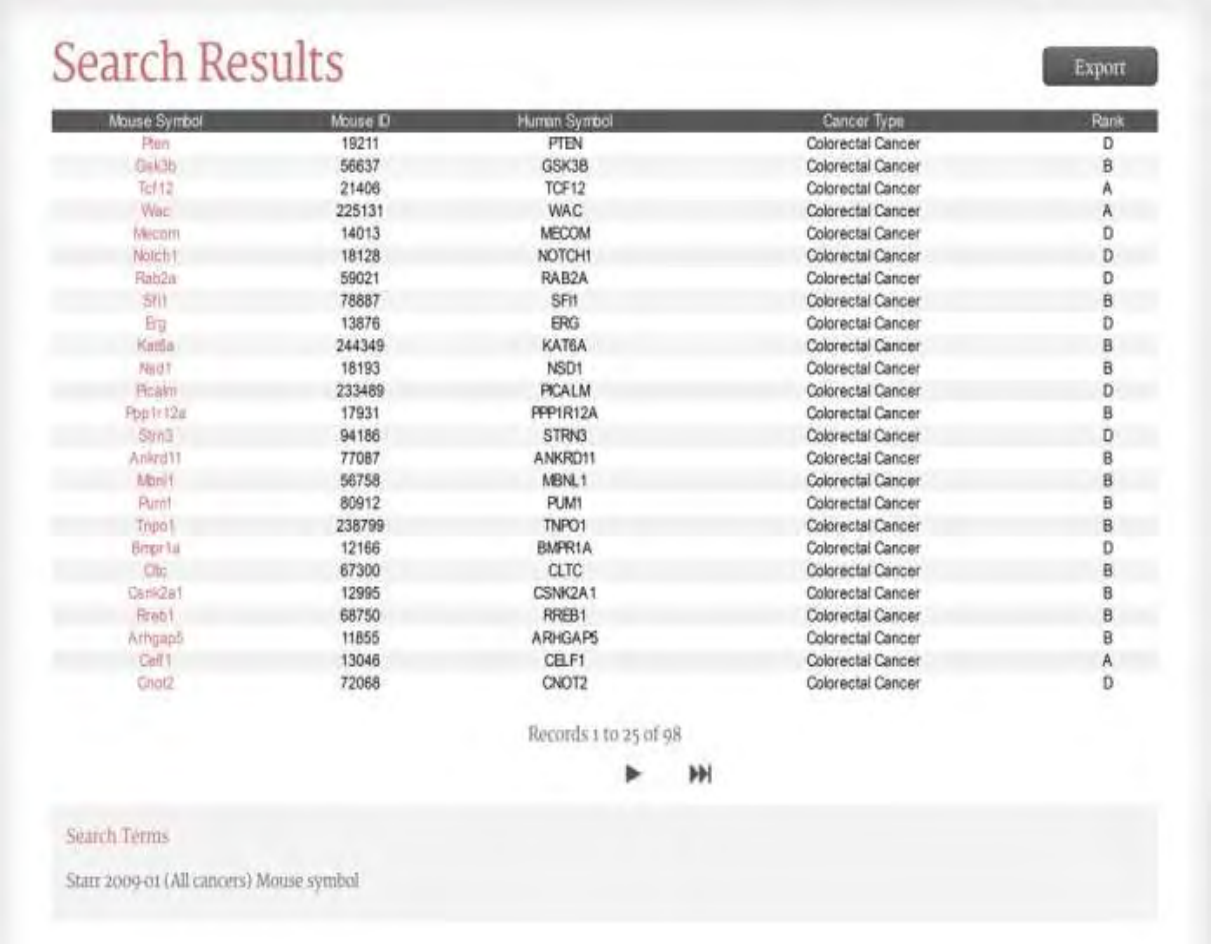

***Figure 3**: CCGD list of driver genes as it is shown in the front page of the website.*

2.4 Differences between the databases.

One of the main differences between the three databases is the way the genes are selected in order to be included in each table. In cosmic cancer gene census, there is a manual approach that helps the capture of very high detail across mutation positions, disease descriptions and other patient and population data (such as age, ethnicity and therapeutic regime). Experienced curators identify inconsistencies or errors in publications, allowing the rejection of untrustworthy, incomplete or unspecific data sources. New genes are included in only when curation of their literature is exhausted, and the mutation patterns are as up-to-date as possible. Complementary to the manual curation effort, a semi-automated approach has been developed for curation of large cancer genome (and exome) data sets. The details of samples and disease descriptions are curated into COSMIC manually, and the mutations, usually supplied as genomic co-ordinates, are annotated automatically via a software pipeline using Ensembl genome annotations. This utilizes custom software similar to the Variant Effect Predictor to identify the positions of coding mutations as well as consequence annotations (Forbes et al. 2015).

On the other hand, IntOGen instead of using mutation recurrence measurements detects other signals of positive selection. For instance, OncodriveFM detects genes in which observed mutations in tumors are biased towards high functional impacting mutations. One important advantage of this method is that it does not depend on the background mutation rate. OncodriveCLUST, on the other hand, detects genes in which mutations are clustered in certain protein regions. An important feature of our methods is that they only require the list of somatic mutations as input and they are computationally inexpensive, both of which are important features when analyzing large cohorts of tumors like we do in IntOGen-mutations (Gonzalez-Perez et al. 2017).

As far it concerns the candidate cancer gene database, CCGD is a manually curated database and the genes are identified in mouse models in all published transposon-based forward genetic screens for cancer. If the frequency of transposon insertions in a given genomic region is higher than expected by chance, the region is called a CIS and the genes within or near this region are identified as candidate cancer genes (Abott et al. 2015).

Another also significant difference between the first two databases, is the number of genes included in each database. In COSMIC, there is a total number of 609 cancer genes in comparison with IntOGen which includes a total number of 459 genes. In the two lists, only 185 genes are in common (Figure 4).

ABL2	ATRX	CDK4	DNMT3A	FGFR2	IDH2	MKL1	NRAS	RAD21	TCF7L2
ACSL3	AXIN1	CDKN1B	EGFR	FGFR3	KAT6B	MLH1	NSD1	RBM10	TET2
ACSL6	AXIN2	CDKN2A	EIF4A2	FIP1L1	KDM5C	MLLT4	NUP98	RHOA	TGFBR2
ACVR1B	B2M	CEBPA	ELF4	FLT3	KDM6A	MTOR	PAX5	RUNX1	THRAP3
AFF4	BAP1	CHD4	EP300	FOXA1	KDR	MYB	PBRM1	SETD2	TP53
AKAP9	BCL11A	CHEK2	ERBB2	FOXP1	KEAP1	MYC	PER1	SF3B1	TSC1
AKT1	BCOR	CIC	ERBB2IP	FUBP1	KIT	MYCN	PHF6	SMAD2	U2AF1
ALK	BLM	CIITA	ERBB3	FUS	KLF4	MYD88	PIK3CA	SMAD4	USP6
APC	BMPR2	CLTC	ERCC2	GATA3	KLF6	MYH11	PIK3R1	SMARCA4	VHL
ARHGAP26	BRAF	CNOT3	EZH2	GNA11	KRAS	MYH9	PLCG1	SMARCB1	WHSC1
ARID1A	BRCA1	CREBBP	FAM123B	GNAS	LCP1	NCOR1	PPM1D	SPOP	WHSC1L1
ARID1B	BRCA2	CRTC3	FAM46C	GOLGA5	MAP2K1	NCOR2	PPP2R1A	SRGAP3	WT1
ARID2	C15orf55	CTCF	FANCI	HLA-A	MAP2K4	NDRG1	PPP6C	STAG2	XPO1
ARNTL	CASP8	CTNNB1	FAS	HLF	MAP3K1	NF1	PRKAR1A	STK11	ZFHX3
ASXL1	CBFB	CUX1	FAT1	HNF1A	MAX	NF2	PRRX1	SUZ12	
ATF1	CCND1	CYLD	FBXO11	HRAS	MECOM	NFE2L2	PSIP1	SYK	
ATIC	CDC73	DDX3X	FBXW7	HSP9OAA1	MED12	NOTCH1	PTCH1	TBL1XR1	
ATM	CDH1	DDX5	FCRL4	HSP9OAB1	MEN1	NOTCH2	PTEN	TBX3	

Figure 4: *List of driver genes that are in common in COSMIC and IntOGen platforms.*

3. Overview of targeted cancer therapies.

3.1 Cancer driver genes as targeted cancer agents.

The discovery of a single driver gene is the first step in the development of targeted or immune therapies. While the discovery of the complete set of driver genes will take time, the identification of new cancer driver genes is a scientific breakthrough with limited benefits for cancer patients unless it can be targeted by a drug. However, designing a drug is even harder than identifying a cancer gene in the first place, and it is getting harder every year.

The cost of developing a new drug doubles approximately every nine years. During the last 20 years less than 200 cancer drugs have been approved by FDA (US Food and Drug Administration). This is due to the varying degrees of toxicity and the cancer heterogeneity in each patient. Targeted cancer therapies are mostly drugs or other substances that interfere with specific targets in order to stop the growth and spread of cancer. They are designed to interact with a specific target (molecular target) and that is why they sometime are called molecularly targeted therapies.

Targeted cancer agents are classified as either monoclonal antibodies or small molecules. Monoclonal antibodies target specific antigens on the surface of the cell in comparison with small molecules that penetrate the membrane to interact with targets inside the cell (interference with the enzymatic activity of the target protein). The FDA has approved multiple targeted drug cancer therapies.

3.2 FDA approved drugs for cancer driver genes.

Current targeted therapies can be categorized in three different approaches: 1) direct targeting in order to inhibit activated cancer driver genes 2) indirect targeting through the inhibition of non-altered proteins connected to the altered drivers and 3) therapies that compensate the activity loss of a tumor suppressor gene. Unfortunately, current therapies cannot replace the function of a tumor suppressive gene. Furthermore, the majority of the approved drugs cannot interfere with drivers that participate in protein complexes. Also, genetic heterogeneity among the cells can have a huge impact in patients' response to targeted cancer therapy.

The FDA has approved over the past 20 years less than 200 cancer drugs due to the cost of developing new drugs in addition with the varying degrees of toxicity, pervasive institutional regulation and patient heterogeneity. In order to identify how many of these cancer drugs have specific targets (cancer driver genes that are included in COSMIC and IntoGen). A list exported from the FDA official site (http://www.fda.gov) comprising a total of 609 genes in COSMIC and 459 genes in IntoGen, was screened leading to the discovery of only 25 genes that are specific targets for cancer drugs that have been approved by the FDA (**Figure 5**).

DRUG	GENES
1. AFANITIB	EGFR, HER2
2. AXINITIB	KIT,PDGFR
3. BELINOSTAT	HDAC
4. BOSUTINIB	ABL
5. CABOZANTINIB	FLT3,KIT,MET,RET
6. CERITINIB	ALK
7. CETUXIMAB	EGFR
8. COBIMETINIB	MEK
9. CRIZOTINIB	ALK,MET,ROS1
10. DABRAFENIB	BRAF
11. DASATINIB	ABL,KIT,SRC
12. ERLOTINIB	EGFR
13. EVEROLIMUS	mTOR
14. GEFINITIB	EGFR
15. IBRUTINIB	BTK
16. IDELALISIB	PI3K
17. IMATINIB	ABL,PDGFR
18. LAPATINIB	HER2
19. NECITUMUMAB	EGFR
20. NINTEDANIB	PDGFR
21. NILOTINIB	ABL
22. OLAPARIB	BRCA1/2
23. OLARATUMAB	PDGFR
24. OSIMERTINIB	EGFR
25. PALBOCICLIB	CDK4/6
26. PANITUMUMAB	EGFR
27. PANOBINOSTAT	HDAC
28. PAZOPANIB	KIT,PDGFR,RAF,RET
29. PERTUZUMAB	HER2
30. PONATINIB	ABL,FGFR,FLT3
31. REGORAFENIB	KIT,PDGFR,RAF,RET
32. ROMIDESTIN	HDAC
33. RUCAPARIB	BRCA1/2
34. RUXOLITINIB	JAK
35. SIROLIMUS	mTOR
36. SONIDEGIB	SMO
37. SORAFENIB	BRAF,FLT3,PDGFR,RAF,RET
38. SUNITINIB	FLT3,KIT,PDGFR,RET

39. TEMSIROLIMUS	mTOR
40. TRAMETINIB	MEK
41. TRASTUZUMAB	HER2
42. VANDETANIB	RET
43. VEMURAFENIB	BRAF
44. VENETOCLAX	BCL-2
45. VISMODEGIB	PTCH,SMO
46. VORINOSTAT	HDAC

Figure 5*. List of current FDA approved drugs targeting specific cancer driver genes. The list includes only drugs that are used against these cancer drivers as primary targets. Cancer drivers that are targeted by these FDA approved drugs as secondary targets via metabolic pathways are not included.*

3.3 Clinical trials.

Clinical trials provide the basic evidence of efficacy and outcome that regulators, providers, and insurance companies rely on when they determine whether or not to approve, prescribe, or reimburse a patient for taking a drug. Trials are incredibly slow and expensive for a myriad of reasons, but perhaps the most important one that relates to our lack of rapid therapeutic progress is the mismatch between disease, trial design, and patient enrollment (Lefkofsky, 2016).

Clinical trials are not often based on the molecular profile of each patient. The main criteria for the enrollment of a patient, is only if it fits the phenotype of the disease that is under examination. But, each patient's response in the trial is often based on these molecular characteristics such as point mutations. Furthermore, trials are so expensive and as a result companies prefer to avoid small or targeted trials such as drugs for specific cancer driver genes. The percentage of people that carry the specific mutations that are under investigation in each trial is very low. So it is rather difficult to find enough patients to conduct the trials. Also, a trial for drug that targets a specific cancer driver gene implies a small market size. The necessary investments for clinical trials are huge, so each company wants a market that can actually support the up-front cost of the clinical trials. But due to cancer's heterogeneity, in the near future in order to be able to treat cancer effectively, we need to design tailor-made drugs, especially for rare cancer driver genes mutations. Though, targeting specific driver genes, can be a huge step in cancer treatment, combination therapies hold even a bigger promise to each patient's unique molecular composition.

Clinical trials determine which new investigational drugs are allowed to enter medical practice, but the unit on which value is measured in a clinical setting is not a drug, but a treat-

ment pathway. Currently, many anticancer agents are studied as potential drugs against cancer drivers (http://clinicaltrials.gov). The table below presents the genes from the 609 genes of COSMIC catalogue and the 459 genes of IntOGen catalogue, associated with a specific anticancer agent that is currently on clinical trials (**Figure 6**).

GENES	DRUGS
AKT	cenisertib, ipatasertib ,afuresertib, uprosertib, ARQ751, AZD5363.
ALK	dalantercept, brigatinib, alectinib(approved in japan), lorlatinib,entrectinib.
ATR	VX-803,VX-970
BCL-2	navitoclax,,obatoclax,oblimerson,rosomindar. AT101
BRAF	ARQ736,BGB-283,RAF265,PLX3603
BRACA1/2	BMN673(talazoparib),
BRD3/4	GSK525762.
BTK	acalabrutinib,cenisertib,AVL-292
CCND1	briciclib
CD27	Varlilumab
CD74	Milatuzumab
CDK4/6	Ribociclib,abemaciclib
CDKN2A	Iloraserib
CHEK2	Rabuserib
DDR2	Sitravatinib
EGFR	AEE788,brigatinib,naquotinib,vandetanib,Icotib,canertinib,rociletinib,epitinib,theliatinib Olmutinib,pelitinib,poziotinib,dacomitinib,,Modotuximab,depatuxizumab,nimotuzumab Duligotuzumab,AZD8931,AC480,BMS-690514
FGFR1/2/3	Masitinib,lucitanib,sulfatinib,dovinitib,Erdafinitib,orantinib,AZD-4547,BGJ398 DEBIO1347,XL228,enmd-2076,HGS1036,LY28774455
FLT3	Quitinib,cenicertib,gilteritinib,lestaurtinib,,Crenolanib,tandutinib,amuvattinib,midostaurin Enmd-2076,PLX3397
HDAC	Resminostat,citarinostat,abexinostat,entinostat,retinostat,ricolinostat,mcetinostat
HER2	AEE788,tucatinib,canertinib,neratinib,XL647,Mubritinib,glycooptimized trastuzumab,margetixumab,MM-111-5422
HSP90	Tanespimycin,onalespib,luminespib,ganetespib,Retaspimycin,debio0932,SNX- 5422,XL888
IDH1/2	AG-120,enasidenib
JAK1/2	PRT062070,AT9283,BMS-911543,XL019,Itacitinib,lestaurtinib,fedratinib,pancritinib, gandotinib.
KIT	Masitinib,motesanib,cenisertib,telatinib,XL820,Sitravatinib,tandutinib,amuvatini, modostaurin
KRAS	BGB-283
MAPK1	BVD-523,MK-835m3
MDM2	idasanutlin
MEK	AZD8330,BI847325,CL-1040,GDC-0623,PD0325901,ROS5126766,TAK-733
MET	AMG208,AMG337,BMS-777607,foretinib,EMD1214063,PF042217903,SAR125844 ABT-700,volitinib,glesatinib,sitravatinib,amuvatinib,emibetuzumab,onartuzumab
MTOR	Ridaforolimus,dactolisib,apitolisib,sapanisertibGedatolisib,voxtalisib,AZD2014,AZD8055 BGT226,CC-223,OSI-027,PF-4691502PI103,PWT33597,SF1126
NOTCH1	brontictuzumab
NRAS	BGB-283

NTRK1/2/3	AZD7451,LOXO-101,PLX7486,TSR-011Sitravatinib,entrectinib,lestaurtinib
PDGFR	Iloraserib,motesanib,pazopanib,sitravatinib,Tandutinib,X-82
PDGFRA/B	Crenolanib, amuvatinib,linifanib,axitinib,Sorafenib,telatinib,regorafenib,sunitinib Orantinib,tovetumab,XL820
PI3K	panlisib,dactolisib,buparlisib,alpelisib,Pictilisib,apitolisib,omipalisib,gedatolisib Pilaralisib,sophoretin,quercetin,PI-103,GSK2636771,PF-4691502,PWT33597 PX-866, SF1126.
PKC	Sophoretin,enzastaurin
RAF	Encorafenib,MLN2840,ROS5126766,XL281
RET	Motesanib,sitravatinib,amuvatinib
ROS1	Iorlatinib,entrectinib
SMO	Patidegib,taladegib
STAT3	OPB-31121
SRC	Sracatinib,ilorasertib,KX2-391,XL228
SYK	Entospletinib,fostamatinib,PRT062070
TGFBR1/2	Galunisertib,IMC-TR1
XPO1	Selinexor

***Figure 6**. List of anticancer agents that are currently on clinical trials against specific cancer drivers as primary targets.*

3.4 FDA versus EMA.

The evolution of European regulation of DADs, by contrast, is much more recent, with significant changes after the formation of EU in 1993. Before EU regulation and marketing approval for DADs fell to each member state. Differences in regulations among the states often impeded marketing and disbursement of DADs across Europe, and in some cases fostered “protectionist” legislation within states to shield sovereign nations' companies from fierce market competition. Among the current 28 member states, many interstate agencies have been reorganized. Clinical trial applications are generally handled in the member state, whereas marketing applications are approved by both state and central agencies in accordance with regulations set forth by the EC. After clinical trials, FDA drug approvals follow a centralized path, whereas European approval can occur through 4 different paths, depending on the nature of the drug and the preference of the manufacturer(Norman et al. 2016) (**Figure 7**).

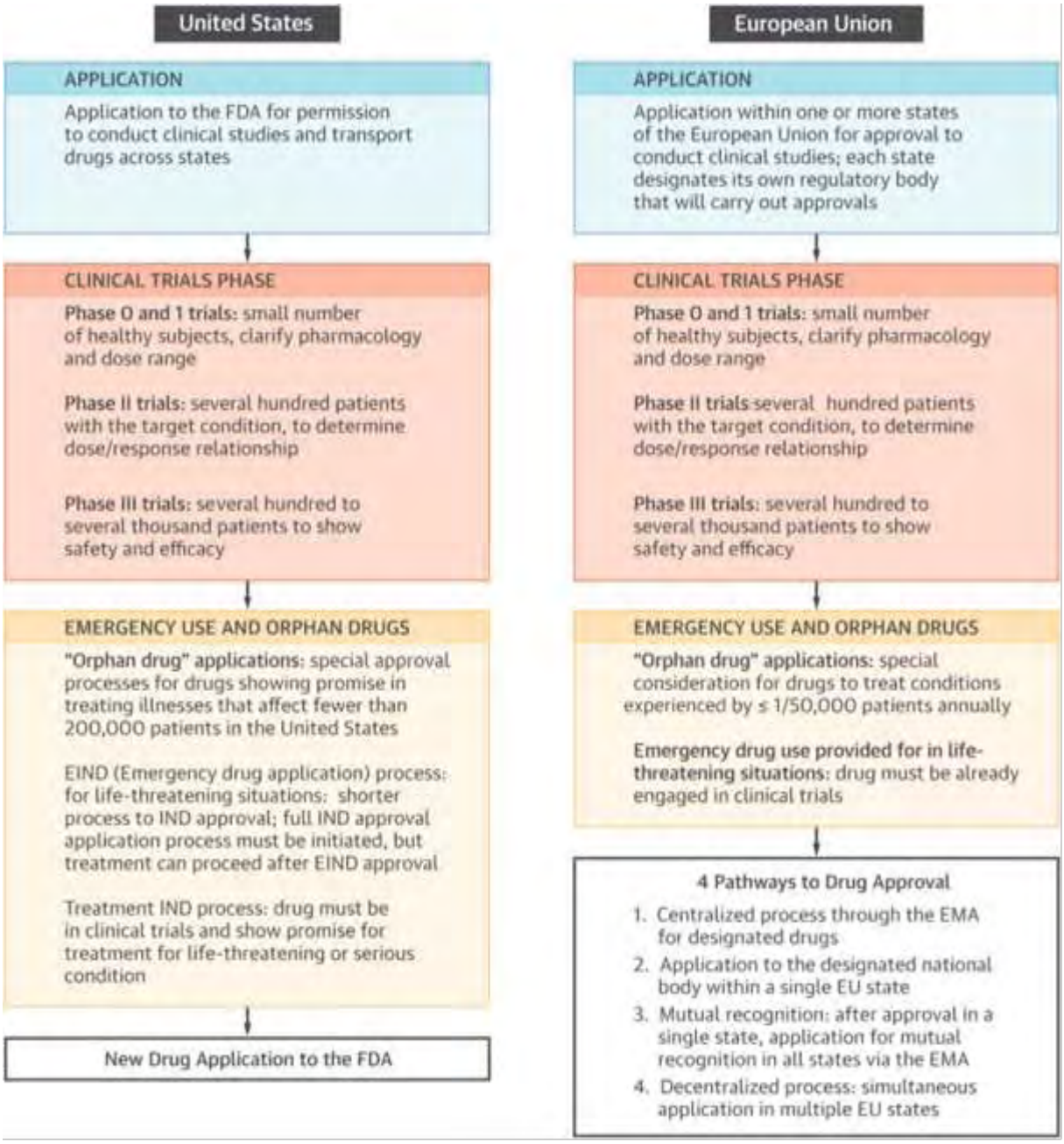

Figure 7*: Differences between the procedures for drug approval in USA and Europe. Taken from Van Norman et al. 2016.*

The FDA and the EMA regulatory systems, share common goals. Despite the main effort of each organization to approve drugs which can reach the public more quickly, the motives behind the approvals appear to have huge dissimilarities. Drugs appear to be approved more quickly in Europe than in the United States. However, analysis indicate, that they actually reach the commercial market more quickly in United States than in Europe. During the past years, efforts to promote transparency and mutual standardization of DAD approval accesses, are currently underway (Norman et al. 2016).

	EMA[a]	FDA
Axitinib	503 (401)	288
Bosutinib	n/a [b]	292
Crizotinib	453 (357)	149 [c]
Dasatinib	312 (252)	182 [c]
Erlotinib	389 (301)	111 [c]
Gefitinib	414 (352)	273 [c]
Imatinib	255 (147)	72 [c]
Lapatinib	614 (434)	181 [c]
Nilotinib	410 (350)	396
Pazopanib	472 (356)	304
Regorafenib	n/a [b]	153 [c]
Ruxolitinib	449 (323)	166 [c]
Sorafenib	315 (232)	166 [c]
Sunitinib	323 (240)	168 [c]
Vandetanib	535 (443)	273 [c]
Vemurafenib	290 (226)	111 [c]
Mean (range)	409.6	205.3 (167.1[c])

***Figure 8**: Time of approval of drugs for specific cancer driver genes. [a]Numbers indicate time to final approval from the European Commission, with time to positive opinion from EMA's Committee for Medicinal Products for Human use given in parenthesis, [b]under review at time of publication, [c]Priority review procedure used. Taken from RR Shah et al. 2013.*

4. Circulating tumor cells.

4.1 CTC's role in metastasis.

The major cause of cancer-associated mortality is tumor metastasis. Cancer cells invade the surrounding tissue of the primary tumor into lymphatic and blood circulation systems co-opting them as a transportation vehicle to ultimately metastasize to other parts of the body (Marchetti et al. 2014). Once in the blood, the vast majority of the circulating tumor cells actually die from natural causes, but a few survive. A few of those also attach at distant organs where they may exit the circulation and migrate further into the distant organ site where they may settle down. A fateful few of those dispersed cells may in turn start to grow to found new tumors (Ericsson et al. 2016). Because this process mostly occurs through the blood, circulating tumor cells (CTCs) are of obvious importance and interest. CTCs might serve as an alternative source of material (liquid biopsy") to diagnose cancer patients, in contrast to current invasive and painful tumor biopsies (Marchetti et al. 2014).

The metastatic process includes the suggestion that a reversible epithelial-to-mesenchymal transition (EMT), describing a major phenotypic change in a subset of cells within the primary tumour, is essential for metastasis to proceed (Thiery and Sleeman,2006; Yang *et al*. 2006; Thiery, 2003). During EMT, epithelial tumour cells lose cell-to-cell to contacts and develop a more motile and invasive mesenchymal phenotype, facilitating their entry into the bloodstream, and revert back to an epithelial phenotype upon extravasation in host tissue (the so-called mesenchymal-to-epithelial transition (MET) (Christiansen and Rajasekaran, 2006; Thiery and Sleeman, 2006). Many CTC detection techniques depend on capture of CTCs based on defined epithelial protein expression (e.g. EpCam and cytokeratins); thus, the very process of CTC detection may be inherently flawed if EMT has occurred. However, it is recognized that EMT is not a homogenous 'black and white' cellular scenario and it seems likely that CTCs can express both epithelial and mesenchymal properties, to varying degrees, giving rise to heterogeneous CTC populations (Christiansen and Rajasekaran, 2006).

Detecting and studying CTCs, has still many obstacles to overcome. CTCs inherit a heterogeneous malignant potential to home in on and generate metastasis into secondary organs. Further, technical challenges in the field persist with regard to identifying and interrogating their heterogeneity, additive to fully capturing these rare tumor cells. Many studies have discussed the clinical impact of detecting CTCs, considering that CTC testing is

being employed currently in over 400 clinical trials worldwide. However, much of the information regarding CTCs is missing, and many challenges must be overcome before their clinical potential as biomarkers and therapeutic targets is fulfilled (Marchetti et al. 2014).

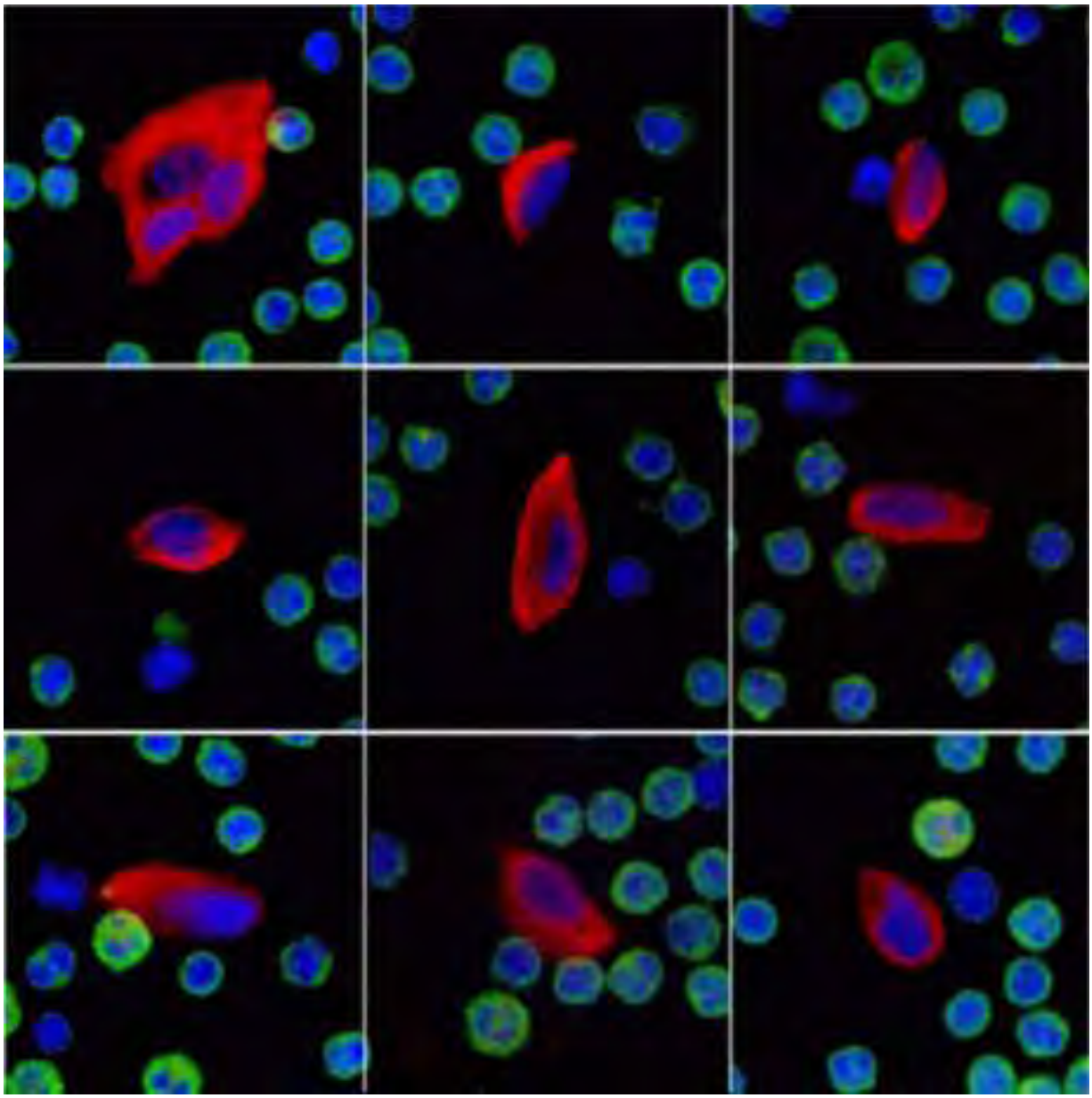

***Figure 9**: Circulating tumor cells, typical of those found in cancer patients. Taken from Marrinucci et al. 2012.*

4.2 CTC's detection methods.

CTCs are captured from the vasculature by using specific antibodies able to recognize specific tumoral markers (usually EpCAM); however this approach is biased by the need for a sufficient expression of the selected protein on the cell surface, event necessary for the enrichment step. Moreover, since EpCAM and other proteins (e.g. cytokeratins) are not expressed in some tumors and can be down regulated during the epithelial to mesenchymal transition (EMT), new enrichment strategies are required (Mikolajczyk et al. 2011). In principle, methods can be divided into nucleic-acid-based and cytometric approaches.

Nucleic-acid-based methods were predominantly adopted throughout the 1990s following the development of the polymerase chain reaction (PCR) and indeed can be very sensitive techniques, relying on the detection of specific DNA or RNA sequences differentially expressed by tumour cells (Alunni-Fabbroni and Sandri, 2010; Paterlini-Brechot and Benali, 2007). However, in recent years there has been a preferential shift toward cytometric assays where cells remain intact, hence morphology can be visualized, cells can be enumerated and further analysis by techniques such as fluorescent *in situ* hybridization (FISH) or even DNA/RNA extraction are practically or theoretically possible, contrasting with the more limited capabilities using nucleic-acid-based methods.

Cytometric approaches use immunostaining profiles to identify and characterize CTCs. These assays need to be highly sensitive, highly specific and highly reproducible if they are to be useful in the clinical setting and used to make patient treatment decisions. Most methods employ an initial enrichment step to optimize the probability of rare cell detection, achievable through immunomagnetic separation, centrifugation or filtration. Cytometric-based techniques subsequently interrogate cells by fluorescence microscopy or immunohistochemistry.

The most widely used cytometric CTC technology currently in clinical testing is the CellSearch™ platform (Veridex LLC, Huntingdon Valley, PA, USA) and is the only technology to have received FDA approval for the enumeration of CTC in whole blood in specific cohorts of cancer patients (Miller *et al*. 2010). The major advantage of this system is its semi-automation and proven reproducibility, reliability, sensitivity, linearity and accuracy (Riethdorf *et al*. 2007; Allard *et al*. 2004). These are features crucial to any biomarker technology to ensure validity of results in clinical testing across multiple sites and have so far been lacking with previous techniques. CellSearch employs immunomagnetic bead-based separation to enrich for CTCs, and the platform characteristics have been described in detail previously (Miller *et al*. 2010; Allard *et al*. 2004).

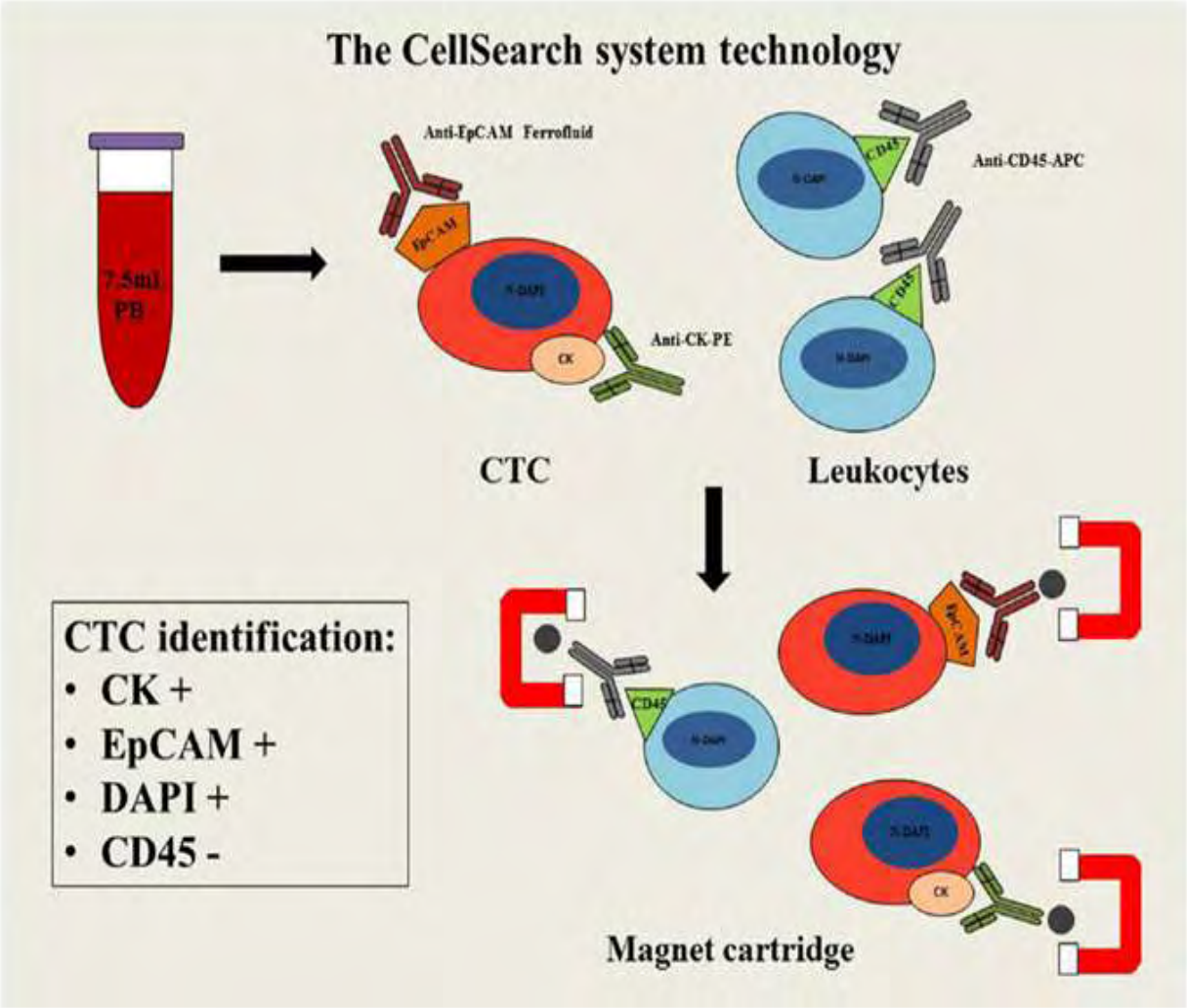

Figure 10: *Using ferrofluid nanoparticles with antibodies that target epithelial cell adhesion molecules, CTCs are magnetically separated from the bulk of other cells in the blood. CTCs are then stained with cytokeratin monoclonal antibodies, which are specific to epithelial cells. A monoclonal antibody stain is used to identify CD45, a marker specific to leukocytes called DAPI is also added to highlight the nuclei of both CTCs and leukocytes. Cells are put in a magnet cartridge that applies a magnetic force that pulls the cells to a single focal depth. The cartridge containing stained CTCs is placed onto the CELLTRACKS ANALYZER II®System for scanning. Once the cartridge has been scanned, the system displays tumor cell candidates that are positive for cytokeratin and DAPI. These candidate cells are presented to an operator for final review. Taken from Truini et al. 2014.*

4.3 CTC's detecting technique using TelomeScan.

Because of the low number of CTCs (1 CTC in 10^6-10^7 leukocytes), there are many problems with the current CTC detection methods. Current methods are based on the concentration of CTCs using anti-CD45 and anti-epithelial adhesion molecule antibodies. The detection is occurred via immunostaining with anti-cytokeratin-8 anti-CK-9 and anti-CK-19 antibodies. Although, the CTCs with these methods can be detected, there are also false negative and false positive results. Many types of tumor cells are negative of EpCAM or the CK molecules. Furthermore, these antigens are also expressed on normal epithelial cells.

A novel method for detecting CTCs ,using a green fluorescent protein (GFP) - expressing conditionally replicating adenovirus(Ad) (Fuminory Sakurai et al. 2015), is developed by Oncolys Biopharma. Telomescan or OBP-401, is an oncolytic adenovirus vector that expresses a GFP gene. OBP-401 contains the human telomerase reverse transcriptase gene promoter. Human telomerase is a complex of template RNA and enzyme subunits including hTERT, which is expressed in cancer and some normal stem cells but not in normal somatic cells.

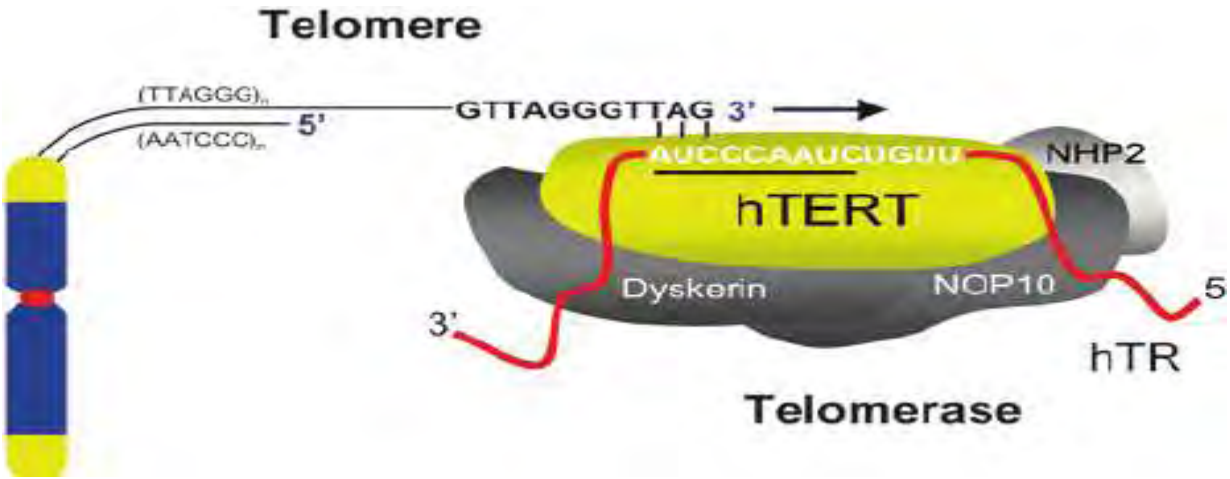

Figure 11: *Structure of telomerase complex.* The human telomerase complex consists of a catalytic sub- unit (hTERT), an RNA component (hTR), dyskerin, NOP2, NHP2, and additional associated proteins. Taken from Roth et al. 2010.

The hTERT gene promoter is contained upstream of the E1 gene in the adenovirus type 5 genome. The expression is highly specific to cancer cells (Takakura et al.1998) and the hTERT promoter has cancer specificity, the virus can express E1 genes preferentially in cancer cells and replicate with much higher efficiency than in healthy cells (Kawashima et al.2004; Takakura et al. 2010).

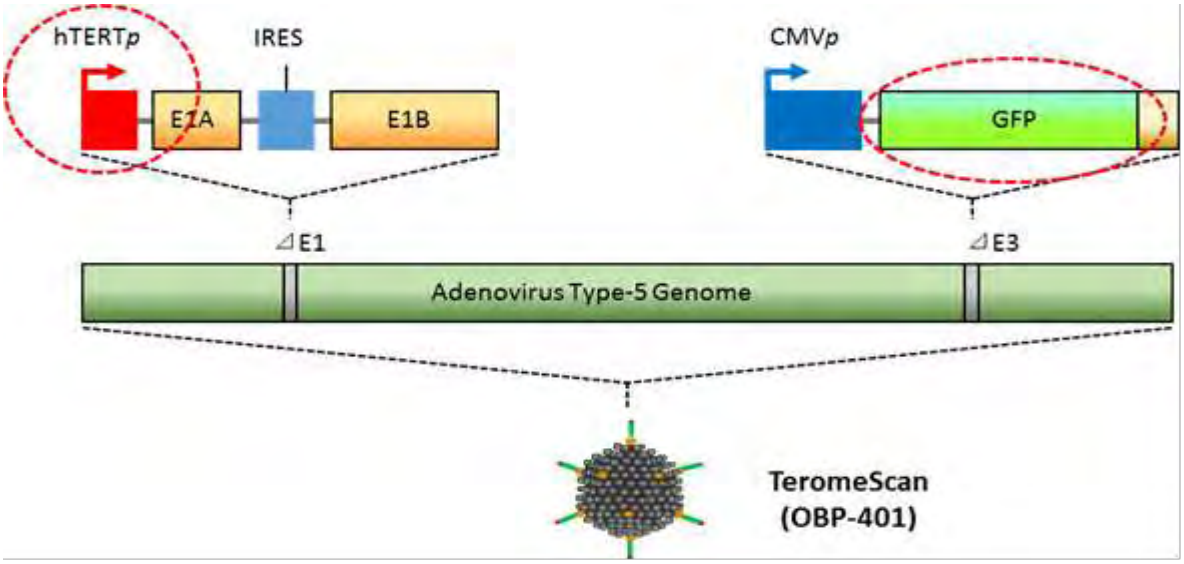

***Figure 12**: Structure of TelomeScan. The hTERT gene promoter has been inserted upstream of the E1A and E1B genes linked with an internal ribosome entry site in the adenovirus type 5 genome, and the GFP gene is inserted under the cytomegalovirus promoter into the E3 region. Taken from Urata et al. 2013.*

Adenovirus enters target cells by binding to the Coxsackie/Adenovirus Receptor (CAR) (Bergelson et al. 1997). After binding to the CAR, the adenovirus is internalized via integrin-mediated endocytosis (Russell, 2000) followed by active transport to the nucleus. Once in the nucleus, the early events are initiated (e.g. transcription and translation of E1 proteins), followed by expression of the adenoviral late genes and viral replication. That expression of the late genes is dependent upon E1.

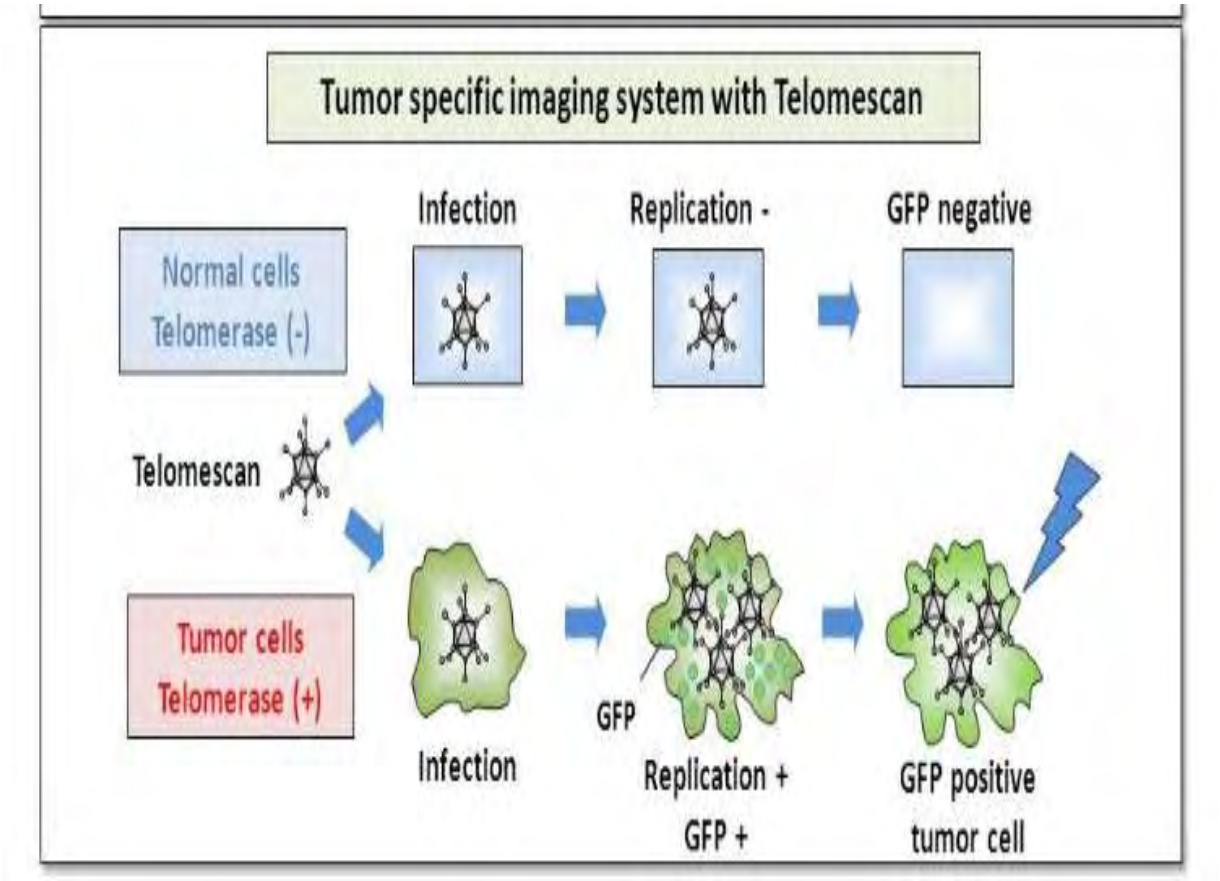

Figure 13:.*Cancer cells are essentially acquired immortalized nature and express strong telomerase activity. Thus, OBP-401 can replicate only in telomerase activity-positive cells such as cancer cells by hTERT promoter activity-dependent E1/E2 gene expression, but not in telomerase activity-negative normal cells. Along with the increase in virus replication, concurrently-expressed GFP accumulates in hTERT-positive cells. Taken from Kishimoto H et al.2006.*

5. Material and methods.

Several apparatuses have been developed for the detection of CTCs including iCellate Medical AB, which offers a service to identify circulating tumor cells from blood samples. In our lab, we perform CTC analyses with our proprietary isolation system IsoPic™ and related downstream verification tests. Most of these assays rely on surface markers which may be susceptible to downregulation, such as that associated with epithelial-mesenchymal transition (EMT). In order to evaluate the efficacy of the detection of CTCs through IsoPic™, we used a novel CTC detection method which is based in one of the hallmarks of cancer (limitless replicate potential). Telomerase is an enzyme that replenishes the ends of chromosomes and it is not expressed in the majority of the normal cells. In contrast, telomerase is upregulated in almost all tumor cells. Oncolys Biopharma, created a telomerase-specific replication selective-adenovirus, which is an oncolytic adenovirus vector that expresses a green fluorescent protein. In this experiment, we evaluated the efficacy of this adenovirus vector to infect breast cancer cell lines (MCF7 and MDA-MB-231) as well as their GFP expression ratios.

5.1 Cell culture.

Two breast cell lines (MCF-7 and MDA-MB-231) were grown in plates with RPMI-1640 medium with 10% (v/v) FBS and 1% (v/v) PS. Cells were harvested at an 80% confluent state using 1x trypsin (Sigma) and resuspended in RPMI-1640 medium. The cells were transferred in a 24 well plate and incubated for 24h.

5.2 Infection without Lipofectamine.

The infection of the cell lines was performed as follows. Firstly, we used one vial of the OBP-401 that is stored in -80 ^{o}C. Materials used in this experiment are:
Virus vial (10~12 μL) OBP-401,
Opti-MEM media,
24-well plate with MCF7 and MDA-MB-231 cell lines.

The experiment was performed in a fully equipped virus laboratory (biosafety level 2). First the virus for the infection was prepared by adding opti-MEM with a final concentration of

100x (12 μL of virus +1.188 μL of Opti-MEM for both rows of the cells). Then the first row of the MCF7 cell line was infected apart from the first well which was used as a control. The first well was used as a control, the second well was infected with an amount of the OBP-401/ Opti-MEM solution of 6 μL for a MOI of 10. The third well with 26 μL for a MOI OF 40. The fourth well with 65.2 μL for a MOI of 100. The fifth well with 130 μL for a MOI of 100. The sixth with 326 μL for a MOI of 500. The procedure was replicated for the first row of the MDA-MB-231 cell line.

MCF7 CONTROL	6μL(virus solution)+ 0.5mL Opti-MEM	26μL(virus solution)+ 0.5mL Opti-MEM	65.2μL(virus solution)+ 0.5mL Opti-MEM	130μL(virus solution)+ 0.5ml Opti-MEM	326μL(virus solution)+ 0.5mL Opti-MEM
MCF7	MCF7	MCF7	MCF7	MCF7	MCF7
MDA CONTROL	6μL(virus solution)+ 0.5mL Opti-MEM	26μL(virus solution)+ 0.5mL Opti-MEM	65.2μL(virus solution)+ 0.5mL Opti-MEM	130μL(virus solution)+ 0.5mL Opti-MEM	326μL(virus solution)+ 0.5mL Opti-MEM
MDA	MDA	MDA	MDA	MDA	MDA

After this step, we added, in each well of the rows that we infected, 0.5 mL of Opti-MEM and we put the 24 well plate for a 12 hour incubation (required time for the infection of the cell lines). After 12h the virus solution was removed and RPMI-1640 medium was added for another 12h. After 24h of infection, the cells expressing GFP were counted under a fluorescent microscope.

5.3 Infection with Lipofectamine.

To compare the efficacy of the infection, the same experiment was performed using Lipofectamine® 2000 Transfection Reagent. The protocol used had one more step in comparison with the infection protocol without the transfection reagent. In addition with the virus solution, a Lipofectamine solution with Opti-MEM was prepared. In total a dilution of 15 μL of Lipofectamine® in 2.947 mL of Opti-MEM was made. After the Lipofectamine solution was added into the virus solution was incubated for 20 min. After 12h of incubation, the virus

solution from the wells was replaced by an RPMI- 1640 medium. After 24h, the GFP expression rates were observed in a fluorescent microscope.

5.4 Virus vector OBP-401.

For this experiment, the amount of virus needed for different MOI (10, 40,100,200) was calculated. For a 24 well plate the cell number is approximately $1.5*10^5$ cells in a single well.

MOI	Virus needed[IU]
10	$1.5*10^6$
40	$6*10^6$
100	$15*10^6$
200	30*106
500	$75*10^6$

To transform the amount of virus needed into μL, we used this general formula: **μL needed=virus needed[IU]/viral titer[IU/μL]**. The virus titer is $2.3*10^7$ IU/μL.

MOI	Adenovirus amount [μL]
10	0.065
40	0.26
100	0.652
200	1.3
500	3.26

Each virus vial contains 10 μL of viral particles. This was diluted 100x, so 6 μL from the vial was used while 594 μL of Opti-MEM was added to total 600 μL. The final concentrations for different MOI are summarized in the table below.

MOI	Adenovirus amount[μL]
10	6
40	26
100	65.2
200	130
500	326

5.5 Cell counting.

To determine cell viability, a TC20^{TM} automated cell counter by BIORAD was used. 1 part trypan blue dye was mixed with 1 part of cell suspension. On parafilm 10μL of the cell suspension was combined with 10μL of trypan blue dye and pipetted up and down to mix. 10 μL of the mixture was pipetted into the opening of either chamber on the counting slide. Then the counting slide was inserted into the slide slot of the TC20 cell counter, which automatically initiated a cell count.

6. Results.

6.1 Titration of OBP-401 virus vector in cancer cell lines without Lipofectamine.

For the first phase of the experiment, a decision was made to infect the cell lines without using a transfection reagent. Before infecting the cell lines the numbers of virus particles needed to infect one cell (MOI) were calculated. The MOI differs between various cell lines. Therefore, because the cell lines were infected for the first time, this approach of infecting without Lipofectamine, was used to determine the MOI necessary for efficient transgene expression in MCF7 and MDA-MB-231 cell lines.

One day before the infection, the MCF7 and MDA-MB-231 cells were seeded into a 24 well plate (first two rows for MCF7 and last two for MDA-MB-231). The following day, the virus particles were thaw on ice. For easier handling the virus was diluted in OPTIMEM and pipetted in higher quantities into the wells. To establish a control, one well was left free of virus particles. The infected cells were incubated for 12h.

After 12h the virus solution was removed from the wells, and RPMI-1640 medium was added. The cells were left for another 12h of incubation at 37^0C before the acquisition of pictures of the cells by fluorescent microscopy. Another set of cell pictures were acquired after 48h of incubation. (**Figure 14**).

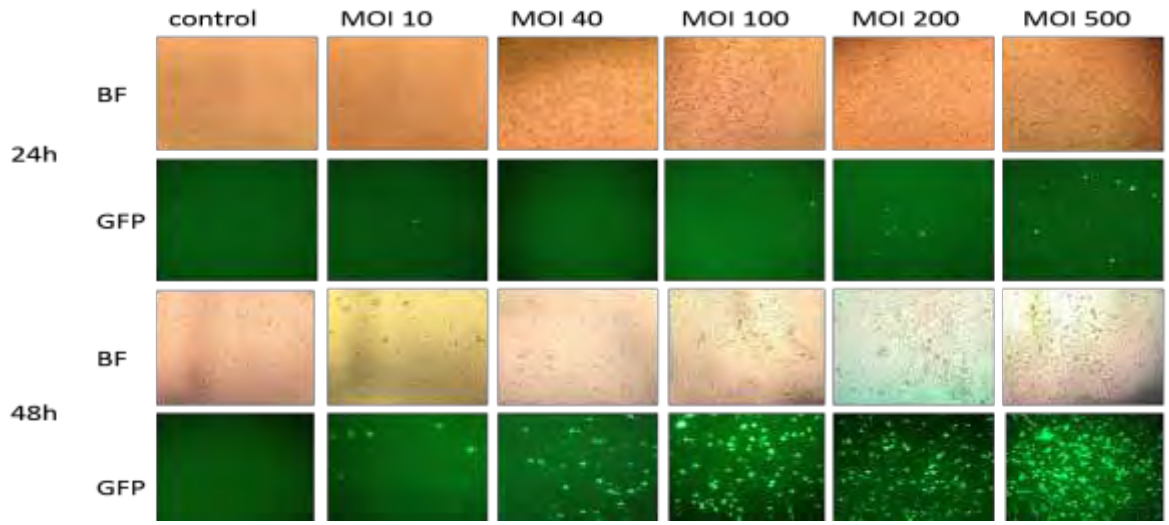

Figure 14*. MDA cell line after infection with the virus vector OBP-401 without using Lipofectamine. For MCF7 data are not shown.*

6.2 Titration of OBP-401 virus in cancer cell lines using Lipofectamine.

In order to determine the efficiency of infection with OBP-401 in MCF7 and MDA-MB-231 cell lines, the cell lines were infected, using this time Lipofectamine as a transfection reagent. Cells were seeded into a 24 well plate and left for a 24h incubation. The virus particle was thaw on ice and the virus solution was prepared (12 μL of virus + 1.188 μL of Opti-MEM). The virus solution was divided in 5 tubes, (based on the amounts for different MOI). Then, another 5 tubes were prepared for the Lipofectamine solution. The calculated amounts for Opti-MEM was added in each tube followed by the addition of 3 μL of Lipofectamine in all of them. Tubes were incubated for 5 minutes, then transferred the Lipofectamine solution from each tube to the one with the virus solution. All were further incubated at room temperature for 20 minutes before infection and further incubation for 12h. Then, the virus-Lipofectamine solution was replaced by RPMI-1640 medium were left for another 12h incubation. Pictures were acquired after 24h and 48h of incubation respectively, using a fluorescent microscope (**Figure 15, 16**).

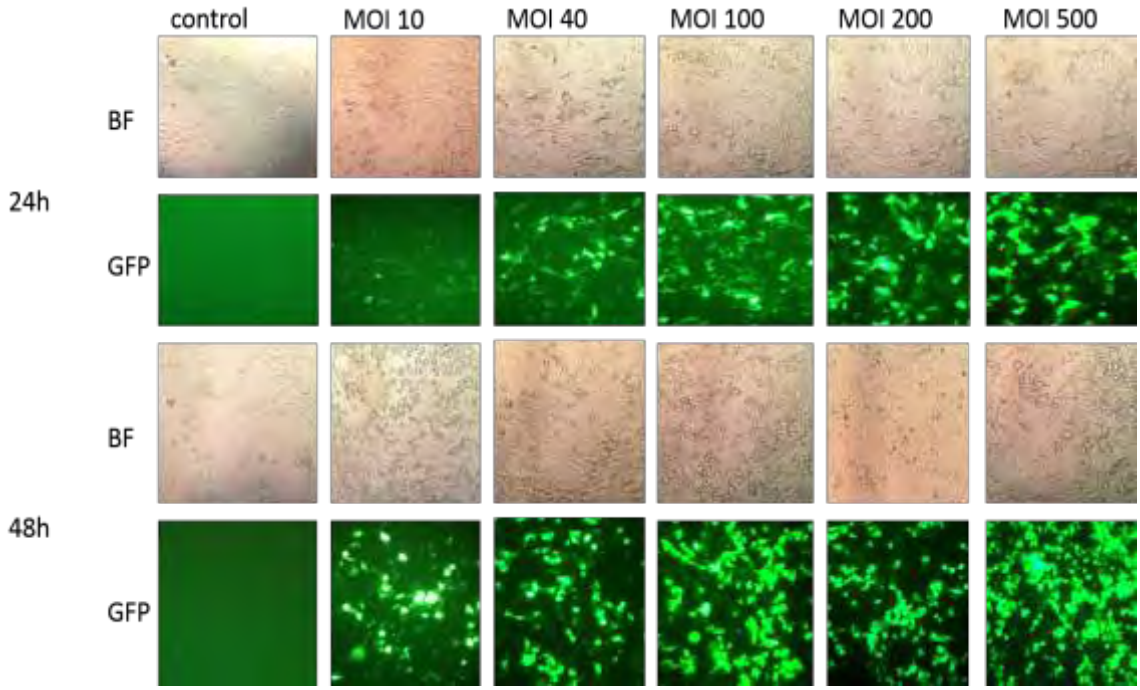

Figure 15*. MCF7 cell line infection results using Lipofectamine. GFP expression ratios were observed using an inverted fluorescent microscope after 24h and 48h.*

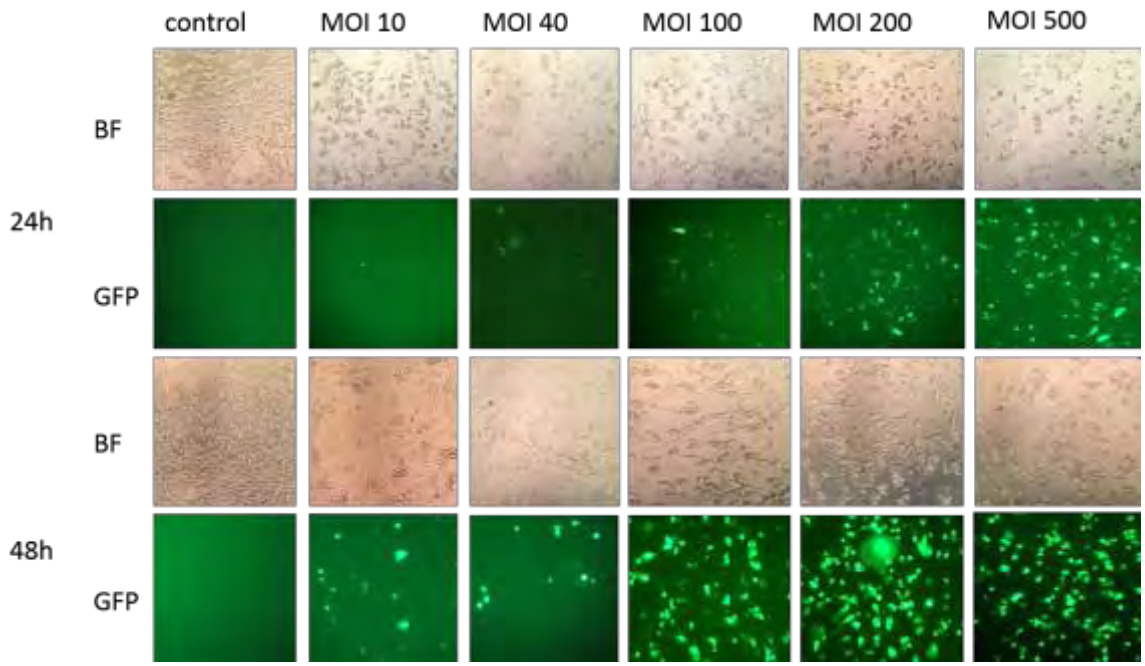

***Figure 16**: MDA-MB-231 cell line infection results using Lipofectamine. GFP expression ratios were observed using an inverted fluorescent microscope after 24h and 48h respectively.*

6.3 Cell counting.

Despite the successful infection of both cell lines with and without a transfection reagent, the rate of the infected cells per well for each MOI and at each time point was necessary to be determined. To evaluate the efficiency of the infection, pictures of the cells were acquired after 72h of incubation. Higher quantities of the virus after 72h (MOI 200,500) lead to cytotoxic effects (**Figure 17**). In order to determine the ratio between live and dead cells into each well, a TC20^{TM} automated cell counter by BIORAD was used. First, the RPMI-1640 from each well was removed and cells were treated with trypsin and then were counted. The results of the cell counting are presented in **Figures 18 and 19**.

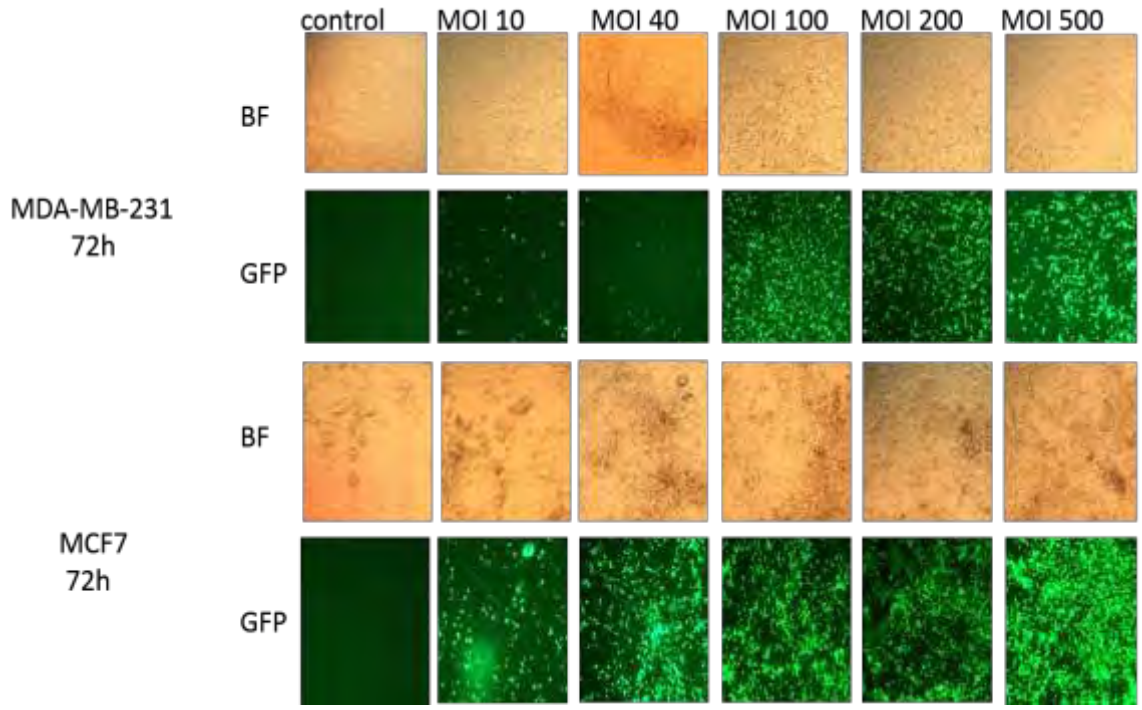

***Figure 17**: MCF7 and MDA-MB-231 cell lines infection results using Lipofectamine. GFP expression ratios were observed using an inverted fluorescent microscope after 72h.*

MCF7	Total count	Live count
CONTROL	$7{,}55 \times 10^4$ cells/mL	$5{,}53 \times 10^4$ cells/mL
MOI 10	$1{,}31 \times 10^5$ cells/mL	$8{,}55 \times 10^4$ cells/mL
MOI 40	$6{,}54 \times 10^4$ cells/mL	$5{,}53 \times 10^4$ cells/mL
MOI 100	$1{,}11 \times 10^5$ cells/mL	$9{,}06 \times 10^4$ cells/mL
MOI 200	Out of range	Out of range
MOI 500	$6{,}54 \times 10^4$ cells/mL	$5{,}03 \times 10^4$ cells/mL

***Figure 18**: Total number of living cells after the infection of MCF7 cells with OBP-401 using Lipofectamine. In the table we can see the number of living cells in comparison with the total amount of cells in each well infected for each MOI after 72h from infection.*

MDA-MB-231	Total count	Live count
CONTROL	$3.62x10^5$ cells/mL	$3.47x10^5$ cells/mL
MOI 10	$2.62x10^5$ cells/mL	$2.52x10^5$ cells/mL
MOI 40	$2.06x10^5$ cells/mL	$2.01x10^5$ cells/mL
MOI 100	$1.91x10^5$ cells/mL	$1.86x10^5$ cells/mL
MOI 200	$1.56X10^5$ cells/mL	$1.36x10^5$ cells/mL
MOI 500	Out of range	Out of range

***Figure 19**: Total number of living cells after the infection of MDA-MB-231 cells with OBP-401 using Lipofectamine. In the table we can see the number of living cells in comparison with the total amount of cells in each well infected for each MOI after 72h from infection.*

In order to determine the rate of infected cells per well for each MOI, the amount of the GFP expressing cells for each cell line and for each MOI after 72h from infection with Lipofectamine were calculated. **Figures 20 and 21**, illustrate the results from testing different MOI on the MCF7 and MDA-MB-231 respectively.

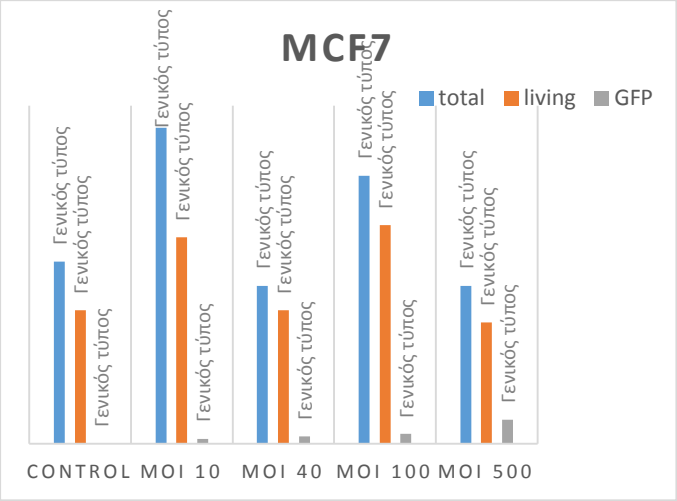

***Figure 20**: Total number of cells in each well in comparison with the number of living cells and GFP positive cells for MCF7 cell lines. The counting of the GFP positive cells was performed manually.*

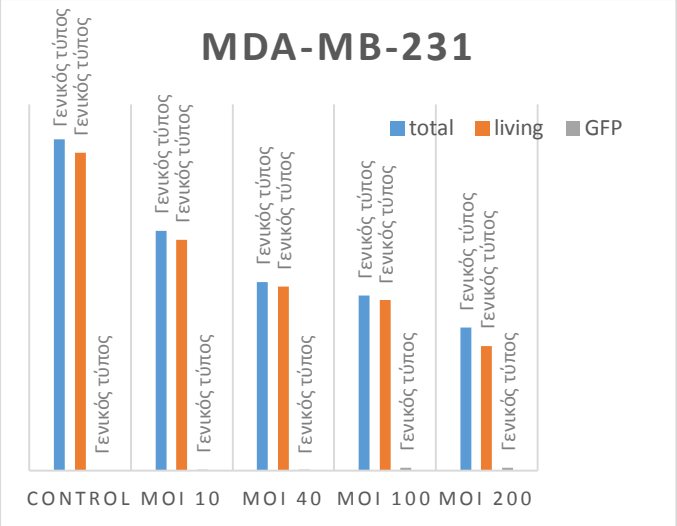

***Figure 21**: Total number of cells in each well in comparison with the number of living cells and GFP positive cells for MDA-MB-231 cell lines. The counting of the GFP positive cells was performed manually.*

6.4 Infection of MCF7 and MDA-MB-231 cells lines with OBP-401 using Lipofectamine in MOI 40 and 100.

The first two experiments were performed to evaluate the efficiency of infection of MCF7 and MDA-MB-231 cell lines with the virus vector OBP-401. It was determined that infection using Lipofectamine as a transfection reagent was more successful in comparison with the infection with a transfection reagent. Furthermore, the MOI necessary for efficient GFP expression without cytotoxic side effects was determined. High quantities of the virus (MOI 200.500) led to artifacts due to non-healthy cells. As a result, the infection of the cell lines was performed with the adjusted MOI 40 and 100.

This time, the cells were seeded and transferred in a 6 well plate. In the first column the MCF7 cells were transferred while in the second column the MDA-MB-231 cells. The plate was incubated for 24h. The following day, the virus solution (2 μL virus particles + 198 μL Opti-MEM) was prepared and the Lipofectamine solution divided in two tubes (one for MOI 40 and one for MOI 100). The tubes were left for a 5 minutes incubation and then the Lipofectamine solution from each tube was transferred to each tube with the virus solution, respectively. After a 20 minute incubation at room temperature cells in each well were infected and incubated for 12h. Then the virus-Lipofectamine solution was replaced by RPMI-1640 medium and another 12h incubation was followed. Pictures were acquired after 24h (**Figure 22**).

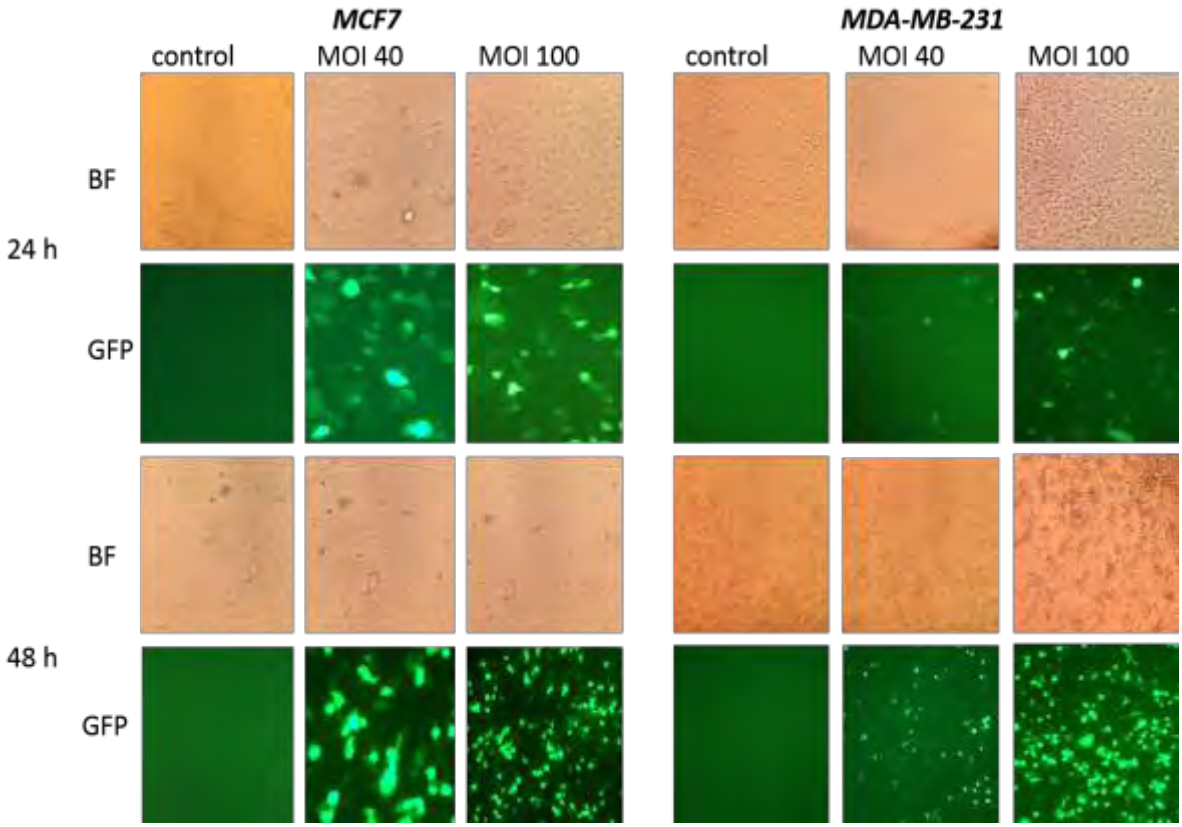

***Figure 22**: MCF7 and MDA-MB-231 cell line infection results for MOI 40 and 100 using Lipofectamine. GFP expression ratios were observed using an inverted fluorescent microscope after 24h and 48h respectively.*

6.5 Identification of GFP positive cells.

To evaluate the GFP expression ratios, in the blood samples that were spiked with the OBP-401 virus vector infected MCF7 and MDA-MB-231 cell lines, a fluorescent microscope was used. The microscopy results are shown in **Figures 23 and 24**.

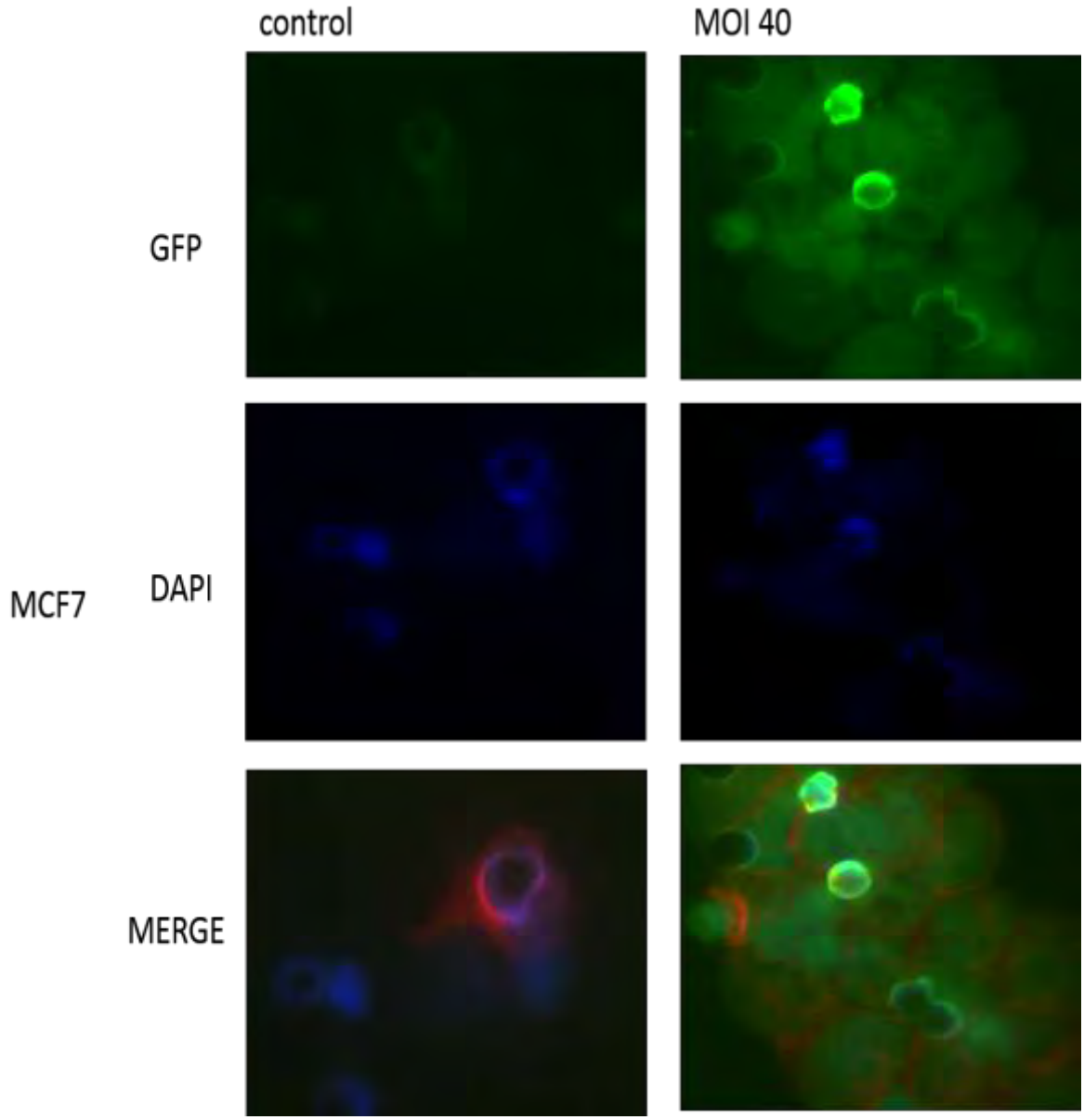

***Figure 23**: Fluorescent microscopy results of blood samples spiked with OBP-401 virus infected MCF7 cell lines. The blood samples were spiked with uninfected MCF7 cells as a control and with MOI 40 virus infected MCF7 cells. In the first row, the GFP expression results are shown. In the second and third row, the DAPI and the merge pictures are displayed.*

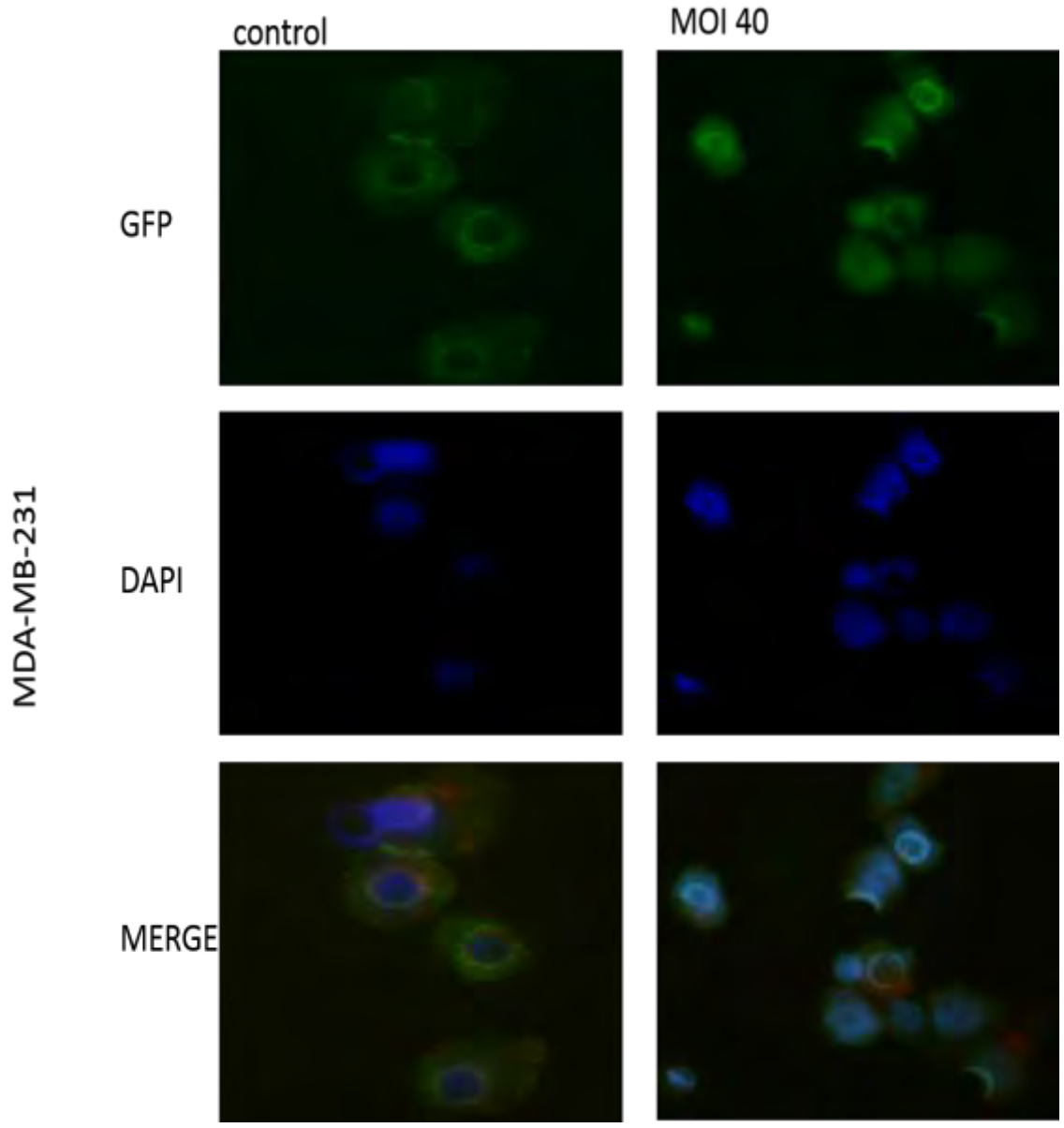

***Figure 24**. Fluorescent microscopy results of blood samples spiked with OBP-401 virus infected MDA-MB-231 cell lines. The blood samples were spiked with uninfected MDA-MB-231 cells as a control and with MOI 40 virus infected MDA-MB-231 cells. In the first row, the GFP expression results are shown. In the second and third row, the DAPI and the merge pictures are displayed.*

7. Discussion.

Current available CTC detection methods are based on immunological recognition of EpCAM, cytokeratin and other epithelial specific cell surface markers. Although the detection with these methods is successful, the rates of detection vary between 14,4 to 90% (He et al. 2008; Judson et al, 2003; Behbakht et al,2011; Poveda et al. 2011). During the epithelial-to-mesenchymal transition (EMT) process, the epithelial markers are known to be downregulated, leading to false negative results (Mikolajczyk et al. 2011). Furthermore, the detected cells using these methods are not always cancer cells. There is a chance that these cells are normal epithelia found into the blood circulation. Lastly, these epithelial markers are also expressed in dead cells (present surface antigens), so the detected cells may not be living cells. OBP-401 was chosen due to his ability to infect and replicate only in cancer living cells. Also the rate of infection is not affected by the EMT process.

OBP-401 virus vector was chosen to evaluate the efficient detection of cancer cells via GFP expression. The MCF7 and MDA-MB-231 breast cancer cell lines were used. MDA-MB-231 cells are very prone to cytotoxic agents because of their lack of DNA repairing capability due to mutations of BCRA or TP53 (Velic et al. 2015). From the infection of the cell lines without Lipofectamine, the successful infection of the cells was observed. After 48h from infection and in MOI 40,100,200 and 500, the expression rates increased periodically.

In higher MOI (200,500), the number of dead cells was higher in comparison with the wells infected with lower MOI. After the observation that cells can be infected successfully without a transfection reagent, the same experiment was performed using Lipofectamine. From the microscopy observations, the difference of the GFP expression ratios between the cell lines was evident. The intensity and the number of the GFP expressed cells are higher in MCF7 cells in comparison with MDA-MB-231 cells.

By comparing the infection results with and without Lipofectamine of the MDA-MB-231 cells, the number of the GFP positive cells was higher while the fluorescence was also more intense using Lipofectamine. In conclusion, it can be estimated that infection with Lipofectamine is more successful than the infection without it. Furthermore, it was discovered that higher quantities of the virus led to cytotoxic effects so it was decided to perform another infection with Lipofectamine with MOI 40 and 100. After the infection, the detection efficiency of the OBP-401 detection system was tested in blood samples. After harvesting the GFP positive MCF7 and MDA-MB-231 cells were spiked into blood samples.

From microscopy results, the CTC detection rate was approximately 45% successful in these clinical samples. The low CTC detection rate, shows the improvements that are needed in current detection systems.

CTC detection could be a really valuable prognostic marker in combination with genetic sequencing. CTC's can be sequenced and give us a panorama of the mutations in each patient's cancer. Blood samples are relatively easily available compared to tissue biopsies and can be taken repeatedly. This can lead to a perpetual monitoring of treatments. As a result, the overgrowth of drug resistant sub clones may be prevented by shifting to alternative therapies. There is a hope that the short list of treatment targets can be extended to include ultimately all, of the several hundred driver genes in cancer.

8. References.

1). http://cancer.sanger.ac.uk/cancergenome/projects/cencus/

2). http://www.intogen.org

3). http://ccgd-starrlab

4). http://www.fda.gov

5). www.mycancergenome.org

6). National cancer institute www.cancer.gov

7). Futreal P.A., Coin L., Marshall M., Down T., Hubbard T., Wooster R., Rahman N., Stratton M.R. A census of human cancer genes. Nat. Rev. Cancer. 2004; 4: 177–183

8) Gonzales-Perez A., Lopez-Bigas N. Functional impact bias reveals cancer drivers. Nucleic Acids Res. 2012; 40: e169.

9). Forbes SA,Beare D,Gunasekaran P,Leung K, Bindal N, Boutselakis H, Ding M, Bamford S, Cole C, Ward S, Kok CY, Jia M, De T, Teague JW, Stratton MR, McDermott U, Campbell PJ. COSMIC: exploring the world's knowledge of somatic mutations in human cancer.Nucleic Acids Res. 2015 Jan; 43(Database issue):D805-11.

10). 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(Database issue):D652-7

11). Bamford S, Dawson E, Forbes S, Clements J, Pettett R, Dogan A, Flanagan A, Teague J, Futreal PA, Stratton MR, Wooster R. The COSMIC (Catalogue of Somatic Mutations in Cancer) database and website.Br J Cancer. 2004 Jul 19; 91(2):355-8

12). Mularoni L, Sabarinathan R, Deu-Pons J, Gonzalez-Perez A, López-Bigas N. OncodriveFML: a general framework to identify coding and non-coding regions with cancer driver mutations.Genome Biol. 2016 Jun 16;17(1):128. doi: 10.1186/s13059-016-0994-0.

13) Tamborero D, Gonzalez-Perez A, Lopez-Bigas N. OncodriveCLUST: exploiting the positional clustering of somatic mutations to identify cancer genes.
Bioinformatics. 2013 Sep 15; 29(18):2238-44. doi: 10.1093/bioinformatics/btt395. Epub 2013 Jul 24.

14) Abel Gonzalez-Perez, Christian Perez-Llamas,Jordi Deu-Pons,DavidTamborero,Michael P Schroeder, Alba Jene-Sanz,Alberto Santos & Nuria Lopez-Bigas. IntOGen-mutations identifies cancer drivers across tumor types. Nature Methods 10, 1081–1082 (2013) doi:10.1038/nmeth.2642

15) Vogelstein B1, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer genome landscapes. Science. 2013 Mar 29; 339(6127):1546-58. doi: 10.1126/science.1235122.

16) Hanahan D1, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011 Mar 4; 144(5):646-74. doi: 10.1016/j.cell.2011.02.013.
17) DouglasHanahan,Robert AWeinberg. The Hallmarks of Cancer. Cell.Volume 100, Issue 1, 7 January 2000, Pages 57-70.
18) https://www.tempus.com
19) www.icellate.se
20) Bin Hong and Youli Zu, Detecting Circulating Tumor Cells: Current Challengesand New Trends. Theranostics .2013 3(6): 377–394.
doi: 10.7150/thno.5195
21) M Takakura, S Kyo,, M Nakamura, Maida, Mizumoto, Bono, X Zhang, Hashimoto, Urata, T Fujiwara, and M Inoue. Circulating tumour cells detected by a novel adenovirus-mediated system may be a potent therapeutic marker in gynaecological cancers. Br J Cancer. 2012 Jul 24; 107(3): 448–454. doi: 10.1038/bjc.2012.27.
22) Kim SJ1, Masago A, Tamaki Y, Akazawa K, Tsukamoto F, Sato J, Ozawa T, Tsujino Y, Noguchi S. A novel approach using telomerase-specific replication-selective adenovirus for detection of circulating tumor cells in breast cancer patients. Breast Cancer Res Treat. 2011 Aug; 128(3):765-73. doi: 10.1007/s10549-011-1603-2. Epub 2011 Jun 1.
23) Sakurai F1, Narii N2, Tomita K2, Togo S3, Takahashi K3, Machitani M2, Tachibana M2, Ouchi M4, Katagiri N4, Urata Y4, Fujiwara T5, Mizuguchi H6. Efficient detection of human circulating tumor cells without significant production of false-positive cells by a novel conditionally replicating adenovirus. Mol Ther Methods Clin Dev. 2016 Mar 2; 3: 16001. doi: 10.1038/mtm.2016.1. eCollection 2016.
24)www.oncolys.com
25) Rubio-Perez C1,Tamborero D1, Schroeder MP1, Antolín AA2, Deu-Pons J1, Perez-Llamas C1, Mestres J2, Gonzalez-Perez A1, Lopez-Bigas N3. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. Cancer Cell. 2015 Mar 9; 27(3):382-96. doi: 10.1016/j.ccell.2015.02.007.
26) Diaz-Cano. Tumor heterogeneity: mechanisms and bases for a reliable application of molecular marker design. Int J Mol Sci. 2012; 13(2):1951-2011. doi: 10.3390/ijms13021951. Epub 2012 Feb 13.
27) Stratton MR1, Campbell PJ, Futreal PA. The cancer genome. Nature. 2009 Apr 9; 458(7239):719-24. doi: 10.1038/nature07943.
28) Marchetti A1, Felicioni L, Malatesta S, Grazia Sciarrotta M, Guetti L, Chella A, Viola P, Pullara C, Mucilli F, Buttitta. Clinical features and outcome of patients with non-small-cell lung cancer harboring BRAF mutations. J Clin Oncol. 2011 Sep 10; 29(26):3574-9. doi: 10.1200/JCO.2011.35.9638. Epub 2011 Aug 8.

29) Thiery JP1, Sleeman JP. Complex networks orchestrate epithelial-mesenchymal transitions. Nat Rev Mol Cell Biol. 2006 Feb; 7(2):131-42.
30) Christiansen JJ1, Rajasekaran AK. Reassessing epithelial to mesenchymal transition as a prerequisite for carcinoma invasion and metastasis. Cancer Res. 2006 Sep 1; 66(17):8319-26.
31) Mikolajczyk SD1, Millar LS, Tsinberg P, Coutts SM, Zomorrodi M, Pham T, Bischoff FZ, Pircher TJ. Detection of EpCAM-Negative and Cytokeratin-Negative Circulating Tumor Cells in Peripheral Blood. J Oncol. 2011; 2011: 252361. doi: 10.1155/2011/252361. Epub 2011 Apr 19.
32) Alunni-Fabbroni, Sandri. Circulating tumor cells in clinical practice: Methods of detection and possible characterization. Methods. 2010 Apr; 50(4):289-97. doi: 10.1016/j.ymeth.2010.01.027. Epub 2010 Jan 29
33) Paterlini-Brechot P1, Benali NL. Circulating tumor cells (CTC) detection: clinical impact and future directions. Cancer Lett. 2007 Aug 18; 253(2):180-204. Epub 2007 Feb 20.
34) Riethdorf S, Mu¨ller V, Zhang L, et al. Detection and HER2 expression of circulating tumor cells: prospective monitoring in breast cancer patients treated in the neoadjuvant GeparQuattro trial. Clin Cancer Res. 2010; 16: 2634-2645.
35) Allard WJ1, Matera J, Miller MC, Repollet M, Connelly MC, Rao C, Tibbe AG, Uhr JW, Terstappen LW. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clin Cancer Res. 2004 Oct 15; 10(20):6897-904.
36) Bergelson JM, Cunningham JA, Droguett G, Kurt-Jones EA, Krithivas A, Hong JS, Horwitz MS, Crowell RL, Finberg RW. Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. Science. 1997 Feb 28; 275(5304):1320-3.
37) Denis Velic , Anthony M. Couturier, Maria Tedim Ferreira, Amélie Rodrigue, Guy G. Poirier , Fabrice Fleury and Jean-Yves Masson. DNA Damage Signalling and Repair Inhibitors: The Long-Sought-After Achilles' heel of Cancer. Biomolecules 2015, 5(4), 3204-3259; doi: 10. 3390/biom5043204.
38) He W, Kularatne SA, Kalli KR, Prendergast FG, Amato RJ, Klee GG, Hartmann LC, Low PS. Quantitation of circulating tumor cells in blood samples from ovarian and prostate cancer patients using tumor-specific fluorescent ligands. Int J Cancer. 2008; 123(8:1968–1973
39) Poveda A, Kaye SB, McCormack R, Wang S, Parekh T, Ricci D, Lebedinsky CA, Tercero JC, Zintl P, Monk BJ. Circulating tumor cells predict progression free survival and overall survival in patients with relapsed/recurrent advanced ovarian cancer. Gynecol Oncol. 2011; 122 3:567–572.
40) Behbakht K, Sill MW, Darcy KM, Rubin SC, Mannel RS, Waggoner S, Schilder RJ, Cai KQ, Godwin AK, Alpaugh RK. Phase II trial of the mTOR inhibitor, temsirolimus and

evaluation of circulating tumor cells and tumor biomarkers in persistent and recurrent epithelial ovarian and primary peritoneal malignancies: a Gynecologic Oncology Group study. Gynecol Oncol. 2011; 123 1:19–26.