THE UNIVERSITY
of EDINBURGH

This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.

Identification of Targets of Human Cytomegalovirus microRNAs by Cross Linking and Analysis of cDNA (CRAC).

CHIWESHE Stephen Masaka

s1161257

College of Medicine and Veterinary Medicine
Royal (Dick) School of Veterinary Studies
The Roslin Institute
University of Edinburgh

Thesis submitted for the fulfilment of the requirements of a Master of Science by Research Degree 2012

Supervisor: Dr Finn Grey

Funded by the BBSRC and the Wellcome Trust

Declaration

I declare that this thesis has been composed by myself, Stephen Masaka Chiweshe, and the work presented is my own original work unless where otherwise indicated and all work of other authors is duly acknowledged. This work does not form part of a thesis presented successfully for a degree in this or another University.

Stephen Masaka Chiweshe

Acknowledgements

I would like to express my sincere and greatest appreciation and gratitude to Dr Finn Grey for giving me the opportunity to work on this project and the extensive immeasurable support he gave me through the year. I would like to thank Dr Natalie Reynolds for the help, support and tips in the laboratory as well as proof reading my thesis. Mr Jonathan Andrew Pavelin for, I don't know - being a member of the group perhaps, mainly preliminary knowledge and skills with the CRAC that was transferred productively as well as support here and there and "a wee bit" of entertainment in the lab. Many thanks to the Dr Amy Buck and group for very useful CRAC tips and emergency reagents supply when we most needed them. A huge thank you to ARK genomics for our cDNA sequencing (at a cost, thanks to Dr Grey and funding for payment) and the on-going results analysis. I would to say thank you to all the other groups in the lab for the help they gave during the year, discussions and cake breaks now and again. Last but not least, a massive shout out to friends and family, with a special mention to Mrs Sue Dixon and Mrs Karen Binch, for the encouragement and mainly for leaving me be when I wanted to be left be, also for the support in different forms and kinds, much appreciated.

List of Abbreviations

2ME	2-Mercaptoethanol
Ago-2	Argonaute 2
α	Alpha
β	Beta
CCMV	Chimpanzee cytomegalovirus
cDNA	Complementary deoxyribonucleic acid
cm	centimetre
CRAC	Crosslinking analysis of cDNA
CTL	Cytotoxic T lymphocyte
DC	Dendritic cell
DNA	Deoxyribonucleic acid
dsDNA	Double Stranded DNA
DMEM	Dulbecco's Modified Eagle Medium
E	Early
EBV	Epstein-Barr virus
FACS	Fluorescence-activated cell sorting
FBS	Foetal Bovine Serum
GFP	Green Fluorescence Protein
gN*	Glycoprotein (N*= B/L/M/N/O)
HCMV	Human cytomegalovirus
HSC	Haematopoietic stem cell
HEK	Human embryonic kidney
HIV	Human immunodeficiency virus
HITS-CLIP	High-Throughput Sequencing of RNAs isolated by Crosslinking Immunoprecipitation
HV	Herpes virus
HSV1	Herpes Simplex Virus type 1
hpi	hours post infection
hrs	Hours
IgG	Immunoglobulin G
IL	Interleukin
IE	Immediate early
IP	Immunoprecipitation
IRS	Internal repeat short
kb	kilobases
kDa	kilo Daltons

KSHV	Kaposi Sarcoma Herpes Virus
L	Late
MHC	Major Histocompatibility Complex
MIE	Immediate early
miRNA	Micro ribonucleic acid
mJ	millijoule
N/A	Not Applicable
*ori*Lyt	Origin of Lytic Replication
µg	microgram
µl	microlitre
ml	millilitres
mRNA	Messenger ribonucleic acid
MT	Microtubule
nm	nanometres
NKG2D	Natural Killer cell protein Group 2D
ORF	Open reading frame
PBM	Peripheral blood mononuclear
Pri-miRNA	Primary miRNA
Pre-miRNA	Precursor miRNA
PAR-CLIP	Photo*a*ctivatable-*R*ibonuclease enhanced Crosslinking and Immuno*p*recipitation
Pol II	Polymerase II
PRG1	Plasticity related gene 1
RISC	RNA induced silencing complex
THBS1	Thrombospondin 1
PCR	Polymerase Chain Reaction
qRT PCR	Real Time quantitative Reverse Transcription PCR
TRL	Terminal repeat long
TRS	Terminal repeat short
UTR	Untranslated region
UL	Unique long
US	Unique short

List of Figures

List of Tables

Abstract

The discovery of a class of small ribonucleic acid (RNA) molecules known as microRNAs (miRNAs) has led to extensive interest in their biological relevance and role. The first miRNA was discovered in *Caenorhabditis elegans* in 1993. Since then studies have shown that miRNAs represent a fundamental mechanism of gene expression regulation, regulating thousands of genes at the post-transcriptional level. Given that viruses are highly adept at exploiting cellular processes, it is perhaps unsurprising they have evolved miRNAs of their own. The majority of known viral miRNAs are expressed by herpes viruses underscoring their importance to this virus family. Identifying the targets of herpes virus miRNAs would aid in elucidating the role played during infection. In this study, we aim to identify and understand the targets of the human cytomegalovirus (HCMV) encoded miRNAs using a cutting edge biochemical technique, Cross-Linking and Analysis of cDNA (CRAC).

HCMV is a member of the beta (β) herpes virus subfamily and is globally distributed causing clinically asymptotic infections in immune competent individuals. However, persistent and recurrent infections in AIDS and organ transplant patients, who have a compromised immune function causes a high degree of mortality and morbidity. Intra-uterine infected infants are also at high risk with infection causing congenital abnormalities and mental retardation. Scientists worldwide are trying to understand one special characteristic of HCMV and herpes viruses in general, which is latency i.e. the presence of an intact viral genome in the cell with a majority of the genes in a dormant or silent state. Viral encoded miRNAs regulate both the viral genome as well as the host's and this has been postulated as a latency inducing mechanism.

To date, there are 22 known HCMV encoded miRNAs. On-going research in our group using techniques such as bioinformatics, RISC immunoprecipitation (RISC-IP) and microarray analysis has identified both viral and cellular targets of HCMV miRNAs. Targets include a crucial viral transactivator thought to play an important role in latency as well as cellular targets involved in a variety of functions including cell cycle control, intrinsic defence and innate cellular defence. However, the majority of HCMV miRNAs have no known function.

New approaches and technologies are required to elucidate the functions of these miRNAs. The initial goal of this project is to establish and optimise the CRAC technique. This will be performed in HEK293 cells transfected with a plasmid expressing a cluster of HCMV miRNAs termed miR-US25. The long-term goal of the project will be to use CRAC technology to identify miRNA targets in infected cells.

Key words: HCMV, miRNA, Latency, CRAC.

Table of Contents

Chapter - 2 Materials and Methods

Chapter 3 - Results

CHAPTER 1- Introduction

1.1 Herpes Viruses

The herpes viruses are a family of viruses that widely infects animals and humans. They are also referred to as *Herpesviradae*, a Greek derived word with a meaning inclining towards "latent" which is a characteristic of members of this family, i.e. latent infections. Members of this viral family have a general structure, illustrated in figure 1.1, made up of the core, capsid, tegument and envelope as the main components and these are key structures that are a prerequisite for a virus to be classified as a herpes virus.

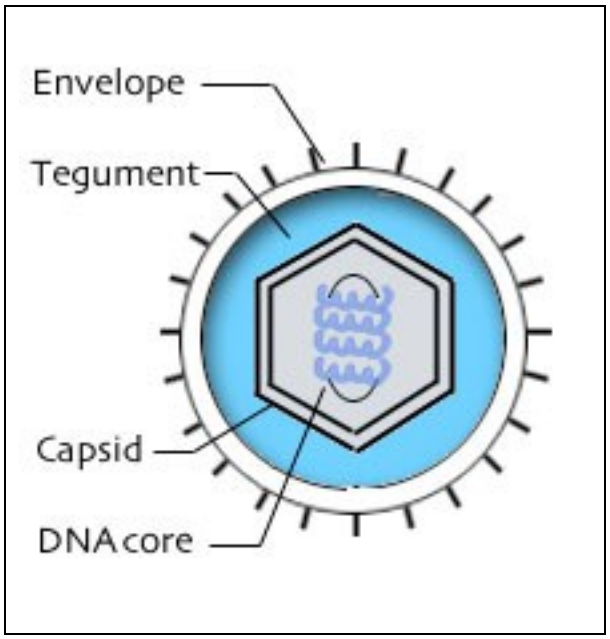


Fig 1.1: Structure of Herpes virus (Grundhoff and Sullivan, 2011).

The core contains the viral DNA in a ring-shaped form known as the torus. The capsid is made up of 162 subunits known as capsomeres and has a diameter of approximately 100nm. The tegument is the component between the capsid and the envelope, and exists in different thicknesses depending on the specific virus as well as the location of the virus within an infected cell. The tegument has the important role of transporting pre-synthesised proteins into the newly infected cells which have an immediate effect on host management, establishing a virus-favourable environment through activities such as shutting down the host protein synthesis, inhibiting infection-triggered cell defences, and stimulating viral gene

expression (Fields et al., 2007). The envelope forms the outer structure of the virus and is composed of lipids, glycoproteins and altered forms of the host membrane.

Herpes viruses have a large double stranded DNA genome which varies from 124 to 239 kb in length and is linear but is known to circularise immediately upon release from the capsid into the infected cells' nuclei. There are eight known human-infecting herpes viruses and they are classified into three subfamilies, *Alphaherpesviranae* (α), *Betaherpesviranae* (β) and *Gammaherpesviranae* (γ) (table 1.1).

Table 1.1: Herpesviruses infecting Humans.

Virus Name	Abbreviation	Vernacular Name	Subfamily	Genome size
Human HV 1	HHV-1	Herpes simplex virus 1	α	152
Human HV 2	HHV-2	Herpes simplex virus 2	α	155
Human HV 3	HHV-3	Varicella-zoster virus	α	125
Human HV 4	HHV-4	Epstein-Barr virus	γ	172
Human HV 5	HHV-5	Cytomegalovirus	β	230
Human HV 6	HHV-6A	HHV-6 variant A	β	159
	HHV-6B	HHV-6 variant B	β	162
Human HV 7	HHV-7		β	145
Human HV 8	HHV-8	Kaposi's sarcoma-associated Herpesvirus	γ	170

HV: Herpes virus. Table adapted from Fields Virology (Fields et al., 2007).

In addition to the structural features, the human herpes viruses have other common characteristics, including the conservation of what is known as the core genes across the α, β and γ subfamilies. These are 40 genes that encode for mainly structural, nucleotide metabolism or DNA replication proteins (Fields et al., 2007). Herpes viruses have evolved means and ways to enable their replication and survival in the highly non-favourable

environment the host presents. One such mechanism is by blocking the induction of apoptosis or the interferon pathway (Fields et al., 2007, Miller et al., 2000). Herpes viruses are also known to employ a process known as molecular mimicry whereby some viral encoded proteins function as immunomodulators. The virus can also block the presentation of antigenic peptides on the surface of antigen presenting cells. These processes in turn result in a delay of the infected cells' elimination, long enough to allow the replication and infection of new cells by the virus as well as the viral evasion of the immune system and persistence during latency period.

1.2 Biology of Human Cytomegalovirus

Human cytomegalovirus (HCMV) is a member of the β herpes virus subfamily and specifically infects humans. HCMV infection in immunocompetent individuals is asymptomatic but is a cause of significant morbidity and mortality in immunocompromised people such as, organ transplant recipients, intrauterine infected infants as well as AIDS patients where it causes opportunistic infections.

1.2.1 Virion Structure

HCMV, like other herpes viruses, have the general structure illustrated in figure 1.1. The virion has a diameter of approximately 200 to 300nm, one of the largest and structurally more complex in the *Herpesviridae* family. The HCMV genome is encompassed by the capsid and together the two components make an icosahedral nucleocapsid with a diameter of approximately 130nm. The tegument surrounds the nucleocapsid and comprises viral proteins, termed tegument proteins, as well as a selection of viral and cellular RNA (Fields et al., 2007). A selection of tegument proteins and the roles they play in HCMV infection are discussed in the lytic replication section below. On the outer layer of the virion is the irregular envelope, a lipid bilayer, which is partly made from altered host cellular membranes and has a vast number of glycoproteins that play a critical role in viral host cell entry.

1.2.2 Genome Structure

HCMV has the largest genome among the herpes viruses, which is in a linear double stranded form, of approximately 230 kb. The genome is made up of unique long (UL) and unique short (US) regions that are flanked by repeat regions as illustrated below.

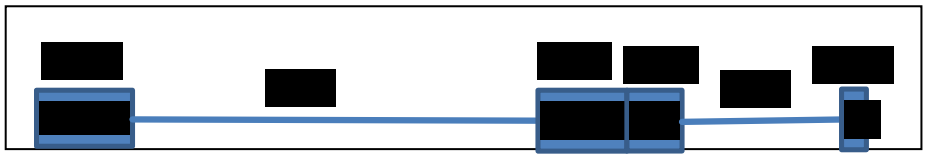


Fig 1.2: HCMV genomic arrangement. TRL: Terminal Repeat Long, IRL: Internal Repeat Long, IRS: Internal repeat Short, TRS: Terminal Repeat Short, UL: Unique Long and US: Unique Short (redrawn from (Fields et al., 2007)).

Recombination during replication leading to the inversion of these genome components generates four different genome isomers in equal amounts, which are all infectious. Work by Chee *et al* (1990) predicted that the HCMV genome encodes for 189 protein coding genes. This figure has since been reduced to a range of 164 to 167 after work comparing the closely related chimpanzee CMV (CCMV) sequence to the laboratory strains AD169 and Toledo genome sequences found that some of the open reading frames (ORFs) were not protein-coding genes (Davison et al., 2003). The exact number of ORFs is still difficult to determine as laboratory strains are obtained after a number of passages and they tend to undergo some genetic loss, mutations and rearrangements as they are passaged.

1.2.3 Lytic Replication

The lytic life cycle of HCMV follows a sequence of controlled events that commence immediately after binding of the virus to host cell receptors. By a mechanism that involves glycoproteins, particularly glycoprotein B (gB), one of the core proteins conserved in all herpes viruses, HCMV attaches to receptors on the cell surface and initiates its entry. Heparin sulphate, which is expressed on the host cell surface, is thought to act as a receptor for gB molecules. HCMV is understood to also use other envelope proteins such as glycoprotein H, L, M and O (gH, gL, gM and gO) for host cell binding. The two most reported mechanisms of entry are fusion between the viral envelope and the plasma membrane for entry into fibroblast cells, and receptor mediated, pH dependent endocytosis for entry into epithelial and endothelial cells (Bodaghi et al., 1999, Ryckman et al., 2008, Fields et al., 2007). For entry into fibroblast cells, the glycoproteins gH, gL and gO have been demonstrated to interact and form a complex, gH/gL/gO, that interacts with host cells surface's integrins leading to the direct fusion at the plasma membrane of the virion and a target host cell. The three proteins, UL128, UL130 and UL131, encoded by the UL128-131 region form pentameric complexes with gH/gL and mediate HCMV entry into epithelial and endothelial cells also via integrin binding (Ryckman et al., 2008, Park et al., 2006). Upon entry of the virus into the host cell, the nucleocapsid is released into the cytoplasm and translocated to the nucleus by a cytoskeletal dependent mechanism and interacts with a nuclear pore allowing the release of the genome into the nucleus.

Lytic replication of the HCMV genome starts with the expression of the viral immediate early (IE) genes as soon as the viral DNA reaches the host cell's nucleus. This is followed by the expression of the early (E) genes, viral DNA replication and then late (L) gene expression. Encapsidation of the viral genome followed by release of an infectious virus concludes the lytic life cycle (Johnson et al., 2001). Efficient and successful delivery of the viral genome to the host's nucleus is facilitated by important tegument proteins which also initiate viral gene expression (Kalejta, 2008). Upon entry of the HCMV nucleocapsid into the cell, the first barrier to reaching the cell's nucleus is the high cytoplasm density as well as the size of the nucleocapsid. To circumvent this, the virus uses the intracellular transport machinery, travelling along microtubules (MTs) (Ogawa-Goto et al., 2003) with the aid of some tightly associated tegument proteins such as the UL32 encoded pp150, as well as the UL47/UL48 protein complex (Bechtel and Shenk, 2002, Kalejta, 2008). The UL47/UL48 protein complex also plays a role in the injection of the viral DNA into the cell's nucleus via the nuclear pore complex. Work performed on herpes simplex virus type 1 (HSV1) found similar complexes formed by the pUL37 and pUL36, also called VP1/2 respectively that migrate along the host cells' MTs (Klupp et al., 2002). These complexes were further associated with nucleocapsids, which were found accumulating together at the nuclear surface (Kalejta, 2008).

pp71, a component of the tegument, initiates viral IE gene expression through interaction with proteins that bind the nuclear components known as the promyelocytic leukemia (PML) nuclear bodies. The PML nuclear bodies' mode of action is not fully understood and it is not clear whether they are pro- or anti-viral. It is however known that the viral genome binds to a subset of these PML nuclear proteins on successful entry into the nucleus (Stinski and Petrik, 2008) and in turn interacts with two other proteins, Daxx and Sp100 (Tavalai et al., 2006). The Daxx protein causes repression of viral gene expression and interaction of pp71 with Daxx counteracts this repression leading to stimulation of IE expression (Cantrell and Bresnahan, 2005).

The major IE genes (MIE), IE86 and IE72, play pivotal roles in driving acute replication of HCMV. Although not fully understood, the MIE proteins function as transactivators of early and late viral gene expression which occurs within the 6 to 24 hours post infection (hpi) window (Fields et al., 2007) and is then followed by viral DNA replication.

HCMV DNA synthesis initiates at the *cis*-acting origin of lytic replication (*ori*Lyt), a sequence that is more than 2.4 kbp in length and is located within the UL genomic region. *ori*Lyt is a bidirectional promoter with a complex structure consisting of repeat elements as well as transcription factor binding sites. The essential regions are the IE2-UL84 responsive promoter region and an RNA/DNA hybrid structure which is a substrate of the UL84 protein (Prichard et al., 1998, Pari, 2008, Anders et al., 1992). IE2 binds to the *ori*Lyt promoter activation element, a 14-bp DNA motif within the *ori*Lyt, and, in cooperation with UL84, activates this promoter. Through a mechanism not fully understood, this binding facilitates or triggers initiation of DNA synthesis via transcriptional activation (Xu et al., 2004). Other elements important for DNA replication include the DNA polymerase (UL54), primase and primase-associated factor (UL70 and UL102), helicase (UL105), p52/DNA processivity factor (UL44) and the single-stranded DNA-binding protein (UL157). These elements form the core replication machinery and are conserved in other herpes viruses. Four proteins encoded by the UL112-113 region via alternative splicing act as transcriptional enhancers and recruit the DNA processivity factor (UL44) to the pre-replication foci. The helicase-primase complex formed by UL105-UL102-UL70 unwinds the dsDNA which is then prevented from reforming into a double strand by the binding of the single stranded DNA binding protein (UL57). DNA polymerase then binds and replicates the DNA with UL44 maintaining its binding to the DNA template (Pari, 2008, Fields et al., 2007, Sarisky and Hayward, 1996, Park et al., 2006).

The circularised HCMV genome is thought to replicate by a process that involves the formation of concatemeric molecules which are in turn cleaved into unit length genomes. Initially, genome inversion and maturation occurs in the nucleus with the inversion leading to the formation of the different genome isomers. The mature DNA is then packaged into pre-formed B capsids with some tegument proteins such as UL56, UL89 and UL104 playing a role in the packing. DNA cleavage and packaging is directed by the highly conserved short sequences found near the genomic termini in all mammalian HVs known as the *pac*1 and *pac*2. Once packaged, the mature virions are then transported in vesicles via the Golgi apparatus to the cell surface until the successful egress from the cells through the exocytic pathway, a process that takes approximately two days from DNA synthesis (Fields et al., 2007, Pari, 2008, Xu et al., 2004, Sarisky and Hayward, 1996).

1.3 Pathogenesis of HCMV

HCMV infection has a widespread distribution with a variable seroprevalence of between 50% and 90% depending on the socio-economic state of a population (Sinclair and Sissons, 2006). HCMV establishes a systemic infection with major cells infected including fibroblast, epithelial and endothelial cells, but leukocytes, dendritic cells, monocytes/macrophages, brain and retinal neurones, gastrointestinal smooth muscle cells and hepatocytes, are also susceptible to infection (Sinzger and Jahn, 1996, Rossini et al., 2012). Individuals infected have a lifelong infection with episodes of reactivation.

HCMV is transmitted through infected bodily fluids such as saliva, urine as well as cervical and seminal excretions during sexual intercourse. The virus can also be transmitted through the placenta from infected pregnant women to their unborn babies with breast milk also reported as a source of transmission. Recipients of blood transfusion, solid organ or hematopoietic stem cell transplantation (HCT) from infected donors also risk HCMV infection (Fields et al., 2007, Dworsky et al., 1983, Gandhi and Khanna, 2004). Primary infection in immunocompetent individuals does not often manifest into clinically obvious disease, mainly as a result of the broad, strong and durable immune response detailed in the following section.

1.3.1 Immune Response to HCMV

HCMV infection triggers a coordinated innate and adaptive immune response. Microarray studies by Boehme *et al*. (2004) demonstrated that cells treated with a recombinant form of the envelope protein gB and with HCMV displayed a profile with a high degree of similarity with both treatments resulting in beta interferon secretion from fibroblast cells. This led to the conclusion that envelope glycoproteins represent at least one mechanism by which HCMV can elicit interferon responses and established that an antiviral state begins before the cell is infected. But interferon secretion is just one way in which the innate immune system responds. A viral attack also triggers an increase in the production of chemokines, cytokines and natural killer (NK) cells (Boehme et al., 2004).

T-lymphocytes and antigen-presenting cells cooperate and provide an adaptive immune response in an effort to counter the viral infection. The T helper (CD4+) cells recognise and interact with antigens presented on the cell surface by the major histocompatibility complex (MHC) class II molecules on the surfaces on infected cells eliciting their metabolic transformation which, in turn, release various cytokines. This leads to the activation of cytotoxic T lymphocyte (CTL) (CD8+) cells and macrophage cells. CTLs target infected cells and release cytotoxins that lead to cell death, preventing the spread of the virus. Similarly, macrophages prevent the spread of infection by targeting and phagocytosing infected cells. The CTLs can also recognise HCMV infected cells by directly interacting with the antigen presenting MHC class 1 molecules on infected cells leading to cytotoxin release (Fields et al., 2007, Kindt et al., 2007). Unlike T-lymphocytes, NK cells lack the antigen-specific receptors but are able to employ NK-cell receptors that have the ability to distinguish abnormalities such as MHC class I expression or different profile of surface antigens displayed by the virus infected cells. Stimulation of NK cells can also be achieved by direct activation of natural killer cell protein group 2 (NKG2D) and Ly49H receptors by cytokines such as interleukins (IL) 12 and 18 as well as INFs (Kindt et al., 2007, Arase and Lanier, 2004). NK cells present the body with an innate immune response to HCMV infection and their activation results in cytotoxicity or cytokine production, depletion of infected cells and prevention of virus spread. In addition to this, there are some viral proteins, IE and E, that induce apoptosis of the host cell. For example US28 was shown to induce apoptosis through the activation of caspases 8 and 10 (Pleskoff et al., 2005), ultimately preventing replication and spread of the virus.

HCMV, like all herpes viruses, has evolved mechanisms to counteract these events to promote a persistent and latent infection despite a robust and fully functional immune system.

1.3.2 Latency

Latency, as a characteristic of all herpes viruses, is defined by a reversibly quiescent state in which viral genomes are maintained, but viral gene expression is highly restricted and no virus is produced (Goodrum *et al.*, 2012). Viruses are however able to reactivate and replicate, mostly when the host's T-cell mediated immunity is weak, for example in AIDS

patients or in solid organ and haematopoietic stem cells (HSC) recipients with a suppressed immunity. Viral replication and reactivation has also been linked to the differentiation and activation state of myeloid cells (Reeves et al., 2005). It is not fully understood how herpes viruses maintain latent and lifelong infection in immunocompetant hosts. Studies on HCMV latency are further complicated by the species specificity of this virus meaning that animal model studies have not been possible to date. Cell culture studies have however given some insight on the subject.

Trying to understand mechanisms underlying the latency of HCMV has led to different lines of enquiry such as site(s) of latency, the viral genes involved in the maintenance of latency and successful evasion of the host's immune system. Taylor-Wiedeman and colleagues studied peripheral blood mononuclear (PBM) cells from healthy seronegative and seropositive individuals. They used PCR to study PBM cells that were highly purified by fluorescence-activated cell sorting (FACS) and detected the presence of HCMV DNA primarily in monocytes (Taylorwiedeman *et al.*, 1991).

The presence of HCMV genome in the bone marrow CD34+ progenitor cells in the absence of a lytic gene expression led to the understanding that these cells are a possible site of HCMV latency. Work by Reeves and colleagues also demonstrated, *in vitro,* that myeloid dendritic progenitors are a site for HCMV latency and that their differentiation results in the reactivation of the virus' lytic replication. They isolated and purified monocytes and CD34+ cells from HCMV seropositive and seronegative donors and analysed for the presence of the viral genome. Their results showed the presence of the viral genome in both the seropositive acquired monocytes and CD34+ cells, and that their differentiation to mature dendritic cells (DCs) led to the induction of lytic replication and an increase in the viral genome copy number (Reeves et al., 2005, Mendelson et al., 1996). Healthy seropositive blood donors can transmit HCMV infection to seronegative recipients and this transmission can be reduced by using leukocyte-depleted blood products (Sinclair and Sissons, 2006), supporting the understanding that HSCs harbour the latent HCMV genome and constantly shed the viruses as they differentiate into leukocytes.

In vitro studies have also identified a number of HCMV genes potentially associated with viral latency. These include the UL111A, UL81-82 antisense transcript (UL81-82ast), US28 and miRNAs.

The UL111A encodes for the cmvIL-10 protein, a homologue of human interleukin 10 (IL10). IL-10 is a cytokine that inhibits immune responses and the ability of HCMV to express a homologue of this cytokine suggest one of the ways the virus manages to avoid host immunosurveillance as it maintains latency in the host. cmvIL-10 has been reported to suppress the production of pro-inflammatory cytokines as well as the expression of the MHC II expression in CD34+ HPCs promoting viral latency (Goodrum et al., 2012, Sinclair and Sissons, 2006).

The UL81-82 locus encodes for a 133 amino acid protein (16KD) that has been termed latent undefined nuclear antigen (LUNA). LUNA is understood to restrict the expression of the tegument protein pp71, a product of the viral gene UL82. pp71 is a transcriptional activator of the MIE promoter and suppression of pp71 by LUNA effectively inhibits the lytic replication of the virus, therefore promoting HCMV latency (Bego et al., 2005). Various studies have also indicated the promotion of latency by some viral miRNA as detailed in section 1.4.3.2.

1.4 miRNA Studies & Identification

1.4.1 General miRNAs Introduction

MicroRNAs are small, single stranded RNA molecules of approximately 20 to 24 nucleotides in length. These non-coding RNAs are involved in the regulation of gene expression in eukaryotes thorough a post-transcriptional mechanism using the RNA silencing induced complex (RISC). Ambros and colleagues in 1993 discovered the first miRNA in *Caenorhabditis elegans.* Their studies revealed two *lin-4* transcripts of 22 and 61 nucleotides; the 22 nucleotide transcript being the result of the processing of the 61 nucleotide transcript. These transcripts contained sequences complementary to a repeated sequence element in the 3' UTR of *lin-14* mRNA and it was demonstrated that the *lin-4* regulates *lin-14* via an antisense RNA-RNA interaction leading to a decrease in the translation of *lin-14* mRNA (Lee et al., 1993). This led to the birth of a unique new family of RNAs known as the miRNAs which have now been found to be ubiquitously expressed in metazoans (Grundhoff and Sullivan, 2011). miRNAs have since been established as a subject of enormous interest with their regulatory roles having been detected in various and diverse cellular processes such as immune function, apoptosis, tumorigenesis , growth, proliferation, phenotype, cell cycle as well as death, and in turn having a major influence on pathophysiological outcomes (Bhatt et al., 2011).

1.4.2 miRNA Biogenesis & Function

miRNAs are produced by gene transcription followed by sequential processing or editing. miRNA coding genes form independent transcription units and are found mainly between protein coding regions, intergenic, or in an antisense orientation of annotated genes. The few that are found within genes are located within introns. Most miRNA genes are transcribed by the RNA polymerase II (pol II) with a minority transcribed via RNA pol III (Grundhoff and Sullivan, 2011, Bhatt et al., 2011, Lee et al., 2004).

Transcription of miRNA genes generates large transcripts of several kilobases in length known as primary miRNA (pri-miRNA) molecules which contain well-defined hairpin structures. These hairpin structures are recognised by a nuclear multi-protein complex, the

microprocessor. The microprocessor initiates the first of several steps, in processing of the transcripts, leading to the formation of a mature and functional miRNA molecule of approximately 20 to 24 nucleotides in length (Bhatt et al., 2011, Kim et al., 2009) (figure 1.3).

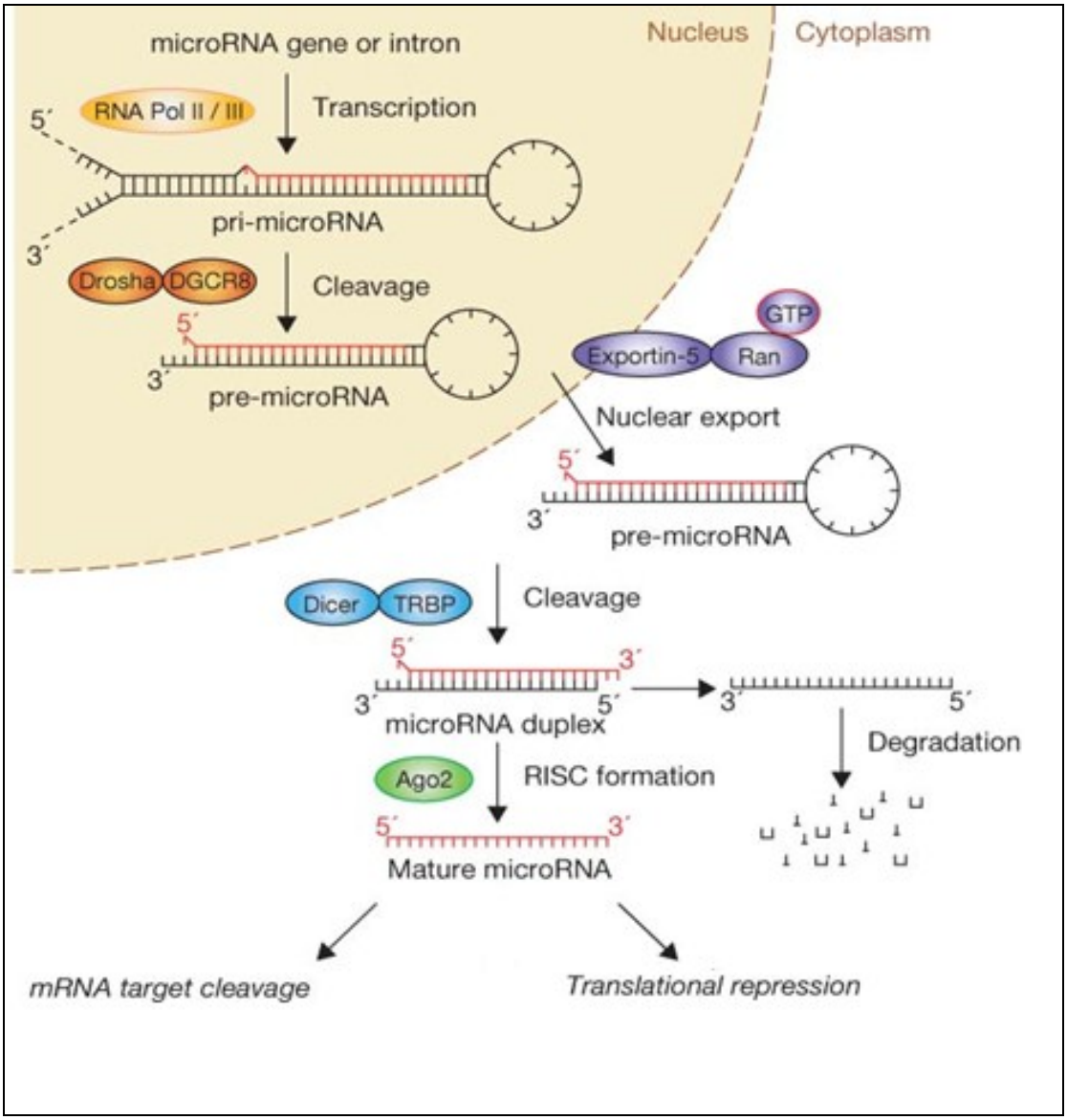


Figure 1.3: miRNA biogenesis (Winter et al., 2009)

The Microprocessor complex consists of two main proteins, an RNase III-like protein known as DROSHA and its co-factor DiGeorge syndrome critical region 8 (DGCR8). The recognition and interaction of the hairpin structure by DGCR8 leads to the recruitment of DROSHA which in turn cleaves the pri-miRNA precisely at the stem loop structure generating a secondary miRNA precursor molecule of approximately 70 nucleotides, known as the precursor miRNA (pre-miRNA) (Bartel, 2004).

The protein exportin-5, which is a nuclear transport receptor family member, recognises the stem-loop structure of the pre-miRNAs and transports them from the nucleus to the cytoplasm where they are further cleaved by an RNase III enzyme known as Dicer, which removes the terminal loop structure generating a mature, double stranded miRNA (Yi et al., 2003). One strand of the duplex is incorporated into the RNA Induced Silencing Complex (RISC), forming a stable interaction with the Argonaute 2 (Ago-2) protein of the RISC. Upon incorporation of the miRNA into RISC, it directs the complex to the target via the seed sequence, which is found between nucleotides 1 to 8 on the 5' end of the miRNA and binds to the 3' UTR of the target gene. Perfect complementarity of the miRNA to its target results in the cleavage of the mRNA whereas a partial match leads to translational repression (Hock and Meister, 2008, Grundhoff and Sullivan, 2011). It must, however, be noted that partial complementarity can lead to degradation of mRNA as well. A single miRNA can bind to and regulate different mRNA target sequences and in the same way, multiple miRNAs can bind and regulate the same target (Lee et al., 2004).

1.4.3 Viral microRNAs

With some of the smallest genomes and miRNA genes taking up relatively little genomic space, miRNAs offers viruses an efficient and convenient way of regulating both their own genomes as well as that of the hosts' to support their life cycle (Dunn et al., 2005, Pfeffer et al., 2004). miRNAs also have the added advantages that a single miRNA can potentially have more than one target and that they do not elicit an immune response from the host (Grundhoff and Sullivan, 2011).

The first viral miRNAs were discovered by Pfeffer and colleagues in 2004, expressed by EBV, a member of the γ herpesvirus sub-family (Grundhoff and Sullivan, 2011, Pfeffer et al., 2004). After cloning small RNAs from a Burkitt's lymphona cell line latently infected with EBV, genomic sequence analysis idenfied 5 miRNAS encoded by EBV (Pfeffer et al., 2004). Further studies have led to the discovery of more viral expressed miRNAs by different viruses. Interestingly, herpes virus encoded miRNAs make up more than 95% of viral miRNAs known to date with 6 of the 8 HHVs having been shown to expresss miRNAs (Grundhoff and Sullivan, 2011, Tuddenham and Pfeffer, 2011, Tuddenham et al., 2012). Enormous interest was raised as it emerged that this class of RNAs play a key role in

enhancing and regulating the herpes virus life cycle, both during lytic replication and latency.

1.4.3.1 miRNAs in the context of HCMV

There are 22 known and annotated HCMV encoded microRNAs (miRNA) and there is still potential for more to be identified. Pfeffer and colleagues first identified HCMV miRNAs in cells undergoing lytic infection. Using small RNA cloning and sequencing they identified nine precursor HCMV miRNAs which gave rise to a total of eleven mature miRNAs (Pfeffer et al., 2005). Three of these miRNAs were also detected by Dunn and colleagues by cloning and sequencing techniques (Dunn et al., 2005).

In 2005 Grey *et al*. used a bioinformatics approach directly comparing HCMV to the closely related Chimpanzee CMV (CCMV) to predict thirteen pre-miRNAs. A bioinformatics program called Stem-loop Finder (SLF; Combimatrix) was initially used to predict potential stem-loop secondary structures forming RNA transcripts from the HCMV genome. The results were further refined using a second algorithm called, MiRscan. MiRscan compares potential candidates from two sequences on the basis of evolutionary conservation of the miRNAs and was used to analyse the HCMV stem-loop transcripts with the CCMV genome for potential homology (Grey et al., 2005). The researchers identified five miRNAs, three of which were previously identified by Pfeffer *et al*. and two novel miRNAs, miR-US4-1 and miR-UL70-1. Using Northern blot analysis, they validated all five miRNAs as well as four previously identified by Pfeffer and colleagues (Grey et al., 2005). The development of high throughput sequencing techniques such as deep sequencing has allowed for the in depth studies of transcript populations. This effectively allows for the yielding a greater representation of a sample or assay as deep sequencing generates reads with a magnitude of millions as opposed to the hundreds that are achieved by small scale sequencing. Stark *et al*. successfully implemented the deep sequencing technique on small RNAs from HCMV lytically infected human fibroblasts cells. They validated ten previously identified pre-miRNAs as well as two novel pre-miRNAs. All but two of these twelve pre-miRNAs resulted in two individual miRNA species leading to the generation of twenty two mature miRNAs (Stark et al., 2012). Their results did not detect the previously annotated pre-mir-UL70 but did reveal the presence of small RNAs outside the annotated HCMV miRNA regions. Further

studies on these regions by bioinformatics using the miResque miRNA prediction algorithm led to the discovery of two novel pre-miRNAs, pre-miR-US22 and pre-miR-US33as, which were validated by Northern blot analysis (Stark et al., 2012). The HCMV encoded miRNAs are summarised in table 1.2. Recently Meshesha *et al*. identified two further pre-miRNAs, pre-miR-UL59 and pre-miR-US29, using deep sequencing. They confirmed both pre-miRNAs by qPCR and proposed that both pre-miRNAs form two mature miRNA. They also detected the formation of one more miRNA, miR-US4-3p, from an already identified pre-miRNA and this potentially brings the total of known HCMV encoded miRNA to twenty seven. Interestingly their results did not reveal any reads for pre-Mir-UL70, agreeing with previous studies by Stark and colleagues.

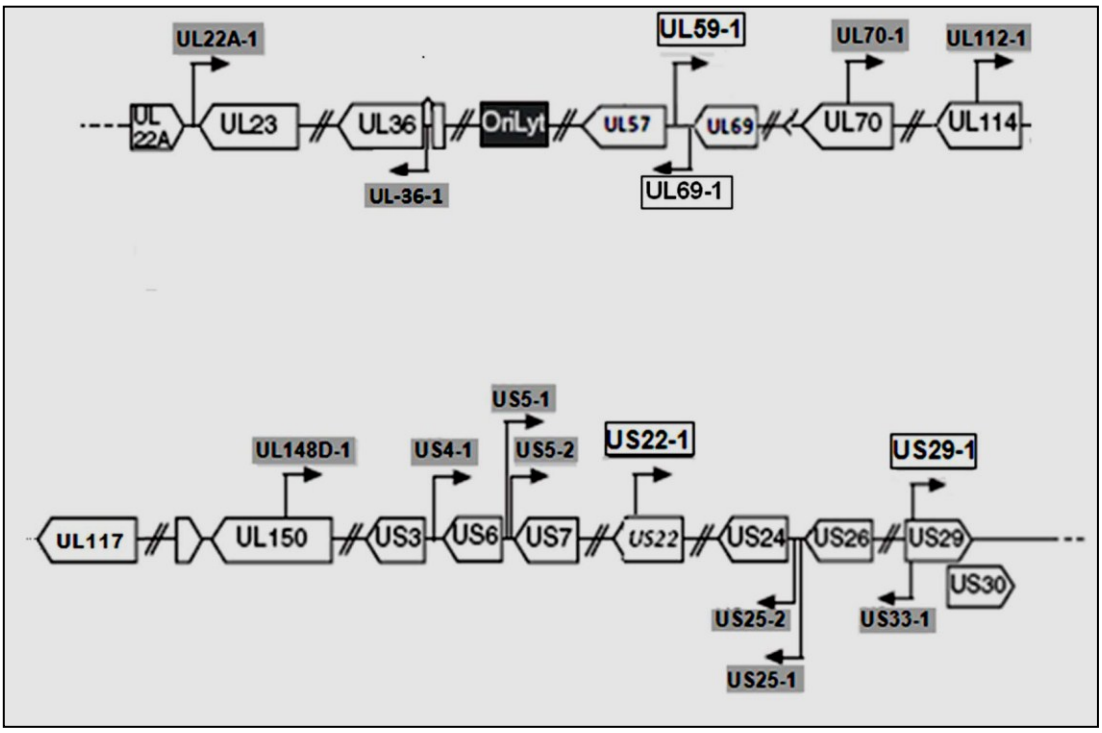


Fig1.4 HCMV miRNA gene map (Meshesha et al., 2012). Also confirmed is the US33as which is not included. Although included, UL70 has been shown to not code miRNAs.

1.4.3.2 HCMV miRNA Potential Targets/Functions - Latency

The three miRNAs, miR-US25-1, miR-US25-2 and miR-UL112-1, have been shown to regulate viral and/or cellular genes in such a way that may promote HCMV latency. Work by Grey *et al*. demonstrated that miR-UL112-1 targets the viral gene IE72 (IE1) that encodes a major *trans*-activating protein. Due to its ability to drive acute replication of HCMV, IE72 along

with IE86, have been suggested to play pivotal roles in establishment and reactivation of the virus. Its inhibition may promote latency (Grey et al., 2007). miR-US25-1 and 2 have been demonstrated by studies from Stern-Ginossar and colleagues to target host cellular transcripts in a mechanism that reduce viral DNA synthesis. By ectopically expressing these miRNAs in HEK293 cells and infecting them with other viruses e.g. HSV 1, they confirmed that these miRNAs were regulating cellular transcript(s) rather than the viral genome as they exhibited the same effect on the growth of these viruses (Stern-Ginossar et al., 2009). On-going work in our lab is trying to establish the role of mir-US25 miRNAs in regulating the cell cycle, particularly cyclin E1 and 2 genes and the resulting effects on latency.

Table 1.2: Known HCMV miRNAs to date.

Annotation and Name	Representative Sequence[1]	Genome Position (5'to 3') [2]	
Annotated in miRBase		From	To
miR-UL112-3p	AAGTGACGGTGAGATCCAGGCT	164557	164578
miR-UL148D	TCGTCCTCCCCTTCTTCACCGT	193587	193607
miR-UL22A-5p	CTAACTAGCCTTCCCGTGAGA	27992	28011
miR-UL22A-3p	TCACCAGAATGCTAGTTTGTAG	28029	28050
miR-UL36-5p	TCGTTGAAGACACCTGGAAAGA	49914	49893
miR-UL36-3p	TTTCCAGGTGTTTTCAACGTG	49870	49851
miR-US25-1-5p	AACCGCTCAGTGGCTCGGACCG	221539	221519
miR-US25-1-3p	GTCCGAACGCTAGGTCGGTTCT	221496	221476
miR-US25-2-3p	ATCCACTTGGAGAGCTCCCGCGGT	221702	221680
miR-US25-2-5p	AGCGGTCTGTTCAGGTGGATGA	221760	221739
miR-US33-3p	TCACGGTCCGAGCACATCCAA	226731	226712
miR-US33-5p	ATTGTGCCCGGACCGTGGGCGC	226768	226750
miR-US4-5p	TGGACGTGCAGGGGGATGTCTG	201376	201395
miR-US5-1	TGACAAGCCTGACGAGAGCGT	202317	202337
miR-US5-2-3p	TTATGATAGGTGTGACGATGTC	202444	202465
Not annotated in miRBase			
miR-UL112-5p	CCTCCGGATCACATGGTTACTCAG	164520	164540
miR-US4-3p	TGACAGCCCGCTACACCTCTCT	201416	201434
miR-US5-2-5p	CTTTCGCCACACCTATCCTGAAAG	202408	202429
miR-US22-5p	TGTTTCAGCGTGTGTCCGCGGGC	216157	216177
miR-US22-3p	TCGCCGGCCGCGCTGTAACCAGG	216195	216216
miR-US33as-5p	TGGATGTGCTCGGACCGTGACG		
miR-US33as-3p	CCCACGGTCCGGGCACAATCAA		

Table details obtained from (Meshesha et al., 2012, Stark et al., 2012). [1]Representative sequence was picked as one with the highest reads in the deep-sequencing. [2]The positions are according to the NCEB database, NC_006273.2 sequence.

1.4.4 Techniques for Identifying miRNA Targets

By identifying viral miRNA targets, we can increase our understanding of their roles and how they influence the viral life cycle. Different tools and techniques have been used to study miRNAs and seeking to improve these will further elucidate miRNA characteristics, targets and function. The techniques previously used, as described below, include bioinformatic studies, microarray analysis, RISC immunoprecipitation, *H*igh-*T*hroughput *S*equencing of RNAs isolated by *C*ross-*L*inking *I*mmuno*p*recipitation (HITS-CLIP) and *P*hoto*a*ctivatable-*R*ibonuclease enhanced *C*ross-*L*inking and *I*mmuno*p*recipitation (PAR-CLIP). We adopted a recently developed technique, *C*ross-*L*inking and *A*nalysis of *c*DNA (CRAC), to study the targets of the HCMV miR-US25 cluster encoded miRNAs.

1.4.4.1 Bioinformatic Studies

Bioinformatics is defined as the application of computational techniques to understand and organise the information associated with biological macromolecules. Bioinformatics is in many ways an ideal approach because of the ease with which computers can handle large quantities of data and probe complex dynamics observed in nature (Luscombe et al., 2001). Bioinformatics equips researchers with a starting point as it identifies potential candidate genes, narrowing down the area and focus point for biochemical studies. A bioinformatics study is mainly a predictive tool with variable limitations depending on the parameters set for the algorithm. There has to be a fine balance in setting up the algorithm parameters in such a way that the stringency levels do not miss out potential targets while at same time minimising the generation of false positives. Further verification steps are taken by the use of wet-bench experimental techniques such as luciferase assays, cloning and sequencing and microarray studies. Bioinformatics remains a strong tool for miRNA studies and can also be used as a validating tool for targets identified by other biochemical techniques.

An example of an algorithm that has been used to identify miRNA targets is RepTar, developed by Stern-Ginossar and colleagues. RepTar was used to identify potential human target genes of miR-UL112-1, a HCMV miRNA. The algorithm searched for repetitive elements in each 3'UTR on the basis that miRNA binding sites can repeat several times in the target's 3' UTR. These studies identified the major histocompatibility complex class 1-

related chain B (MICB) gene as a top candidate target for HCMV-miR-UL112-1 (Stern-Ginossar et al., 2007).

Grey *et al*. (2007) reported the use of a comparative bioinformatics approach to identify viral targets of miR-UL112-1. Using an online target identification algorithm RNAhybrid, they identified 32 potential targets for miR-UL112-1 from 37 HCMV ORFs. However, by comparing these results to the potential targets of the closely related chimpanzee cytomegalovirus (CCMV), 14 genes were predicted to be targeted by miR-UL112-1 in both CCMV and HCMV genomes. Target sites for miR-UL112-1 were confirmed in three viral genes, IE72, UL120/121 and UL112/113 and later validated using luciferase assays and western blot analysis (Grey et al., 2007).

These studies demonstrate that bioinformatic methods can play an important role in target identification and further studies may elucidate future discoveries on viral miRNA function.

1.4.4.2 Microarray Studies

Microarray studies provide a valuable tool for studying and identifying miRNA targets based on gene expression profiling and its ability to simultaneously establish the activity of a large number of genes. Microarrays are used for their ability to identify transcripts subjected to degradation as a result of being targeted by miRNAs of interest (Lim et al., 2005).

In microarray assays, DNA probes on a microchip are used to detect the presence of a transcript by hybridisation, and the amount of mRNA bound to each site on the array indicates the expression level of the coding gene. Typically, the levels of transcripts are compared between a cell line transfected with a miRNA of interest and a control cell line usually transfected with an empty vector. miRNA targeted transcripts will therefore be expected to show a reduction of expression due to cleavage when compared to the negative control sample. This provides an effective way of identifying miRNA targets particularly for plants as they exhibit perfect complementarity with their target, primarily resulting in the cleavage of the transcript whereas with viral and animal miRNAs, the level of degradation is very low as mentioned in section 1.4.2. This makes the detection of transcript level variation difficult with small changes as low as 2- fold having been previously reported (Ziegelbauer et al., 2009).

Lim and colleagues used the microarray technique to demonstrate the tissue specific expression of the human miRNAs miR-1 and miR-124. By transfecting HeLa cells with miR-124 RNA duplexes of the miRNAs, the expression profile shifted towards that of the brain where the miRNA is usually expressed. The same was noted for miR-1 with the expression profile shifting towards that of the heart and skeletal muscle where it is usually expressed. Microarray has also been used in the identification of cellular genes targeted by KSHV-encoded miRNAs (Lim et al., 2005). Samols and colleagues transfected HEK293 cells with a plasmid encoding ten KSHV miRNAs or an empty vector for control samples. They isolated the RNA from the cell lines and generated cDNA for microarray profiling. Their results showed an alteration in the expression of eighty one genes with eight of these having a decrease greater than 4-fold. These were investigated further by qRT-PCR, bioinformatics and luciferase assays with the final results confirming the KSHV miRNA dependent inhibition of the 3' UTRs of the SPP1 (osteopontin), PRG1 (plasticity related gene 1) and THBS1 (thrombospondin 1) genes as well as the targeting of the THBS1 by multiple KSHV miRNAs with the major miRNA species being miR-k12-1, miR-k12-3-3p, miR-k12-6-3p and miR-k12-11 (Samols et al., 2007).

Microarray studies, in combination with deep sequencing and luciferase assay analysis by Suffert *et al*. (2007) led to the identification that three KSHV miRNAs, miR-K12-1, miR-K12-3 and miR-K12-4-3p target caspase 3, an effector caspase involved in apoptosis control. These studies, conducted in B lymphocyte DG-75 and endothelial EA.hy296 cell lines, demonstrated the down-regulation of transcripts with 3' UTRs that possessed KSHV miRNA seed-matches. Deep sequencing of cell lines transfected with KSHV miRNA clusters were used to determine the relative abundance of miRNA within the cluster and determine transcripts with a seed-sequence match within their 3' UTRs. Transcripts showing decreased expression were validated as miRNA targets by luciferase assays (Suffert et al., 2011).

It is fair to say, therefore, that microarray studies provide a solid starting point to identify potential targets which can be investigated further by detailed analyses. However, there is a potential loss of valid targets which would appear as false negatives, for example as result of expression at the RNA level not being altered due to miRNA binding and also those that are affected or down-regulated through secondary effects would score as false positives (Karginov et al., 2007). As a result, microarray usage has been combined with other tests

such as qRT-PCR and luciferase assays to validate the results as well as a validating tool when used in combination with other miRNA studies techniques such as bioinformatics and RISC-IP.

1.4.4.3 RISC Immunoprecipitation

RISC Immunoprecipitation (RISC-IP) is a biochemical technique for isolating miRNA targeted transcripts relying on immunoprecipitation (IP) of the RISC complex. Cell lines stably expressing a tagged component of the RISC complex, Ago-2, allow for the IP of these complexes and the transcripts can then be identified by microarray analysis. The Ago-2 protein is tagged with a c-myc epitope and antibodies conjugated to agarose beads are used to isolate these complexes by IP (Grey et al., 2010, Karginov et al., 2007, Dolken et al., 2010). After transfecting a cell line stably expressing this tagged version of the Ago-2 protein with the miRNA of interest in the form of a plasmid, a sample of the lysate is taken for quantification of RNA levels, which represents the total fraction. IP is then conducted and the RNA from the IP complexes are quantified by microarray or quantitative PCR analysis. On the basis that the association of a specific mRNA with the RISC complex is represented by quantitative enrichment of the mRNA in the IP fraction relative to the total (pre-IP) fraction, the miRNA target is determined from the microarray analysis data (Grey et al., 2010). The cell line can also be infected by a virus and by comparing the IP fraction of uninfected cells to those of infected cells the viral miRNA targets can be identified. The uninfected IP fraction analysis will identify the host's miRNA target and the infected fraction will identify the viral miRNA targets in addition to the host's target. The specific viral miRNAs will therefore have a different enrichment profile which allows their identification. The RISC-IP procedure is schematically represented in figure 1.5.

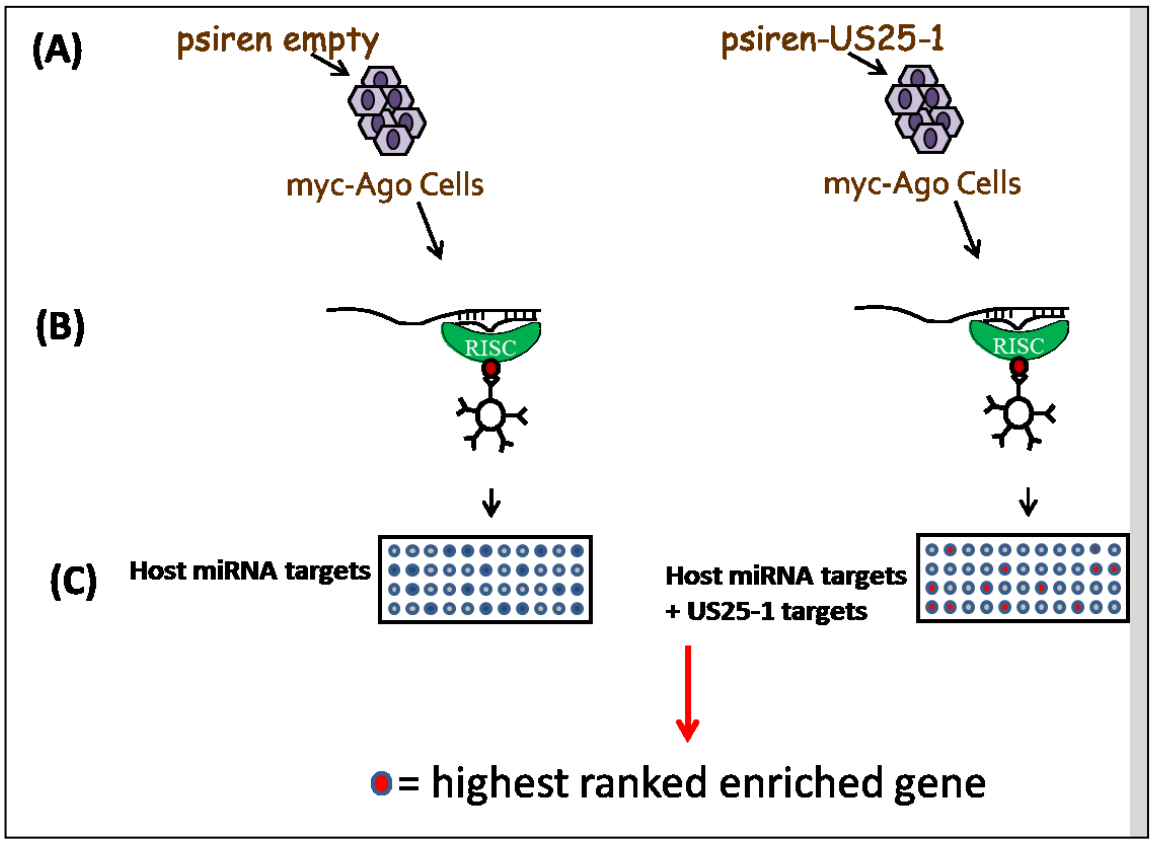


Figure 1.5: A schematic representation of the RISC-IP procedure (Courtesy of Dr Reynolds and Jon Pavelin). (A) miRNAs of interest are transfected into a cell line expressing a tagged Ago-2 protein and an empty vector is used to generate an uninfected IP sample. The miRNAs are incorporated into the RISC, directing mRNA targeting via seed sequence complementarity. (B) Cells are lysed and an anti-Ago-2 antibody is used to pull down RISC/miRNA/target mRNA complex. (C) RNA is extracted from RISC and analysed, in this illustration by microarray but this can also be conducted by qRTPCR.

The identification of multiple cell cycle genes, particularly cyclin E2, as targets of the HCMV encoded miR-US25-1 miRNA is an example among many of the successful implementation of RISC-IP (Grey et al., 2010).

1.4.4.4 HITS-CLIP

In 2005, Ule *et al*. developed the HITS-CLIP technique to identify protein-RNA interaction sites in living cells. The resulting technique reduces false positives that may occur with microarray or IP by the high stringency purification it allows due to the ultra violet (UV) radiation cross-linking of the RNA to proteins before IP. The cross-linking results in the formation of irreversible covalent bonds between the protein and the RNA molecules. An overview of the HITS-CLIP technique is shown in figure 1.6. In the procedure, a cell culture is transfected with a plasmid expressing the miRNA of interest or a vector expressing a

negative control miRNA. Cells are UV-irradiated on ice (b) followed by immunoprecipitation of the RISC complex using an antibody raised against the human Ago-2 protein (c/d). With the miRNA targeted transcript covalently bonded to the protein as well as the miRNA, the remaining parts of the mRNA is digested using an RNase enzyme (e). The RNA is then dephosphorylated, ligated to a 3' RNA linker (f) and radioactively labelled by 5' γ32P (g). The RISC complex is isolated by SDS-PAGE and the complexes are transferred to a nitrocellulose membrane (h). The membrane is then exposed to an X-ray film allowing for the mapping of the region of the membrane containing the RISC complex (h). After excision, the membrane is treated with proteinase K (i) which leads to the dissociation of the RNAs from the protein. A 5' linker is then ligated to the RNA (j) followed by reverse transcription generating a cDNA library which is further amplified by PCR (k). The cDNA is then sequenced (l) and the reads correspond to the miRNA and mRNA originally bound to the RISC complex (Wilbert and Yeo, 2011, Ule et al., 2005).

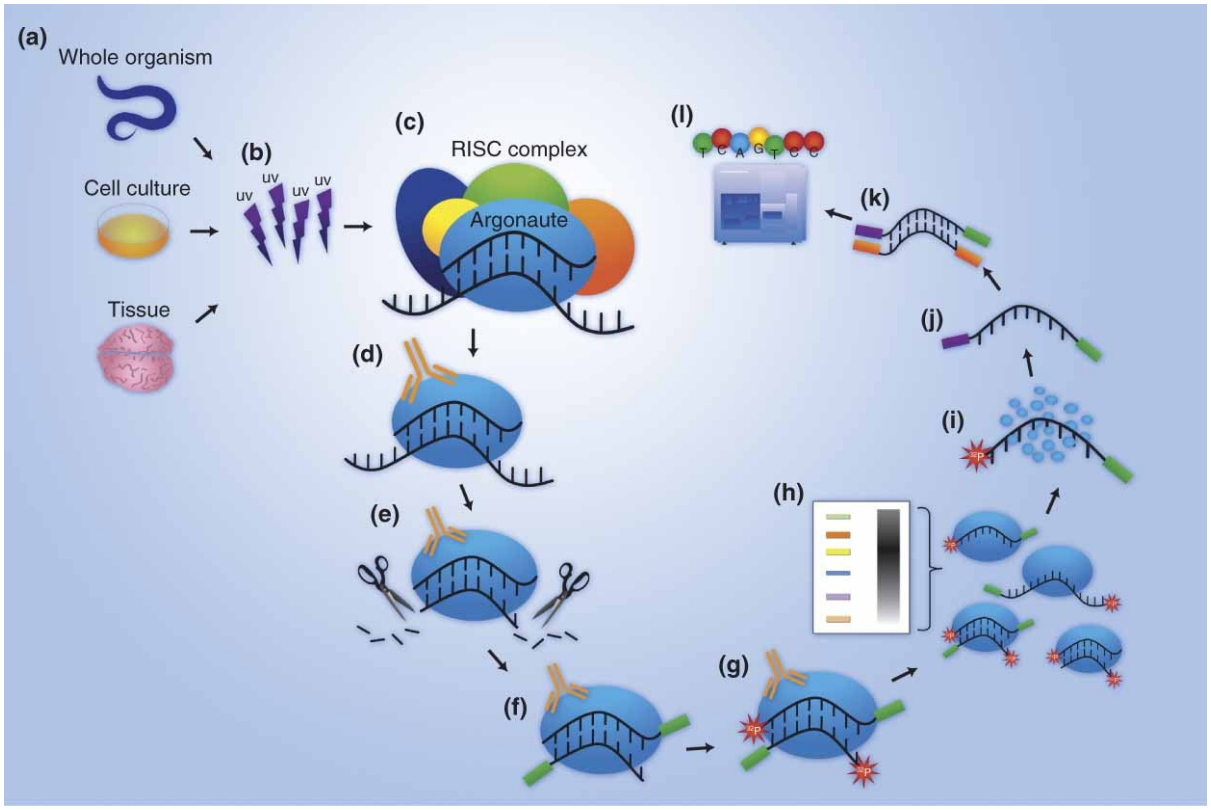


Fig 1.6: Schematic representation of HITS-CLIP procedure (Wilbert and Yeo, 2011)

Riley *et al.* (2012) successfully used HITS-CLIP technique to study EBV miRNA targets. Using a Jijoye Burkitt's lymphona cell line, which expresses 43 of the 44 known EBV miRNAs, they identified 3 viral 3'UTRs and 1503 human 3'UTRs to be co-targetted by both EBV and cellular

miRNAs. 161 human 3' UTRs were also found to be targeted by EBV miRNAs only (Riley et al., 2012). High throughput sequencing results from the HITS-CLIP assay showed that of the EBV miRNAs, 12 of them were highly expressed and represented 90% of the total EBV miRNA reads and these were selected for further detailed analysis. Analysis by bioinformatics revealed that the EBV early lytic cycle protein's, BHRF1, 3' UTR had miRNA binding sites for the human miR-142-3p and miR-17 family as well as the EBV BART-10-3p miRNA. Similarly, the latent membrane protein 1 (LMP1) had two dinstictive 3'UTR binding sites for human miR-17 family, EBV miRNAs BART-19-5p and BART-5-5p. These results were validated by luciferase assays. By using bioinformatics and focusing on the highly expressed EBV and human miRNAs, the same studies reaveled that the EBV and human miRNA co-targeted mRNAs are most particularly involved in transcription, apoptosis and cell cycle pathways. EBV miRNAs targeted genes involved in transcription, apoptosis, Wnt signalling and cell cycle control (Riley et al., 2012).

1.4.4.5 PAR-CLIP

Tushl and colleagues modified the HITS-CLIP technique and developed PAR-CLIP. In this procedure, cells are incubated with a photoactivatable ribonucleoside4-thiouridine before the protein-RNA UV-crosslinking. This added step facilitates the identification of the exact cross-linked sites in the cDNA as the thymidine becomes cytidine in the sequenced cDNA (Hafner et al., 2010, Wilbert and Yeo, 2011). The additional step theoretically results in an improved cross-linking and the refinement of target sites as well as less background noise. An illustration of this procedure is shown in figure 1.7.

PAR-CLIP technique has been successfully used to study miRNA targets for EBV and KSHV viruses. Using primary effusion B cell lines, Gottwein and colleagues revealed that KSHV miRNAs directly target more than 2000 cellular mRNAs and identified miR-K15 to be a mimic of the hematopoietic miRNA miR-142-3p. Their studies also confirmed the encoding of a miR-K11 and its mimicking properties of the hematopoietic miRNA miR-155. KSHV miRNA targets included transcriptional regulation, signal transduction, vesicular trafficking, and the regulation of cell cycle and apoptosis (Gottwein et al., 2011). However, in their report, they clarify that "it is important to point out that PAR-CLIP merely captures interactions, including those that may be transient or may not result in functionally relevant levels of

regulation" (Gottwein et al., 2011). Skalsky *et al.* (2012) investigated EBV miRNA targets using B-95-8 (EBV strain) infected lymphoblastoid cell lines (LCLs), an EBV model for latency studies. In their studies, they combined PAR-CLIP and bioinformatics to identify 531 EBV target sites in cellular 3' UTRs and confirmed 24 of these by luciferase reporter assays. Their findings showed that EBV miRNAs mainly target cellular transcripts which are involved in innate immunity, cell survival and cell proliferation and do so during latent infection (Skalsky et al., 2012).

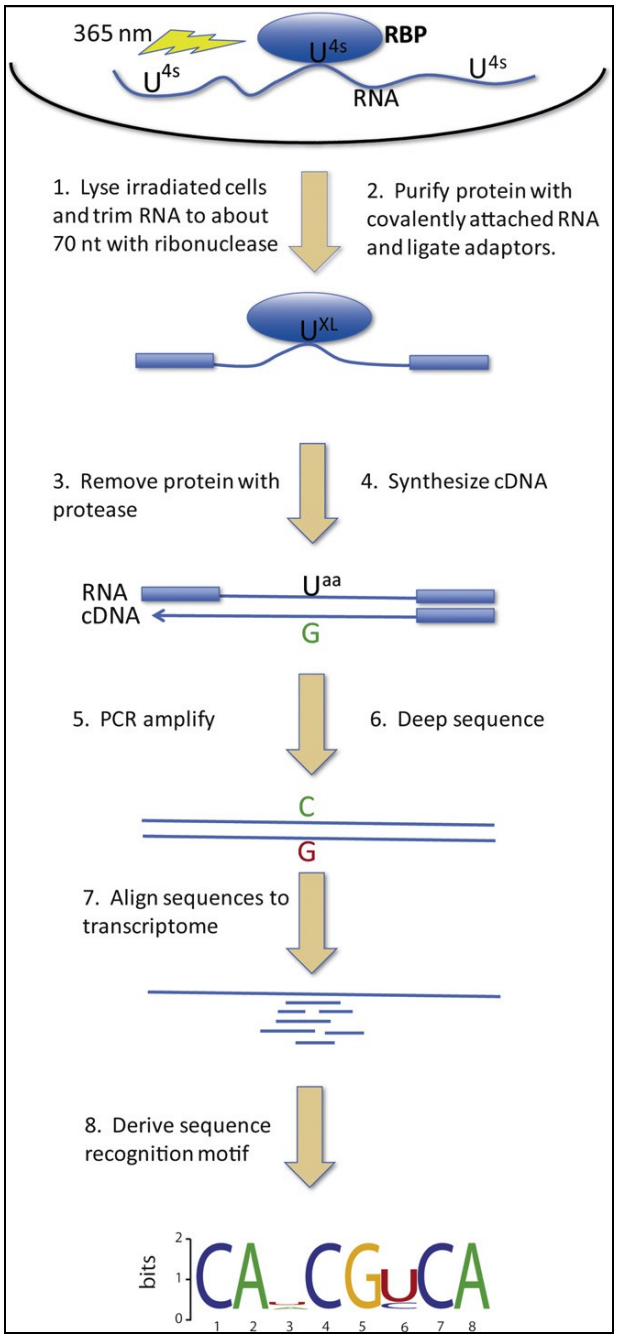


Fig 1.7: Overview of the PAR-CLIP procedure (Corden, 2010).

The discussed bioinformatic and biochemical techniques have been effective in miRNA studies but they have limitations. Bioinformatics as mainly a prediction tool can generate false positives and targets can be missed as false negatives. RISC-IP is dependent on the enrichment of transcripts as well as the microarray analysis of transcripts which is limited by the pre-designed probes. HITS-CLIP and PAR-CLIP techniques offer an improved method to RISC-IP but they lack the definitive target identification. As a result, most of the miRNA targets and functions are unknown highlighting the need of more improved and effective techniques. In this thesis, CRAC, an improvement from the CLIP techniques is adopted and will be described with preliminary work to establish and optimise the technique reported.

Chapter 2: Materials and Methods

2.1 Cell culture

All cell lines were grown in Nunclon delta coated 175 cm^2 flasks in Gibco® Dulbecco's modified eagle medium (DMEM) supplemented with 10% heat-activated foetal bovine serum (FBS) and penicillin/streptomycin (0.5 Units/mL).

2.2 Immortalised cell lines

Human embryonic kidney (HEK) 293T cells were passaged 1:4 when confluent using 0.25% Trypsin-EDTA solution. A HEK293 cell line constitutively expressing Ago2 carrying a dual Protein A and His tag were created using the Invitrogen Flip-in™ system (http://biochem.dental.upenn.edu/GATEWAY/Vector_manual/flpinsystem_man.pdf). These cells were a generous gift from David Tollervey.

2.3 Plasmid DNA

pCMVUS25 was constructed by PCR amplifying a region from the HCMV genome encoding the miR-US25 miRNAs, US25-1, US25-2-3p and US25-2-5p using the primers listed below GCGGAATTCTACAACAGCGGATTCCC, ATAGCGGCCGCACGAGTGTCGTAGGCG. This region corresponds to nucleotide coordinates 214665 to 216130 of AD169 genome (accession number X17403.1). The PCR fragment was cloned into the EcoRI and NotI sites replacing the eGFP cassette of peGFP C1 (Clontech). Expression of all three miRNAs was confirmed by RT-PCR analysis. In control transfections peGFP C1 (Clontech) or pSIREN (Clontech) were used.

2.4 Optimisation Steps

2.4.1 Transfection

To study the targets of specific HCMV miRNAs, plasmid DNA expressing the miRNA of interest was introduced into the HEK293 cell line by transfection using Fugene HD™ (Promega). Six well plates were set up and different amounts of Fugene and plasmid DNA where checked. The ratios resulting in the best transfection levels were determined using fluorescent microscopy. A GFP plasmid (peGFP or pSIREN) was used for this procedure and

the Carl Zeiss Colibri Illumination System microscope used analyse transfection levels. Each well was seeded with 1 x 10^6 cells and transfected after 24 hrs with the amounts of Fugene and plasmid DNA detailed in table 2.1.

Table 2.1: Fugene:DNA ratios used for transfection optimisation.

Fugene (µl)	DNA (µg)
3	1
5	1.5
5	2
7	1.5
9	1
9	2

2.4.2 Induction

The tagged Ago-2 protein expression is doxycycline inducible. To check successful induction and the amount of doxycycline required for induction varying amounts of doxycycline were used and the level of expression analysed by Western blotting. 5 hrs post transfection of a six well plate, 500mg/ml, 250mg/ml, 125mg/ml, 50mg/ml, 5mg/ml or 0mg/ml of doxycycline was added to a well. The cells were harvested 24 hrs from transfection using a gel 1 x loading dye and analysed for the presence of Ago-2 by western blotting.

2.4.3 Dynabeads-IgG Conjugation & Lysate Concentration (Small Scale IP)

To determine the effectiveness of Dynabeads conjugation by IgG antibodies and the effect of lysate concentration on the quality of immunoprecipitation a small scale IP was performed. Quantities representative of a 100 fold less, than those used for the CRAC assay, of the conjugated Dynabeads and lysate volume, 300µg of beads and 100µl of lysate, were used for the IP following the CRAC assay protocol up to RNase IT digestion exclusive. The conjugation of Dynabeads with IgG's effectiveness was checked by comparing the quality of

IP to that of commercially obtained rabbit IgG conjugated agarose beads (Sigma). Three different lysate concentrations were checked, and these were, as obtained from harvesting (100%) concentrated, 1 in 4 dilution of the harvested lysate (25%) and 1 in 10 dilution (10%). The complexes were eluted from the beads using a protein gel loading dye (x1) and these served as western blot samples.

2.4.5 Western Blot Analysis

15µl of each sample was loaded, run on 10% acrylamide gels (appendix 5.2.1.3) for 1.5 hrs at 150 V using the Bio-Rad Mini-PROTEAN® Tetra system, and transferred onto a PVDF membrane (Millipore) by semi-dry transfer using the Bio-Rad Trans-Blot® Turbo™ Transfer system. Blocking of the membrane to minimise non-specific antibody binding was conducted using 5% milk powder (Marvel) in TBST solution (appendix 5.2.1.4). Protein levels were detected using peroxidase-anti-peroxidase (PAP) from Sigma, an antibody specific for Ago-2 protein and do not require a primary antibody. The membrane was washed 3 times, 20 mins each time, using TBST buffer solution. ECL plus agent (Amersham) was used for the detection of protein presence following manufacturer's guidelines and the membrane was visualised under UV using the FluorChem® HD2 gel box.

2.5 Lysate production

Four 15 cm dishes were used for lysate production per every miRNA of interest. 3 x 10^8 cells were seeded per dish and transfected with plasmid DNA after 24 hrs.

2.5.1 Plasmid DNA Transfection and Induction

Following the optimisation steps, 30µg of plasmid DNA was transfected per 15cm dish of 70% confluent HEK-293 cells using 120µl of Fugene HD™ (Promega), as a carrier and Optimem media following the manufacturer's guidelines. Tagged Ago2 expression was induced 5 hours post transfection using 500mg/ml of doxycycline. Supplemented DMEM was changed prior to transfection and also 24 hours post transfection. Transfection levels were checked using the Carl Zeiss Colibri Illumination System Microscope by checking for the green fluorescence on the pGFP transfected cells 24 hours post transfection.

2.5.2 UV Crosslinking and harvesting

Each plate was washed with PBS and immediately UV cross-linked on ice using the UV Stratalinker 1800 at 400 mJ/cm^2. Cells were lysed directly on the plate by adding 2.5 ml of cooled lysis buffer (see appendix section 5.2.4) and then scrapped off. The lysates were incubated on ice for 10 mins, centrifuged at 4500 rpm and 4^{o}C to remove cell debris and stored at -80^{o}C.

2.6 IgG-Dynabeads Conjugation

300mg of Dynabeads® M270 Epoxy from Invitrogen were conjugated using 3, 525µl (from a stock solution of 14mg/ml) of rabbit IgG from Sigma following the optimised and established protocol from Rout laboratory (http://commonfund.nih.gov/pdf/Conjugation-of-Dynabeads.pdf). Preparation and recipes of the required buffers is detailed in appendix section 5.2.1.1 and the protocol followed in section 5.2.2. In summary, Dynabeads were washed 3 times using PBS and incubated with the IgG antibody mix at 30^{o}C with gentle agitation for 18 to 24 hours. Following incubation, conjugated Dynabeads were washed with glycine, Tris, Triethylamine and PBS prior to storing in PBS + 0.02% Sodium azide at 4^{o}C.

2.7 CRAC

A detailed protocol of the steps, procedures and materials used for the CRAC is detailed in the appendix, section 5.2.3. Buffer recipes are shown in section 5.2.1.4. Where changes were made to the procedures or materials used, it has been detailed in the result section prior to the achieved results (chapter 4).

Gel electrophoresis purification of the complexes following radioactive labelling was conducted on a NuPage 4-12% precast gel (10 well, product NP0335BOX, Invitrogen) using the Invitrogen™ Novex® Mini-Cell system. The complexes were transferred onto a PVDF membrane (Millipore) by wet transfer using the Bio-Rad Mini Protean II™ system.

2.8 BLAST (Basic Local Alignment Tool)

The link below was used for the analysis of the CRAC sequence results. The parameters: Human genomic + transcript and Highly similar sequences (megablast), were selected.

http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_PROGRAMS=megaBlast&PAGE_TYPE=BlastSearch&SHOW_DEFAULTS=on&LINK_LOC=blasthome

2.9 Deep Sequencing

The libraries were checked for their quality and quantity using the Agilent Bioanalyser DNA 1000 chip(Agilent Technologies UK Ltd, Stockport, Cheshire) and the KAPA Illumina SYBR Universal Lib Q. Kit (Anachem Ltd, Luton Beds) respectively. The libraries were pooled and the pool concentration was adjusted to 10nM for input to the Illumina Truseq cluster generation kit. The flow cell was prepared using the Illumina Truseq PE Cluster kit V3 (Illumina Ltd., Little Chesterford Essex) following the manufacturer's recommendations. The library pool was loaded onto the flow cell at a final concentration of 9pM using the Illumina cBOT instrument. The sequencing, 100 cycles single end, was carried out using an Illumina HiSeq 2000 instrument with Truseq SBS v3 chemistry.

Chapter 3: RESULTS

3.1 Introduction

A total of 27 HCMV encoded miRNAs has been reported, but the targets and functions of only a few are known. Identifying these miRNAs target transcripts and studying their functions would elucidate why these transcripts are targeted and help us understand the miRNAs functions. As previously mentioned, studies in the Grey laboratory have used a combination of bioinformatics and biochemical techniques to identify targets of two HCMV miRNAs, miR-UL112-1 and miR-US25-1. However, these techniques have limitations to their effectiveness. Therefore there is need for additional technologies for the identification of novel miRNA targets. The aim of this project is to optimise and establish a newly developed technique for miRNA studies, the CRAC assay. CRAC is a cutting edge technique with the ability of identifying a miRNA and its target sequence enabling us to unravel and understand the science behind HCMV miRNA functions.

CRAC follows a similar procedure to that outlined in the HITS-CLIP technique (figure 1.6) with two additional modifications aimed at improving the immunoprecipitation (IP) quality and target transcript identification. The first modification is the double tagging of the Ago-2 protein, as represented in figure 3.1, allowing for a two-step affinity purification. The Ago-2 protein is N-terminally tagged with protein A and histidine tags. The protein-A tag and histidine tag (his-tag) binds to the IgG antibodies and nickel beads respectively.

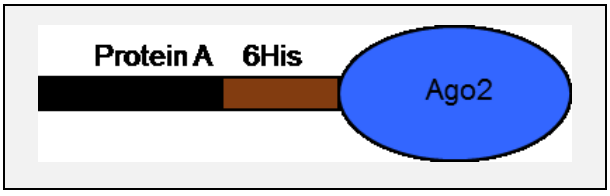


Fig 3.1: Double tagged Ago-2 protein. Six histidine molecules form the his-tag, represented in brown, which is attached to the N-terminal of Ago-2 protein on one end and a protein A tag in black, on the other end.

In addition, an intramolecular-ligation step has been introduced which ligates the end of the RISC associated miRNA to the associated target mRNA resulting in the formation of a miRNA-mRNA hybrid. Sequencing of the hybrid RNA would therefore reveal both the

targeted mRNA and the targeting miRNA. Here I will give more details and illustrate in figures the step by step procedure of the CRAC assay but the figure will focus on the Ago-2 protein and it is worth noting that this is part of the RISC.

3.1.1 UV Crosslinking

The tagged Ago-2 expressing cells are UV irradiated to enable the crosslinking of the RNAs with proteins. This results in the establishment of the stable and irreversible covalent bonding between the RNAs and the proteins. It has been reported that the covalent bonds form between the nucleic acid pyrimidine bases and specific amino acids such as lysine, cysteine, phenylalanine, tryptophan and tyrosine with the RNA-protein complexes needing to be in direct contact (approximately over single angstrom distances) (Budowsky *et al.*, 1976, Chi *et al.*, 2009, Ule *et al.*, 2005).

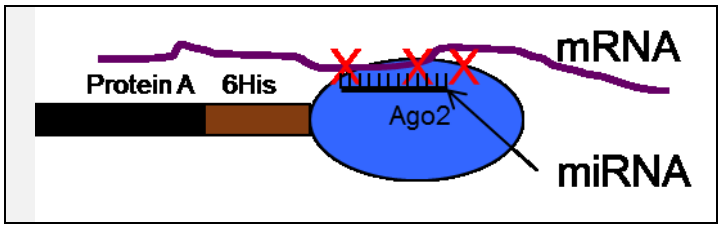


Fig 3.2: UV Crosslinking. A target transcript bound to the miRNA is UV irradiated leading to the formation of covalent bonds between the RNAs and the Ago-2 protein represented as red crosses.

This step allows for stringent purification conditions such as the high salt and denaturing washes conducted throughout the assay while conserving the binding of the miRNA to the target transcript.

3.1.2 Protein A-tag mediated IP-RNase Digestion- His-tag mediated IP

IgG antibody (IgG Abs) conjugated Dynabeads are used for the first IP with the IgG Abs binding to the protein A tag of the Ago-2 protein, Fig 3.3 (a). A magnet is the then used to isolate the IgG bound complexes from the solution followed by low and high salt buffer washes of the IP to eliminate non-specific binding. Limiting RNAse digestion is then conducted to remove the overhanging transcripts leaving the RISC incorporated mRNA which corresponds to the miRNA target (b). A guanidine rich buffer, which causes protein denaturing, is used to elute the complexes from the IgG conjugated Dynabeads. A second IP

is performed using nickel beads that bind to the His-tag of the Ago-2 protein (c). With the complexes still immobilised on the nickel beads, an intra-molecular ligation reaction, resulting in the formation of a hybrid between the miRNA and the mRNA, is performed using a T4 RNA Ligase 1 enzyme (d). T4 RNA Ligase 1 is a single stranded RNA ligase that catalyses the formation of a phosphodiester bond, through ATP hydrolysis, between a 5' phosphoryl-terminated nucleic acid acting as a donor with a 3' hydroxyl-terminated nucleic acid acceptor.

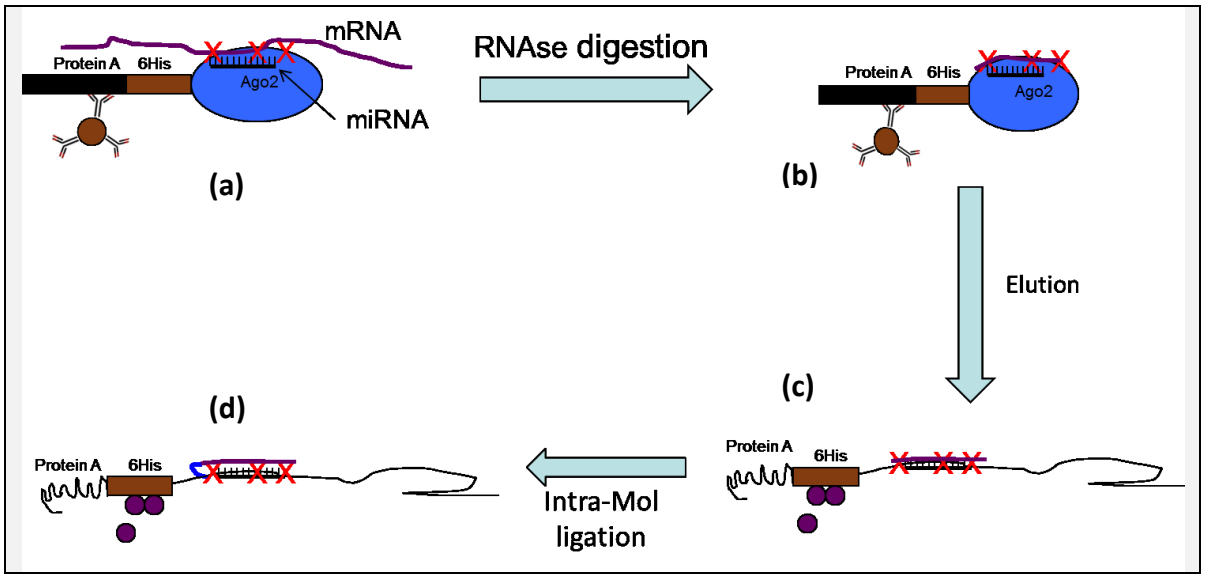


Fig 3.3: Modified steps of CRAC. **(a)** First pull down using the IgG antibody targeting the Protein A tag, **(b)** Second pull down using the nickel beads that target the His-tag, **(c)** Intra-molecular ligation step generating a RNA-miRNA hybrid.

A 3' linker of known sequence is ligated with the complex still immobilised on the nickel beads. This reaction is catalysed by a truncated K227Q T4 RNA Ligase 2 enzyme, a point mutant of T4 RNA ligase 2 that ligates pre-adenylated 5' end of DNA or RNA to a 3' hydroxyl end of RNA without the need of ATP. T4 RNA Ligase 2 mutation reduces the formation of undesired ligation products such as concatemers and circular products. Following ligation the samples are radioactively labelled at the 5' ends of the RNA using γ-^{32}P ATP (figure 3.4).

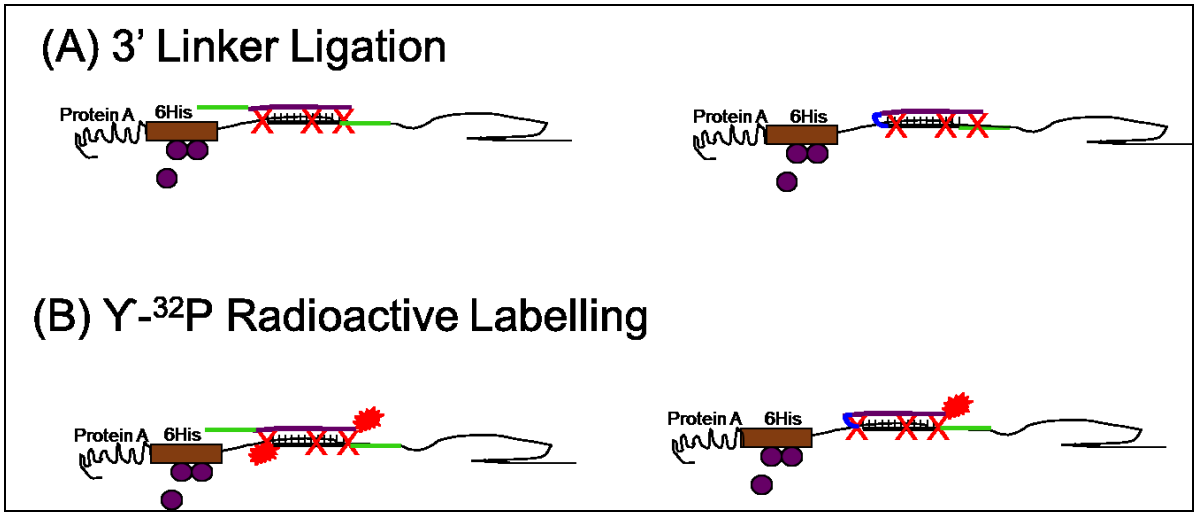


Fig 3.4: Representation of the Intra-molecular ligation reaction and radioactive labelling. (A) - represents the possible outcomes of intra-molecular ligation where a mixed population of non-ligated (left panel) and ligated shown with blue linker (right panel) is generated. 3' Linker is shown as a green line. **(B)** - represents the radioactive labelling of the 5' ends ( ).

An imidazole containing buffer is used to elute the complexes from the nickel beads via competitive binding to the His-tag. Following elution, the same steps, as detailed in the PAR-CLIP technique from (h) to (l), are performed. These include SDS polyacrylamide gel purification of the complexes after radioactive labelling followed by transfer to a membrane and extraction of the RNA by proteinase K digestion. Proteinase K fragments the proteins leading to the release of miRNAs and target transcripts. A 5' linker of known sequence is then ligated to the IP products using T4 RNA ligase 1. The 5' linker contains an additional known variable sequence referred to as a barcode. Ligation of 5' linker oligonucleotides with different barcode sequences allows for subsequence multiplexing of deep sequence reactions. The 3' and 5' linkers allows for the amplification of final cDNA CRAC product with primers designed to match linkers of known sequence. The PCR primers have additional nucleotides that create overhanging known sequences referred to as adapters. Adapters are used in deep sequencing to bind the immobilised solid support chip oligos for cluster generation as explained below. A typical cDNA fragment from a CRAC assay, including its features, is illustrated in figure 3.5.

Fig 3.5: An example of a cDNA fragment for deep sequencing. PCR primers are underlined, 5'Linker highlighted, 3' Linker highlighted, Bar code in italics and the RNA insert is shown in bold. Non-highlighted and underlined sequences represent the adapters from PRC primers.

The deep sequencing procedure is best partitioned into three steps, library preparation, cluster generation and sequencing. The cDNA library is prepared for deep sequencing by purification and size selection of fragments with 200 to 300 base pairs. Clusters are then generated by isothermal amplification of single molecules in a flow cell on a solid support chip cluster system. The chip is densely coated with a network of oligos on its surface that binds to the fragments' ligated adapters. Successfully bound oligos are extended, creating copies that are covalently bound to the flow cell surface. Clonal amplification through a series of extension and bridge amplification leads to the generation of hundreds of millions of unique clusters. The reverse strands are cleaved and washed off followed by the blocking of fragment ends and hybridisation of the sequencing primer and at this stage the libraries are ready for sequencing. Hundreds of millions of clusters are sequenced simultaneously, base by base leading to the generation of results in a short period of time (up to a week). Four fluorescently labelled and reversibly terminated nucleotides are used to sequence the templates in parallel. The bases compete for template binding and this natural competition facilitates accuracy. After each round of synthesis, the clusters are excised by a laser emitting a colour that identifies the newly added base. The fluorescent label and blocking group are then removed for the addition of a new base (Ansorge, 2009). Deep sequencing generates millions of reads which are then analysed by technical bioinformatic techniques. The ability of deep sequencing to sequence clusters simultaneously allows for the sequencing of a mixture of different samples at the same time. The bar code within the 5' linker will then be used to identify specific samples within the sequence results. As figure 3.5 illustrates, analysis of the sequence results involves the identification of the 5' and 3' linkers in order to determine the insert. These linkers are also useful in the reverse transcription and PCR stages of the assay as primers are designed from their known sequences.

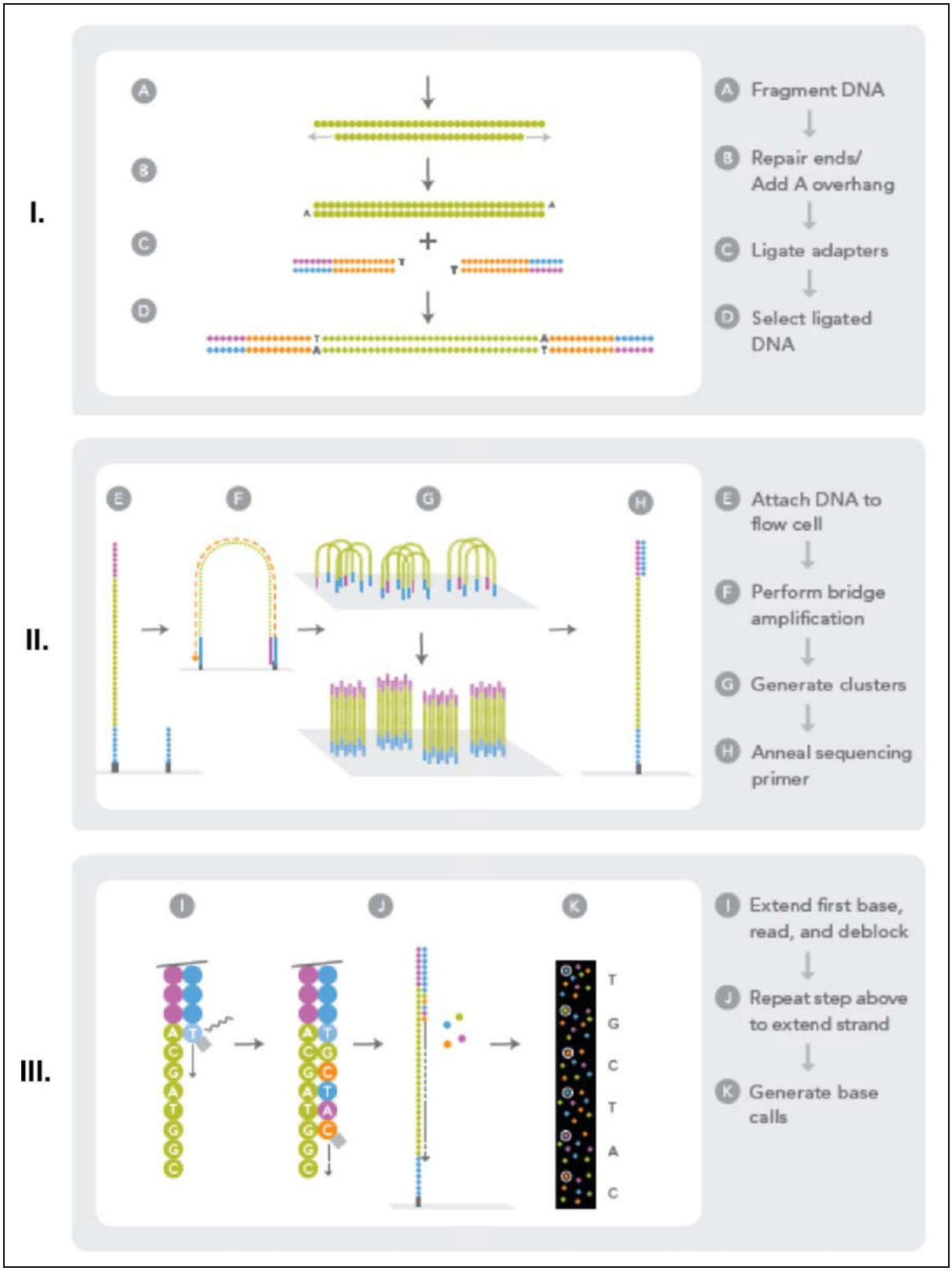


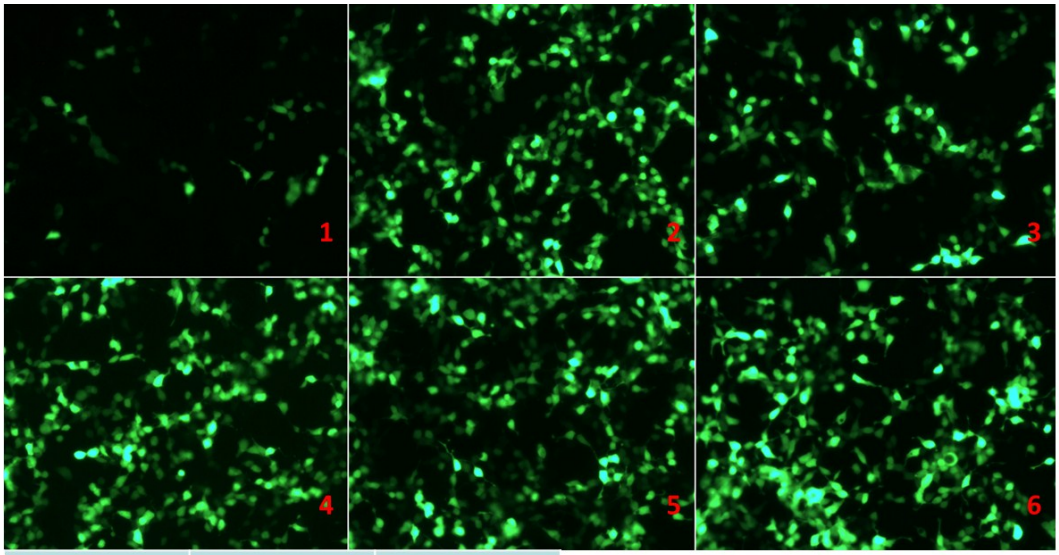
Fig 3.6: Outline of the Illumina Genome Analyser workflow (Ansorge, 2009). The three steps involved are the library preparation (I), Cluster generation (II) and sequencing-by-synthesis reactions (III).

3.2 CRAC Optimisation

Initial experiments were carried out to optimise and establish the best conditions for efficient transfection of the HEK-Flp293 cells with the plasmid expressing viral miRNAs of interest and induction of the tagged Ago-2 protein. HEK-Flp293 is a cell that stably expresses the doxycycline inducible tagged Ago-2 and was created using the Flip in™ expression vector system. The efficiency of IgG antibody conjugation to the Dynabeads was checked against commercially conjugated agarose beads.

3.2.1 Transfection

Previous experiments in the Grey laboratory have demonstrated that transfection of HEK293 cells with Fugene HD results in efficient introduction of plasmid DNA with little or no cytotoxic effects on the cells. To establish the best transfection conditions, a six well plate was seeded with 1 x 10^6 cells per well and transfected with varying ratios of GFP plasmid DNA and Fugene as detailed in the table in figure 3.7.



Panel	Fugene (ul)	DNA (µg)
1	3	1
2	5	1.5
3	5	2
4	7	1.5
5	9	1
6	9	2

Fig 3.7: Transfection optimisation results. Fluorescent Microscopy images demonstrate different transfection levels from using varying Fugene and plasmid DNA amounts as detailed in the table below image.

The fluorescent microscopy pictures in fig 3.7 indicate the different levels of transfection from using different amounts of Fugene and plasmid DNA. The Fugene:DNA ratio manufacturers' recommended to use per well of a six well plate are 3µl:1µg. Few green cells were observed from transfecting the cells with the recommended amounts. Our results however show that optimal levels of transfection were achieved with higher Fugune:DNA ratios. For reagent and cost efficiency, a 5:1.5 (Fugene:DNA) ratio was chosen for future experiments.

3.2.2 Induction

The Ago-2 protein in HEK 293flp cells is doxycycline inducible. Using different doxycycline concentrations, ranging from 0µg/ml to 500µg/ml, the best concentration for efficient induction was determined. Following induction of a six well plate, the levels of expression of the Ago-2 protein was confirmed by western blotting using an antibody that recognises the Protein A tag.

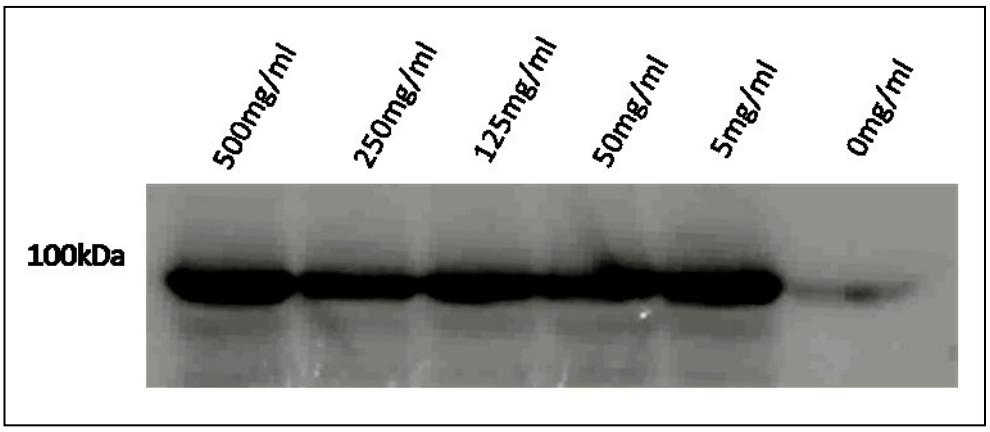


Fig 3.8: Induction check. Western blot results for checking levels of induction using different doxycycline concentrations.

Western blot results were indicative of successful induction of the Ago-2 protein with an exception of when 0µg/ml of doxycycline was used. There was no difference in the resulting expression from the rest of the varying doxycycline amounts used. The cell line was therefore positively inducible by doxycycline.

3.2.3 Conjugation of IgG to Dynabeads

Previous studies by the Tollervey laboratory have demonstrated that IgG conjugated Dynabeads produce the best results in the context of CRAC assays with good efficiency of pull down with low levels of non-specific binding. An initial procedure is required to conjugate purified IgG antibody to magnetic Dynabeads (see materials and methods). The IgG antibodies were conjugated to Dynabeads and checked against commercially produced agarose beads for successful conjugation and pull down efficiency. The effect of lysate concentration on the IPs was also investigated to determine an optimum starting lysate concentration for IP. For a full scale CRAC assay, four 15cm dishes are seeded with 3×10^7 HEK-293 cells, transfected and harvested into 10ml of lysate. 30mg of conjugated Dynabeads are used for IP. For this optimisation procedure, a small scale IP experiment where 300µg of conjugated Dynabeads were used for an IP from 100µl, lysate was set up. Following IP, complexes were eluted from the Dynabeads using a protein loading dye. These served as samples for analysis using western blotting. Three different lysate concentrations were checked, as obtained from harvesting (100% concentrated), 1 in 4 dilution of the harvested lysate (25%) and 1 in 10 dilution (10%).

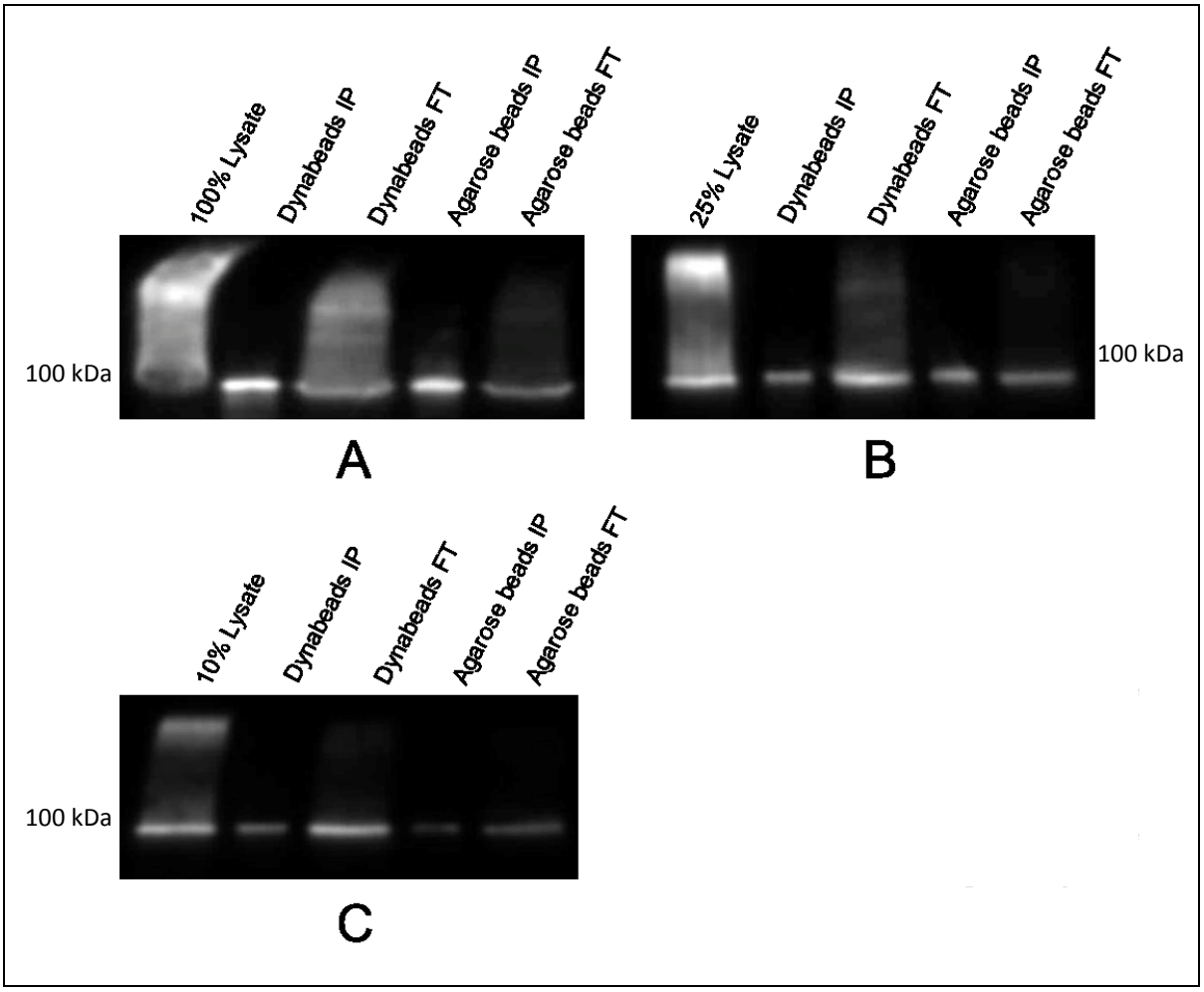


Fig 3.9: Dynabeads conjugation check and effect of lysate concentration in IP effectiveness. Panel A shows IP results from 100 lysate, B from 25% lysate and C from 10% lysate. IP = immunoprecipitation sample and FT = flow through sample.

The western blot results detected the presence of the Ago-2 protein in the lysates and IP samples. Although the results are not quantitative, it is clear that the conjugated beads resulted in a relatively effective IP compared to agarose beads. The same western blot results also indicate that the extent of complex isolation has a positive correlation with lysate concentration and therefore the best results were from the IPs conducted from higher lysate concentration (100 %).

3.2.4 Discussion

Following these optimisation steps, it was decided that transfection levels obtained from using the Fugene:DNA ratio of 5:1.5, equating to 120µl of Fugene and 30µg of plasmid DNA for a 15 cm dish, were sufficient for CRAC sample production. Ago-2 induction was to be

conducted using 500µg/ml of doxycycline. IPs were also to performed from 100% lysate using IgG conjugated Dynabeads for the first IP.

3.3 CRAC

Having established the conditions summarised in section 3.2.4, we followed the protocol as detailed in the appendix section 5.2.3. In this project, using a plasmid encoding for the miR-US25 cluster (pCMV US25) we aim to study the HCMV US25 miRNAs. The specific miRNAs expressed by this plasmid (figure 3.10) are miR-US25-1-5p, miR-US25-2-3p and miR-US25-2-5p. Several transcripts have been identified to be targeted by the miR-US25-1, therefore using this plasmid, we can refer back and confirm these findings as well as potentially identify new target transcripts for any of the three miRNAs.

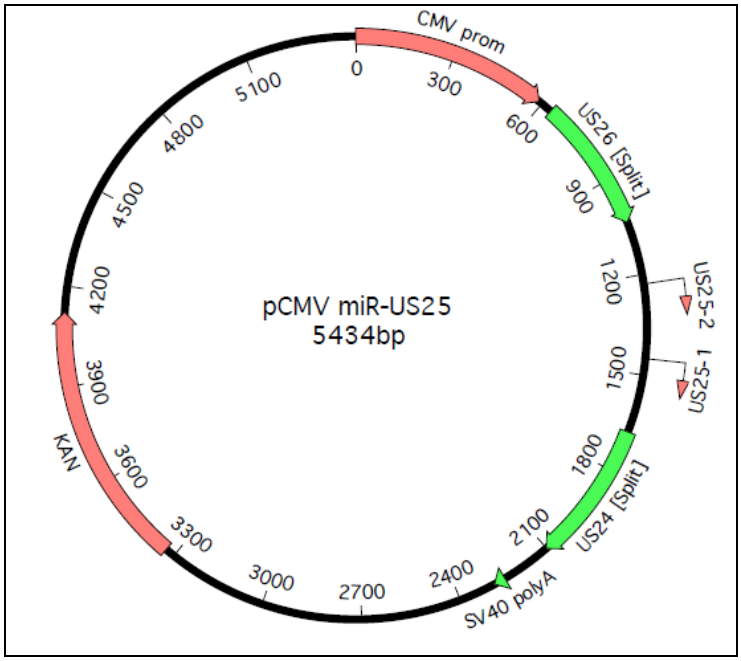


Fig 3.10: miR-US25-1 and -2 expressing plasmid with a CMV promoter upstream of the miRNA genes.

pCMV US25 was transfected into HEK-293 cells and cellular lysates generated as detailed in the section 2.5 to obtain samples for CRAC analysis. A parental plasmid with GFP was used for the generation of a negative control sample. Three other HCMV miRNAs, miR-US5-2, miR-UL22A-1 and miR-UL112-1 were also studied using the CRAC assay.

In the initial CRAC attempt, a 3,000 Ci/mmol γ-P^{32} radioactive isotope was used for the radioactive labelling, but unfortunately no signal was generated on the film after an overnight exposure. A higher activity γ-P^{32} isotope of 6,000 Ci/mmol was then used for the proceeding assays.

3.3.1 CRAC Attempt 2

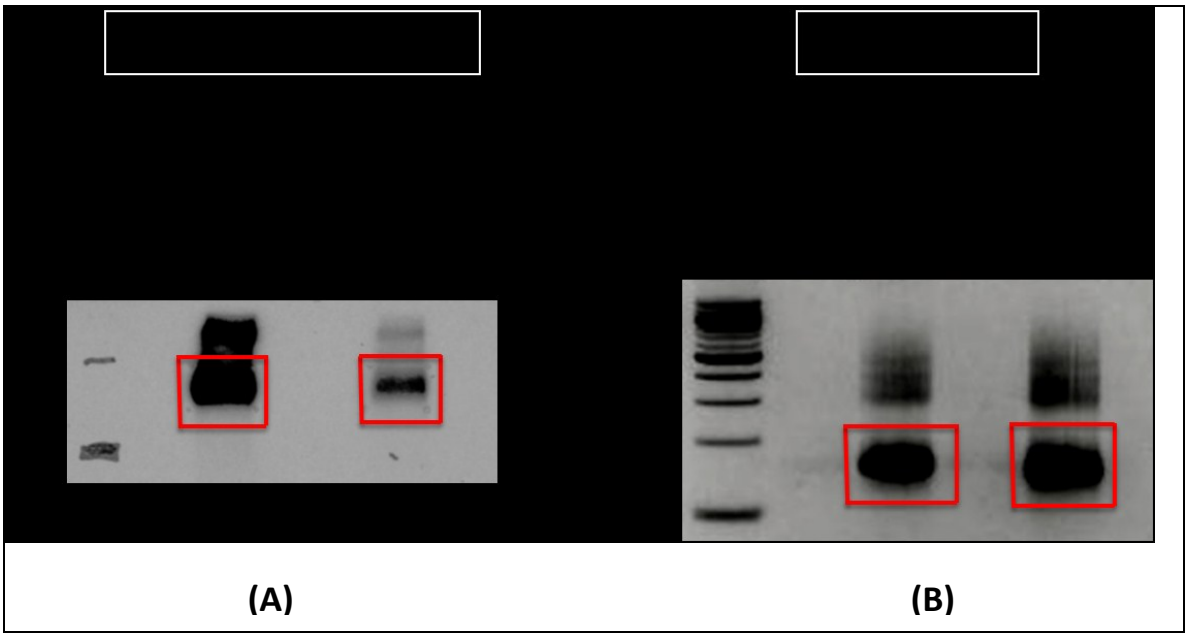


Fig 3.11 (A) developed film from membrane exposure and (**B)** the proceeding PCR products gel results. Red boxes represent excised sections of the membrane and gel.

After gel purification, the complexes were transferred onto a membrane which was exposed to film overnight. The results from the developed film are shown in fig 3.11(A). Clearly, the results show a difference in the amounts of complexes present from the two IPs with the pSiren GFP sample showing the presence of more signal compared to pCMV25 sample. This could be a result of loss of material during precipitation stage of the protocol that leads to the generation of a highly unstable pellet. The presence of other species of a higher molecular weight that leads to the smear above the band of desired size is also difficult to explain. A likely possibility could be less effective stringent washes during the IP and purification stages.

Ago-2 has a molecular weight of approximately 100 kDa and a band representing this was identified (figure 3.11). Regions of membrane indicated by the red boxes were excised and the remaining procedures of the protocol were conducted. cDNA from reverse transcription

is PCR amplified and the products separated by gel electrophoresis using a 3% metaphor agarose gel and the results are illustrated in Figure 3.11(B). PCR gel purification results show a characteristic prominent band at approximately 150 base pairs with a smear above the band. It is difficult to explain what the contents of the smear are, but it is clear from the sizes of the contents, approximately 300 to 600 bp, that they are not the expected products. The prominent band however is within the expected products' size range, therefore the section indicated by the red box was excised from the gel and products were extracted for further analysis.

To determine the quality of these results, cDNA was TA cloned using the One Shot TOP10 Invitrogen Topo™ plasmid kit. Following transformation of bacteria, individual clones were picked, mini prepped and sequenced. Twelve samples derived from cells transfected with pCMV miR-US25 were sequenced and the results are summarised in table 3.1. At this stage, results would confirm successful ligation of the 3' and 5' linkers. Inserts of 20 to 50 bases would be expected, demonstrating successful IPs. However, due to the nature of the small scale sequencing, a full picture of the contents of the cDNA library is limited compared to the results deep sequencing would generate. Therefore, the presence of longer sequences representing miRNA-RNA hybrids formed from the intra-molecular ligation step would be difficult to detect from small sequencing results. Analysis of the sequencing results involved locating the insert sequence flanked by both 5' and 3' linkers, as shown in fig 3.12.

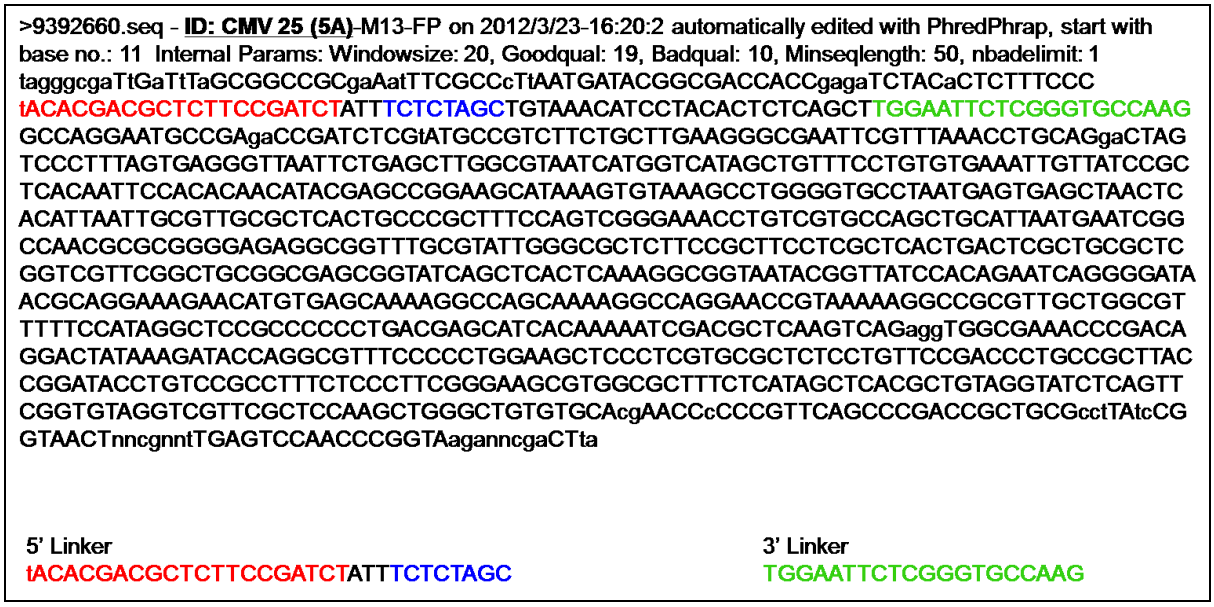


Fig3.12: Demonstration of a small scale sequencing result from a TOPO cloned cDNA sample. 5' linker is shown in red and the barcode (within the 5' linker) in blue. 3' linker is shown in green. The sequence between the linkers is identified as the insert and using the NCBI BLAST tool can be analysed and identified.

Table 3.1: Small sequencing results.

CRAC 2				
Sample	**Insert Sequence**	**Seq length**	**% match**	**BLAST Result**
CMV25 (1)	TGGAGAGAAAGGCGGTTCCTG	21	95% (20/21)	MIR185-5p
CMV25 (2)	ACCCGCGATCTTCGGTTTACG	21	67% (14/21)	Homo sapiens solute carrier family 9, member 11
CMV25 (3)	T	1	N/A	N/A
CMV25 (4)	GTCCGAGCCACTGAGCGGTT	20	75% (15/20)	inositol 1,4,5-trisphosphate receptor, type 2
CMV25 (5)	ACGATCCGAGCCACTGAGCGGTT	23	65% (15/23)	inositol 1,4,5-trisphosphate receptor, type 3
CMV25 (7)	GGTCCGAGCCACTGAGCGGTT	21	71% (15/21)	inositol 1,4,5-trisphosphate receptor, type 4
CMV25 (8)	AGACCCAGTAGCCAGATGTAGCT	23	100% (23/23)	MIR222-3p
CMV25 (9)	G	1	N/A	N/A
CMV25 (10)	TTCCCCATGGGGAGCT	16	N/A	No significant homology.
CMV25 (11)	No 3' Linker detected	N/A	N/A	N/A
CMV25 (12)	AGACCCAGTAGCCAGATGTAGCT	25	100% (25/25)	MIR222-3p

Analysis of the sequencing results revealed the presence of an insert between the linkers. This suggests the RISC-IPs were successful also supported by the BLAST results, which

indicated that three of the sequenced samples were miRNAs - miR-185-5p and miR-222-3p. The detection of transcripts, inositol 1,4,5-triphosphate receptor and the homo sapiens solute carrier, from the BLAST results was also encouraging. As expected, longer insert sequences were not detected as 23 bases was the longest sequence achieved. The shorter sequences obtained from this assay could however be reflective of the extent of RNase IT enzyme digestion. Two sequences had a single nucleotide insert and a 3' linker was not detected from one of the sequences. Therefore the effect of enzyme concentration was studied by using less concentrated RNase IT. For this assay, the stock RNase IT of 5,000 units/ml was diluted by a factor of 10 (2µl + 18µl H_2O) to give 0.5 units/µl and the final amount used for the reaction was 0.5 units (1µl of the diluted enzyme).

3.3.2 CRAC Attempt 3

The protocol was followed as in the appendix (section 5.2.3) but the amount of RNase IT for the on-bead digestion was altered. The stock RNase IT was diluted by a factor of 20 and the final amount used for the reaction was 0.25 units.

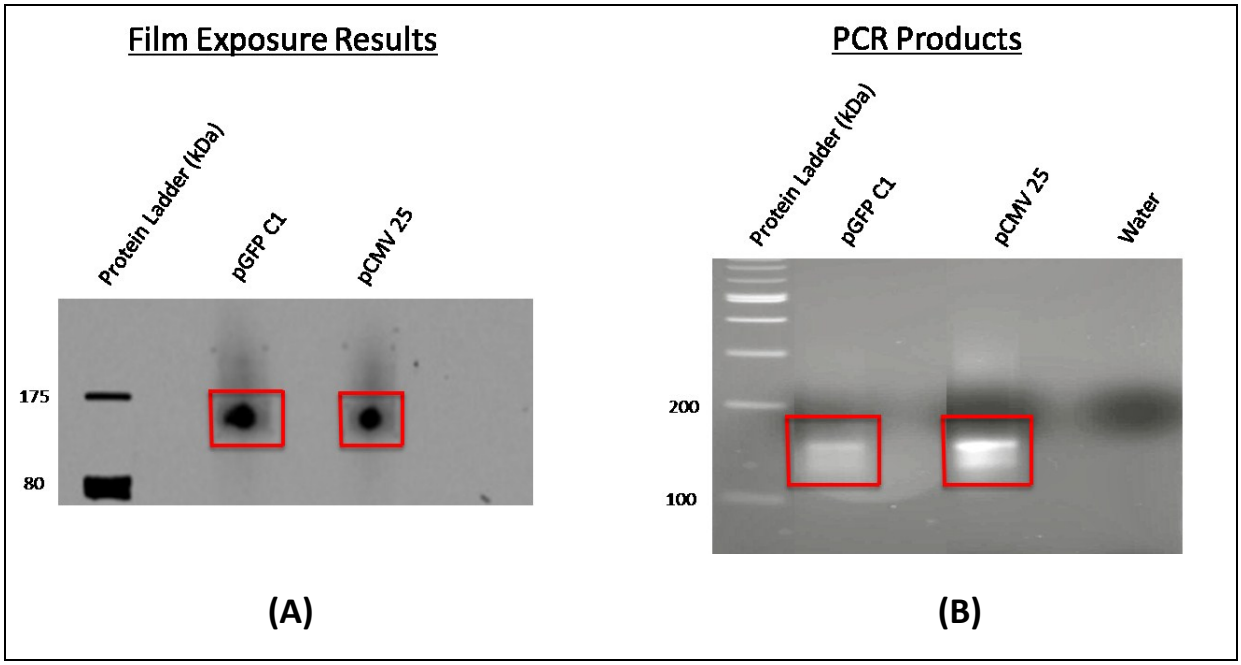


Fig 3.13: Developed film from membrane exposure (A), and the proceeding PCR gel results (B).

The film was developed after an hour's exposure of the membrane to the film. The band density can be viewed as weak but is proportional to the duration of membrane exposure to

the film, i.e. 1 hour. The membrane picture could also be representing a more successful stringency IP and purification during the procedure as there is less smearing.

Bands representing the Ago-2 protein were excised as represented by the red boxes in (A) and RNAs extracted. The gel from PCR purified products of the samples is shown in (B). The bands were, however, faint without the characteristic smearing expected from the samples. Various reasons could explain the resulting faint bands including loss of material during some of the purification stages, particularly the sensitive precipitation stage. Gel loading was also challenging as samples proved to be volatile resulting in loss during loading of samples into wells. The products were nonetheless extracted from the marked area, cloned and DNA was purified from twelve individual clones for small scale sequencing.

Table 3.2: Small sequencing results.

CRAC 3				
Sample	**Sequence**	**Seq length**	**% match**	**BLAST Result**
CMV25 (1A)	TGCAAATCTATGCAAAACTG	20	100% (20/20)	microRNA 19a (MIR19A)
CMV25 (3A)	TACGGGCCACC	11	N/A	Sequence too short
CMV25 (4A)	TCCCTGGTGGTCTAGTGGCTAGG	23	100% (23/23)	hypothetical LOC100505703
CMV25 (5A)	TGTAAACATCCTACACTCTCAGCT	24	100% (24/24)	microRNA 30C-1 (MIR30C1), microRNA 30C-2 (MIR30C2)
CMV25 (6A)	TAGCAGCACATCATGGTTTACA	22	100% (24/24)	microRNA 15b (MIR15B)
CMV25 (7A)	TAGCTCAGCGGTT	13	N/A	Sequence too short
CMV25 (8A)	GTCCGAACGCTAGGTTGGTTCT	22	21/22 (95%)	HCMV miR-US25-1-3p
CMV25 (9A)	GAACGCTAGGTCGGTTCT	18	77% (14/18)	NADH dehydrogenase (ubiquinone) 1 alpha subcomplex 4
CMV25 (10A)	TCCCTGGTGGTCTAGTGGC	19	100% (19/19)	hypothetical LOC100505703
CMV25 (12A)	TAGCAGCACGTAAATATTGGCG	22	100% (22/22)	microRNA 16-1 (MIR16-1)

Following analysis of the sequences, all twelve samples showed the presence of an insert. The minimum length of the inserts was eleven nucleotides, an improvement from the previous assay where the 3' linker was not detected in one sample and two other samples had an insert of one nucleotide. Six of the sequences showed 100% homology to the BLAST results. Of great interest was the presence of an HCMV miRNA, highlighted in red, as one of

the inserts. This is an encouraging result showing successful processing of the transfected pre-miRNA and incorporation into the RISC complex. The result further demonstrates that miRNAs transfected into the cells can be isolated together with cellular miRNAs.

3.2.4 CRAC Attempt 4

The CRAC assay protocol was followed with the amount of RNase IT further altered. The stock RNase IT was diluted by a factor of 160 and the final amount used for the reaction was 3.125 x 10^{-2} units. In this assay expression constructs for 3 other viral miRNAs, pUS5-2, pUL22A-1 pUL112-1, were used.

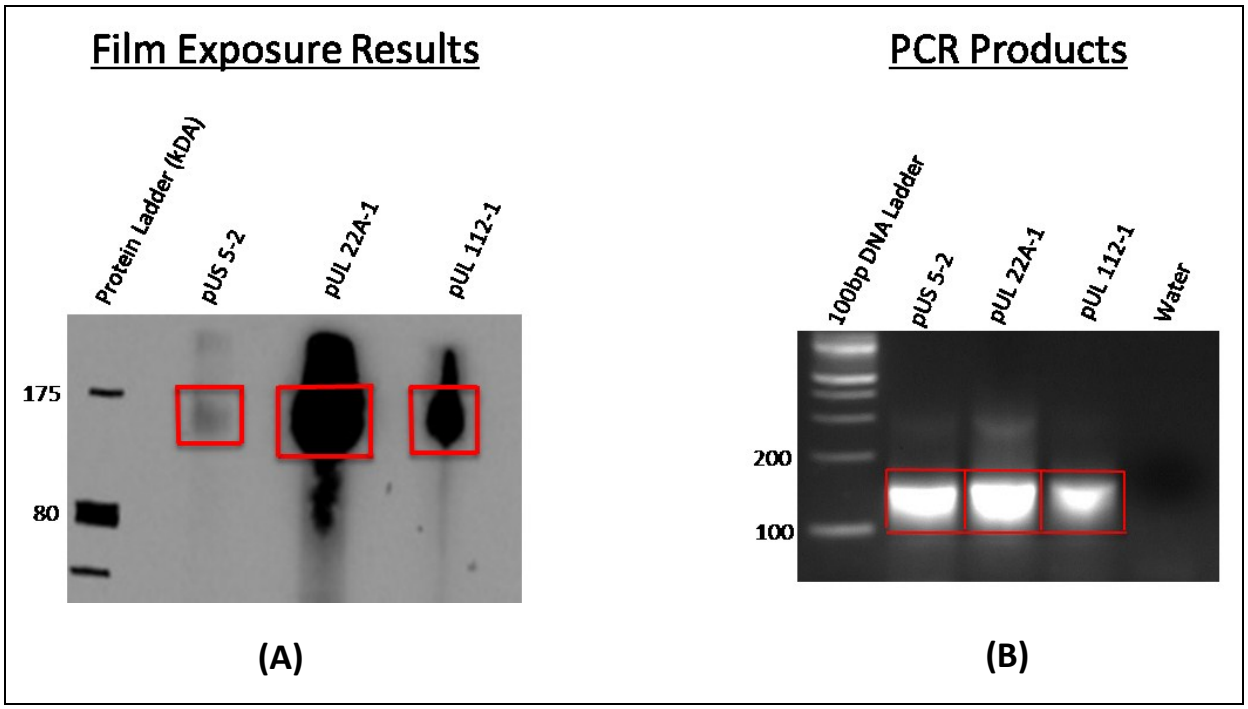


Fig 3.14: Developed film from membrane exposure and the proceeding PCR gel results.

The film was developed after overnight membrane exposure. The membrane picture clearly shows the presence of products in different amounts. pUS5-2 had the most faint band. pUL22A-1 and pUL112-1 also showed the presence of material of lager size than expected. Smaller products were also present in the pUL-22A-1 sample as shown by the smear below the main band.

Bands representing the Ago-2 protein were excised as represented by the red boxes in (A) and RNAs extracted. The gel from PCR purified products of the samples is shown in (B). Despite less signal in with the pUS5-2 sample from the immunoblot, equivalent levels of PCR

product were obtained, suggesting amplification reached saturated levels. Gel loading problems with the samples "jumping-out" of the wells could have led to the resulting loss material from the pUL-112-1 sample. Although the bands were prominent and well defined, the characteristic smearing expected from the samples was only present, to a small extent, in the pUL22A-1 sample. The products were nonetheless extracted, from the marked area, and the pUL22A-1 products were cloned and DNA was purified from twelve individual clones and sequenced.

Table 3.3: Small scale sequencing results

CRAC 4				
Sample	**Sequence**	**Seq length**	**% match**	**BLAST Result**
UL22A - 1 (2)	TAACTAGCCTTCCCGTGAG	19	100% (19/19)	HCMV miR-UL22A-1
UL22A - 1 (3)	AACTAGCCTTCCCGTGAGA	19	100% (19/19)	HCMV miR-UL22A-1
UL22A - 1 (5)	AGTGATTTGTTGTT	14	100% (14/14)	protein (Cdc42/Rac) - activated kinase 2 (PAK2)
UL22A - 1 (6)	CCCTGGTGGTCTAGTGGCTAGGAT	24	100% (24/24)	hypothetical LOC100505703
UL22A - 1 (7)	TAACTAGCCTTCCCGTGAGA	20	100% (20/20)	HCMV miR-UL22A-1
UL22A - 1 (8)	CAGAATGCTAGTTTGTAG	18	N/A	No significant homology
UL22A - 1 (9)	AACTAGCCCCGTGAGA	16	100% (16/16)	Toll-interleukin 1 receptor (TIR)
UL22A - 1 (10)	TAAAGTGCTGACAGTGCAGA	20	100% (20/20)	microRNA 106b (MR106B)
UL22A - 1 (11)	ACTGCTGAGCTAGCACTTCCCGA	23	100% (23/23)	microRNA 93 (MIR(93)
UL22A - 1 (12)	ATGCTAGTTTGTAG	14	N/A	No significant homology

Sequencing results indicate some change in the length of the inserts, with the shortest inserts having 14 nucleotides compared to results from the previous assay, which had the shortest insert of 11 nucleotides. 25% of the sequenced samples returned the presence of the HCMV miR-UL22A-1 miRNA. Sequences that returned results on BLAST search tool had a 100% match and transcripts plus a miRNA were detected.

3.2.5 CRAC Attempt 5

This assay aimed at investigating the reaction time during on bead digestion. The CRAC assay protocol was followed, as well as altering the amount of RNase IT, the reaction was also altered. The stock RNase IT was diluted by a factor of 80 and the final amount used for the reaction was 6.25 x 10^{-2} units. To investigate the effect of reaction time, RNase IT on-bead digestion was carried out for 3 mins as opposed to the 5 mins on the protocol.

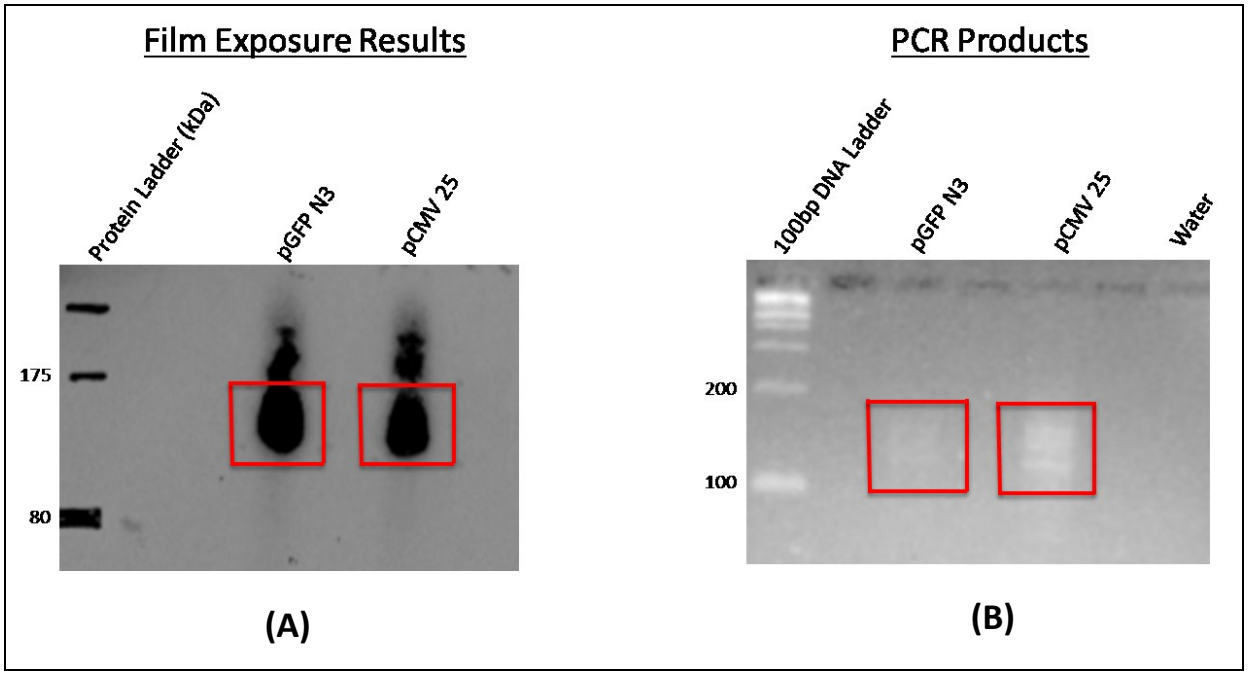


Fig 3.15: Developed film from membrane exposure (A) and the proceeding PCR gel results (B).

The membrane was exposed to the film for an hour and the results from the developed film are shown in (A). The resulting PCR gel however indicted the presence of the products in small amounts. They were however cloned and 12 individual clones picked for each sample, DNA purified and sequenced. The results are summarised in table 3.4.

Table 3.4: Small sequencing results.

CRAC 5				
Sample	**Sequence**	**Seq length**	**% match**	**BLAST Result**
CMV25 (1B)	TAACTGCTCAGTGGCTCGGACCG	23	21/23	HCMV miR-US25-1-5p
CMV25 (2B)	TAAAACTCA	9	N/A	Sequence too short
CMV25 (3B)	No Insert	0	N/A	N/A
CMV25 (4B)	TATTGCACTTGTCCTGGCCTG	21	95% (20/21)	microRNAs 92a-1, 92a-2 & 17-92 (MIR92a-1, 92a-2, MIR17-92)
CMV25 (5B)	No Insert	0	N/A	N/A
CMV25 (6B)	No Insert	0	N/A	N/A
CMV25 (7B)	AACCGCTCAGTGGCTGGAC	19	15/19 (79%)	hypothetical LOC100507663
CMV25 (8B)	AATTCCCCGACGGGTT	16	13/16 (91%)	Homo sapiens phosphohistidine phosphatase 1
CMV25 (9B)	TCGGGTATGAGCC	13	13/13 (100%)	Homo sapiens CDC42 effector protein (Rho GTPase binding) 4
CMV25 (10B)	CTCAGGGTCGTGGGTTCGAGCCCC ACGTTGGGCGCC	36	N/A	No significant homology
CMV25 (11B)	TAAAACTCAA	10	N/A	Sequence too short
CMV25 (12B)	No Insert	0	N/A	N/A
pGFP N3 (1)	5' Linker Barcode not detected	N/A	N/A	N/A
pGFP N3 (3)	GAATCCCAAAAGCAGCT	17	17/17 (100%)	microRNA 191 (MIR191)
pGFP N3 (4)	ATACTGCCATGCTAGAGGACT GGAGAGGTG	30	N/A	No Significant homology
pGFP N3 (6)	GGGGAGTACGGTCGCAAGACTAA AACTCA	29	N/A	No Significant homology
pGFP N3 (8)	5' Linker Barcode not detected	N/A	N/A	N/A
pGFP N3 (9)	AAGGGAGAAATCAAGTAAATT ATCCAAGG	29	29/29 (100%)	Chromosome 12 genomic contig, GRCh37.p5 Primary Assembly
pGFP N3 (10)	No Insert	N/A	N/A	N/A
pGFP N3 (11)	GAGCATGAGACTCtTAAT	18	17/18 (94%)	COMM domain containing 2 (COMMD2)
pGFP N3 (12)	GAcTAAAACTCA	12	N/A	Sequence too short

Sequencing results from this assay (table 3.4) indicates the presence of inserts with a broader range, from 0 to 36 nucleotides in length. Four different sequences of lengths 29,

29, 30 and 36 were encouraging. However analysis of the sequences has failed to determine what they are composed of. As with previous assays, a HCMV miRNA was detected as well as cellular miRNAs and transcripts. The same conditions were maintained for the subsequent assay.

3.2.6 CRAC Attempt 6

Using duplicate samples, 2 lysate samples for pCMV25 and 2 for pGFP transfected HEK-293 cell lines, the protocol was followed with the conditions in CRAC attempt 5 maintained, i.e. 6.25 x 10^{-2} units of RNase IT and a reaction time of 3 mins.

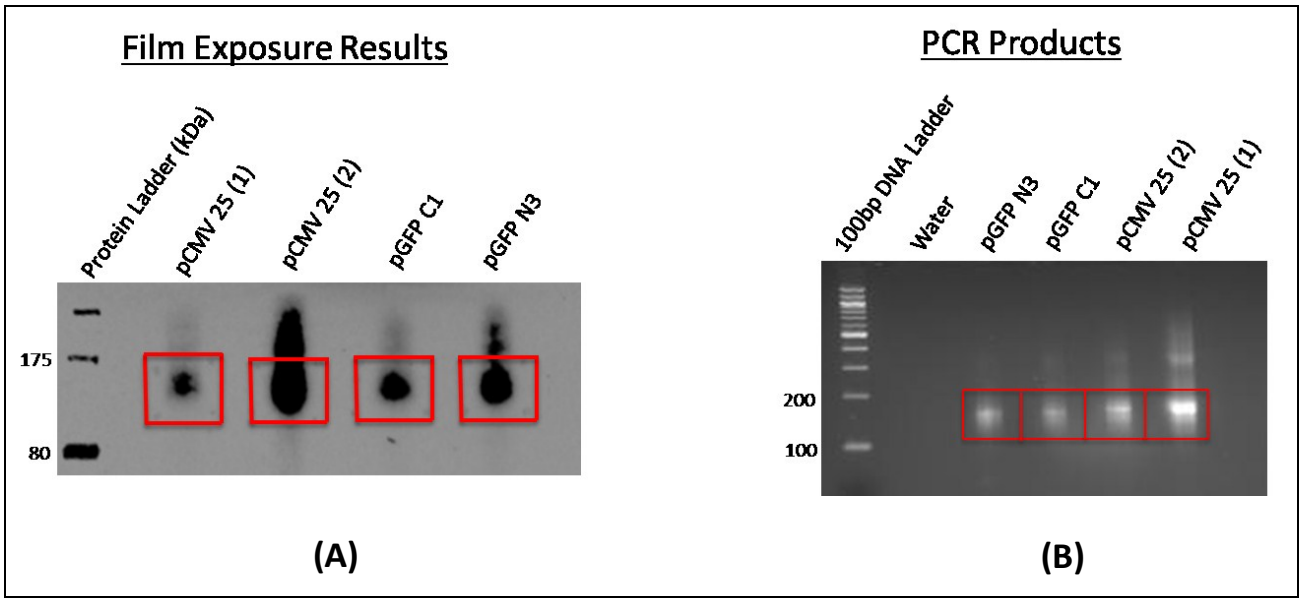


Fig 3.16: Developed film from membrane exposure and the proceeding PCR gel results.

The membrane was exposed to the film for an hour and the results from the developed film are shown in (A). Products were present on the membrane in different quantities. The gel picture (B) indicates the presence of products in the samples from purified PCR products. A prominent band was present at approximately 150 bp, as well as the characteristic smearing that is a result of the presence of products of different lengths were clear on the gel results. Sections marked with red boxes were excised from the gel and products were extracted. These were TOPO cloned and cDNA from six individual colonies from each sample were picked, cultured and mini prep purified for sequencing. The analysed sequencing results are summarised in table 3.5.

Table 3.5 Small sequencing results.

CRAC 6				
Sample	**Sequence**	**Seq length**	**%Match**	**Blast Result**
CMV25 (1C)	CAAAGTGCTCACAGTGCAGGTAG	23	22/23 (96%)	microRNA 20b (MIR20B), microRNA 17 (MIR17), miR-17-92 cluster host gene
CMV25 (2C)	5' Linker Barcode not detected			
CMV25 (3C)	CAAAGTGCtcacaGTGCAGGTAG	23	22/23 (96%)	microRNA 20b (MIR20B), microRNA 17 (MIR17), miR-17-92 cluster host gene
CMV25 (4C)	TAGCAGCACGTAAATTGG	18	15/18 (83%)	microRNA 16-1 (MIR16-1), microRNA 16-2 (MIR16-2)
CMV25 (5C)	TAGATCCGAATTTGT	15	15/15 (100%)	microRNA 10a (MIR10A)
CMV25 (6C)	TATTGCACTtgtCCCGGCctg	21	21/21 (100%)	microRNA 92a-2 (MIR92A2), microRNA 92a-1 (MIR92A1), miR-17-92 cluster host gene
CMV25 (1D)	AACCGCTCAgTGGCTTGGACC	21	15/21 (76%)	inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2)
CMV25 (2D)	AACCGCTcagTGGCTCGGAC	20	15/20 (80%)	inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2)
CMV25 (3D)	AACCGCTCAGTGGCTCGGACC	21	15/21 (76%)	inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2)
CMV25 (4D)	TAAGGTGCATCTAGTGCAG	19	19/19 (100%)	microRNA 18b (MIR18B), microRNA 18a (MIR18A), miR-17-92 cluster host gene
CMV25 (5D)	AACCGCTCAGTGGCTCGGA	19	15/19 (78%)	inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2)
CMV25 (6D)	AACCGCTCAGTGGCTCGGAC	20	15/20 (80%)	inositol 1,4,5-trisphosphate receptor, type 2 (ITPR2)
pGFP C1 (1)	ACAGGgtAGAACCACGGA	18	100% (18/18)	microRNA 140 (MIR140)
pGFP C1 (2)	No Insert detected	0	N/A	N/A
pGFP C1 (3)	No sequence result	N/A	N/A	N/A
pGFP C1 (4)	TTtgGCAATGGTAGAACTCACAct	24	100% (24/24)	microRNA 182 (MIR182)
pGFP C1 (5)	AGtGGTTAG	9	N/A	Sequence too short
pGFP C1 (6)	TACCCTGTAGATCCGAATTGT	21	100% (21/21)	microRNA 10a (MIR10a)
pGFP N3 (1A)	GTCCGAACGCTAGGTCGGTTC	21	N/A	No sequence homology

pGFP N3 (2A)	ACGCGGGAGaCCGGGGTT	18		Homo sapiens nuclear receptor subfamily 0, group B, member 1 Homo sapiens oligonucleotide/oligosaccharide-binding fold containing 1 (OBFC1), mRNA Homo sapiens HtrA serine peptidase 3 (HTRA3), mRNA Homo sapiens PDZ domain containing 1 pseudogene 1 (PDZK1P1)
pGFP N3 (3A)	AGCTACATTGTCTGCTGGGTT	21	21/21 (100%)	(NR0B1), mRNA
pGFP N3 (4A)	AGtgGCTCGGACCG	14	13/14 (93%)	Homo sapiens DC-STAMP domain containing 2 (DCST2)
pGFP N3 (5A)	GGGgAgCCTC	10	N/A	Sequence too short
pGFP N3 (6A)	No Insert detected	N/A	N/A	N/A

Other than four samples, where either no insert or linker were detected or sequencing failed, the insert sequence lengths were consistent with results from the previous CRAC assay attempt. A BLAST search returned results with both miRNAs and transcripts present.

The averages of the insert sequence lengths from all CRAC attempts were calculated and are detailed in table 3.6. The sequence lengths from each assay were also plotted (figure 3.17) for a clear visualisation of their ranges.

Table 3.6: Summary of CRAC Assay results and conditions used.

CRAC Attempt	**Conditions used**		**Average Sequence Length**
	RNase Dilution factor	Reaction time (mins)	
1	10	5	17.2
2	20	5	19.4
3	160	5	18.8
4	80	3	20.14
5	80	3	18.8

The averages of the sequence lengths indicate that a reduction in the reaction would potentially increase the sequence lengths. A more effective and comparative way of studying these results is illustrated in figure 3.17.

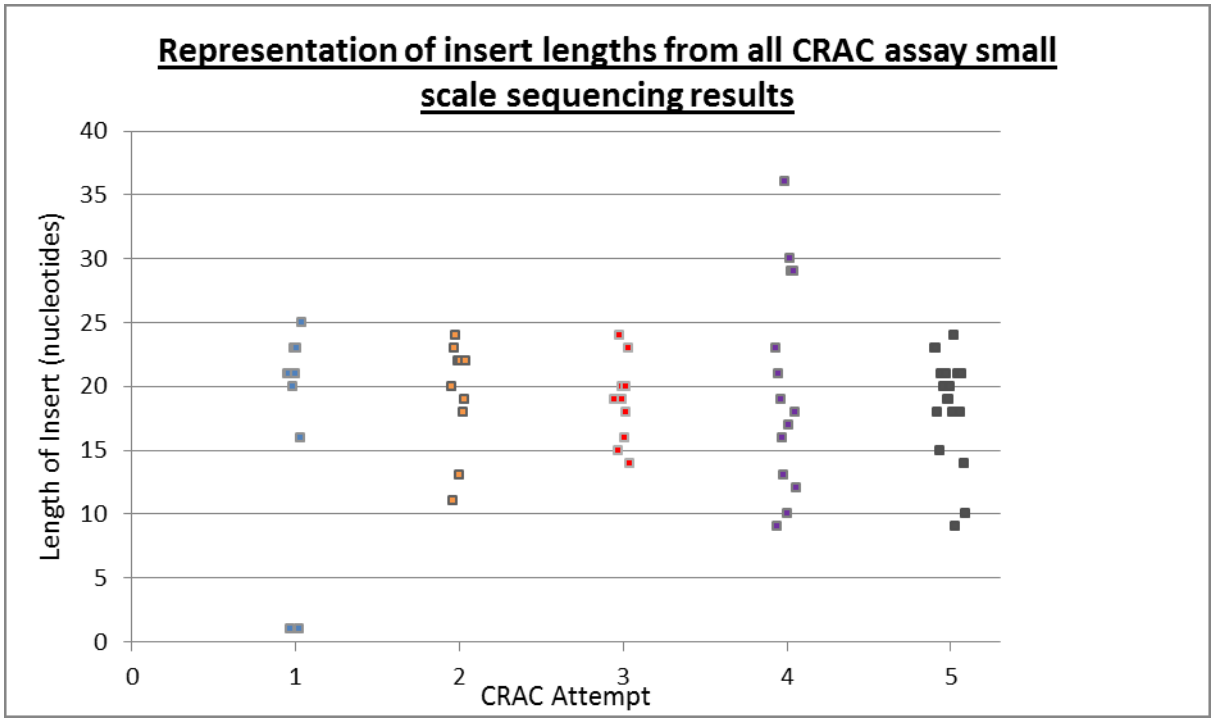


Figure 3.17: CRAC Assay sequence lengths.

The gel and sequencing results collectively show that the CRAC assays worked to some extent. The results are encouraging as longer sequences specifically from CRAC attempt 4 were achieved. Also, the detection of transcripts, miRNAs as well as HCMV miRNAs initially transfected into the cells is indicative of the functionality of the assays. The four samples generated in CRAC attempt 6 were submitted for deep sequencing analysis. Preliminary results from this analysis are discussed in the conclusion section.

CHAPTER 4: Conclusions

The identification of HCMV miRNA targets will provide greater understanding and knowledge of the biology of herpes viruses. Previous studies have demonstrated that herpes virus miRNAs' targets both host and viral transcripts. Identified targets include transcripts involved in apoptosis, innate immunity, cell survival and proliferation as well as mimicking cellular miRNAs with resulting effects favouring the biological life cycle of the viruses. The viral MIE trans-activating region has been shown to be targeted by the viral miRNAs, for example, targeting of the HCMV IE72 gene by the miR-UL112-1. With the majority of the miRNA targets not identified, questions remain as to what their targets are, why those transcripts would be targeted and at what stage of the viral biological life cycle these miRNAs control target genes? The CRAC technique will improve on previous techniques and offer an extensive and valuable platform for these studies. The CRAC assay improves the RISC-IP assays, specifically by reducing background noise by the provision of a 2 step IP and the high stringency conditions in which the assay can be conducted due to UV crosslinking. Also, target transcripts can be identified together with the targeting miRNA due to the intra-molecular ligation reaction. miRNA targets on both viral and cellular genomes can also be investigated simultaneously with the use of deep sequencing which offers an in-depth analysis on a large scale as opposed to microarray which is dependent on the pre-designed probes.

The extent of success of the CRAC assays from these studies will be revealed by the deep sequencing results that are currently under analysis. However, as illustrated in this report's detailed results, we investigated and established the best conditions for transfection of miRNA expression plasmids and induction of tagged Argonaute for the production of quality CRAC lysate samples. We also demonstrated that antibody conjugation of Dynabeads was efficient and capable of pulling down target complexes, a fact supported by the small sequencing result analysis.

To some extent, the small sequencing results strongly indicate the success of the CRAC assay protocol on the whole. The incorporation of the 5' and 3' linkers was confirmed by the small scale sequencing. The presence of cellular transcripts and miRNAs as well as transfected

plasmid encoded miRNAs in some of the assay results was immensely encouraging. It would have been interesting if we had managed to establish what exactly the longer inserts of 30 and 36 nucleotides from CRAC attempt 5 were composed of, or achieved longer sequences with a miRNA and target transcript together indicating that the intra-molecular ligation was achieved. Deep sequencing results will elucidate the success of this step as it was anticipated that this would be a rare event and small sequencing results would most likely miss the resulting miRNA/mRNA hybrid reads.

We investigated the concentration of RNase IT and the digestion reaction time to establish how these factors would affect the insert lengths. Although, admittedly from the present results, RNase IT digestion effects arguably may not be the cause of smaller inserts obtained so far, deep sequencing results will be more informative. It will be interesting to investigate the effect of 1:160 RNase IT dilution with the digestion reaction set up for 3 mins. Also, the recommended amounts of RNase IT, 0.5 units, can be used but altering the time of reaction to a longer time period and at a very low temperature, for example $4^{o}C$, profoundly slowing down the activity of the enzyme.

In these studies, we have demonstrated the immense promise CRAC has in offering a cutting edge technology for studying viral miRNA targets. Although at a preliminary stage of analysis, initial results from deep sequencing indicate that over 100 million reads were captured between the four samples submitted. Of those, approximately 41 million reads contain 5' and 3' linker sequences with identifiable bar codes. Just less than 10 million reads correspond to miR-US25 miRNAs, indicating that the viral miRNAs were successfully incorporated into cellular RISC. Further analysis is on-going to determine insert lengths, composition of inserts and whether specific populations of sequence tags correspond to the samples transfected with the viral miRNAs. However, these results are encouraging and once this technique has been established it will be implemented in studying HCMV miRNAs

Future studies will be aiming to take advantage of the deep sequencing and its ability to provide high speed and throughput results. Therefore assays will be conducted on HCMV infected cells enabling the study of all HCMV encoded miRNA functions, through identifying their targets, as well as potentially discovering novel miRNAs.

CHAPTER 5 - References

5.1 Bibliography

http://www.youtube.com/watch?v=l99aKKHcxC4&feature=related [Accessed 29/08/2012 2012].

ANDERS, D. G., KACICA, M. A., PARI, G. & PUNTURIERI, S. M. (1992). Boundaries and structure of human cytomegalovirus oriLyt, a complex origin for lytic-phase DNA replication. *J Virol,* **66,** 3373-84.

ANSORGE, W. J. (2009). Next-generation DNA sequencing techniques. *New Biotechnology,* **25,** 195-203.

ARASE, H. & LANIER, L. L. (2004). Specific recognition of virus-infected cells by paired NK receptors. *Reviews in Medical Virology,* **14,** 83-93.

BARTEL, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell,* **116,** 281-97.

BECHTEL, J. T. & SHENK, T. (2002). Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. *J Virol,* **76,** 1043-50.

BEGO, M., MACIEJEWSKI, J., KHAIBOULLINA, S., PARI, G. & ST JEOR, S. (2005). Characterization of an antisense transcript spanning the UL81-82 locus of human cytomegalovirus. *Journal of Virology,* **79,** 11022-11034.

BHATT, K., MI, Q. S. & DONG, Z. (2011). microRNAs in kidneys: biogenesis, regulation, and pathophysiological roles. *American Journal of Physiology-Renal Physiology,* **300,** F602-F610.

BODAGHI, B., SLOBBE-VAN DRUNEN, M. E. P., TOPILKO, A., PERRET, E., VOSSEN, R. C. R. M., VAN DAM-MIERAS, M. C. E., ZIPETO, D., VIRELIZIER, J. L., LEHOANG, P., BRUGGEMAN, C. A. & MICHELSON, S. (1999). Entry of human cytomegalovirus into retinal pigment epithelial and endothelial cells by endocytosis. *Investigative Ophthalmology & Visual Science,* **40,** 2598-2607.

BOEHME, K. W., SINGH, J., PERRY, S. T. & COMPTON, T. (2004). Human cytomegalovirus elicits a coordinated cellular antiviral response via envelope glycoprotein B. *Journal of Virology,* **78,** 1202-1211.

BUDOWSKY, E. I., SIMUKOVA, N. A., TURCHINSKY, M. F., BONI, I. V. & SKOBLOV, Y. M. (1976). Induced Formation of Covalent Bonds between Nucleoprotein Components .5. Uv or Bisulfite Induced Polynucleotide-Protein Crosslinkage in Bacteriophage-Ms2. *Nucleic Acids Research,* **3,** 261-276.

CANTRELL, S. R. & BRESNAHAN, W. A. (2005). Interaction between the human cytomegalovirus UL82 gene product (pp71) and hDaxx regulates immediate-early gene expression and viral replication. *J Virol,* **79,** 7792-802.

CHI, S. W., ZANG, J. B., MELE, A. & DARNELL, R. B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. *Nature,* **460,** 479-86.

CORDEN, J. L. (2010). Shining a New Light on RNA-Protein Interactions. *Chemistry & Biology,* **17,** 316-318.

DAVISON, A. J., DOLAN, A., AKTER, P., ADDISON, C., DARGAN, D. J., ALCENDOR, D. J., MCGEOCH, D. J. & HAYWARD, G. S. (2003). The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *Journal of General Virology,* **84,** 17-28.

DOLKEN, L., MALTERER, G., ERHARD, F., KOTHE, S., FRIEDEL, C. C., SUFFERT, G., MARCINOWSKI, L., MOTSCH, N., BARTH, S., BEITZINGER, M., LIEBER, D., BAILER, S. M., HOFFMANN, R., RUZSICS, Z., KREMMER, E., PFEFFER, S., ZIMMER, R., KOSZINOWSKI, U. H., GRASSER, F., MEISTER, G. & HAAS, J. (2010). Systematic Analysis of Viral and Cellular MicroRNA Targets in Cells Latently Infected with Human gamma-Herpesviruses by RISC Immunoprecipitation Assay. *Cell Host & Microbe,* **7,** 324-334.

DUNN, W., TRANG, P., ZHONG, Q., YANG, E., VAN BELLE, C. & LIU, F. (2005). Human cytomegalovirus expresses novel microRNAs during productive viral infection. *Cell Microbiol,* **7,** 1684-95.

DWORSKY, M., YOW, M., STAGNO, S., PASS, R. F. & ALFORD, C. (1983). Cytomegalovirus infection of breast milk and transmission in infancy. *Pediatrics,* **72,** 295-9.

FIELDS, B. N., KNIPE, D. M. & HOWLEY, P. M. 2007. *Fields virology,* Philadelphia, Wolters Kluwer Health/Lippincott Williams & Wilkins.

GANDHI, M. K. & KHANNA, R. (2004). Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *Lancet Infect Dis,* **4,** 725-38.

GOODRUM, F., CAVINESS, K. & ZAGALLO, P. (2012). Human cytomegalovirus persistence. *Cell Microbiol,* **14,** 644-55.

GOTTWEIN, E., CORCORAN, D. L., MUKHERJEE, N., SKALSKY, R. L., HAFNER, M., NUSBAUM, J. D., SHAMULAILATPAM, P., LOVE, C. L., DAVE, S. S., TUSCHL, T., OHLER, U. & CULLEN, B. R. (2011). Viral MicroRNA Targetome of KSHV-Infected Primary Effusion Lymphoma Cell Lines. *Cell Host & Microbe,* **10,** 515-526.

GREY, F., ANTONIEWICZ, A., ALLEN, E., SAUGSTAD, J., MCSHEA, A., CARRINGTON, J. C. & NELSON, J. (2005). Identification and characterization of human cytomegalovirus-encoded microRNAs. *J Virol,* **79,** 12095-9.

GREY, F., MEYERS, H., WHITE, E. A., SPECTOR, D. H. & NELSON, J. (2007). A human cytomegalovirus-encoded microRNA regulates expression of multiple viral genes involved in replication. *PLoS Pathog,* **3,** e163.

GREY, F., TIRABASSI, R., MEYERS, H., WU, G. M., MCWEENEY, S., HOOK, L. & NELSON, J. A. (2010). A Viral microRNA Down-Regulates Multiple Cell Cycle Genes through mRNA 5 ' UTRs. *Plos Pathogens,* **6**.

GRUNDHOFF, A. & SULLIVAN, C. S. (2011). Virus-encoded microRNAs. *Virology,* **411,** 325-43.

HAFNER, M., LANDTHALER, M., BURGER, L., KHORSHID, M., HAUSSER, J., BERNINGER, P., ROTHBALLER, A., ASCANO, M., JUNGKAMP, A. C., MUNSCHAUER, M., ULRICH, A., WARDLE, G. S., DEWELL, S., ZAVOLAN, M. & TUSCHL, T. (2010). Transcriptome-wide Identification of RNA-Binding Protein and MicroRNA Target Sites by PAR-CLIP. *Cell,* **141,** 129-141.

HOCK, J. & MEISTER, G. (2008). The Argonaute protein family. *Genome Biol,* **9,** 210.

JOHNSON, R. A., WANG, X., MA, X. L., HUONG, S. M. & HUANG, E. S. (2001). Human cytomegalovirus up-regulates the phosphatidylinositol 3-kinase (PI3-K) pathway: inhibition of PI3-K activity inhibits viral replication and virus-induced signaling. *Journal of Virology,* **75,** 6022-32.

KALEJTA, R. F. (2008). Functions of human cytomegalovirus tegument proteins prior to immediate early gene expression. *Curr Top Microbiol Immunol,* **325,** 101-15.

KARGINOV, F. V., CONACO, C., XUAN, Z., SCHMIDT, B. H., PARKER, J. S., MANDEL, G. & HANNON, G. J. (2007). A biochemical approach to identifying microRNA targets. *Proceedings of the National Academy of Sciences of the United States of America,* **104,** 19291-19296.

KIM, V. N., HAN, J. & SIOMI, M. C. (2009). Biogenesis of small RNAs in animals. *Nat Rev Mol Cell Biol,* **10,** 126-39.

KINDT, T. J., GOLDSBY, R. A., OSBORNE, B. A. & KUBY, J. 2007. *Kuby immunology,* New York, W.H. Freeman.

KLUPP, B. G., FUCHS, W., GRANZOW, H., NIXDORF, R. & METTENLEITER, T. C. (2002). Pseudorabies virus UL36 tegument protein physically interacts with the UL37 protein. *J Virol,* **76,** 3065-71.

LEE, R. C., FEINBAUM, R. L. & AMBROS, V. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. *Cell,* **75,** 843-54.

LEE, Y., KIM, M., HAN, J., YEOM, K. H., LEE, S., BAEK, S. H. & KIM, V. N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J,* **23,** 4051-60.

LIM, L. P., LAU, N. C., GARRETT-ENGELE, P., GRIMSON, A., SCHELTER, J. M., CASTLE, J., BARTEL, D. P., LINSLEY, P. S. & JOHNSON, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature,* **433,** 769-773.

LUSCOMBE, N. M., GREENBAUM, D. & GERSTEIN, M. (2001). What is bioinformatics? A proposed definition and overview of the field. *Methods Inf Med,* **40,** 346-58.

MENDELSON, M., MONARD, S., SISSONS, P. & SINCLAIR, J. (1996). Detection of endogenous human cytomegalovirus in CD34+ bone marrow progenitors. *J Gen Virol,* **77 (Pt 12),** 3099-102.

MESHESHA, M. K., VEKSLER-LUBLINSKY, I., ISAKOV, O., REICHENSTEIN, I., SHOMRON, N., KEDEM, K., ZIV-UKELSON, M., BENTWICH, Z. & AVNI, Y. S. (2012). The microRNA Transcriptome of Human Cytomegalovirus (HCMV). *Open Virol J,* **6,** 38-48.

MILLER, D. M., ZHANG, Y. X., RAHILL, B. M., KAZOR, K., ROFAGHA, S., ECKEL, J. J. & SEDMAK, D. D. (2000). Human cytomegalovirus blocks interferon-gamma stimulated up-regulation of major histocompatibility complex class I expression and the class I antigen processing machinery. *Transplantation,* **69,** 687-690.

OGAWA-GOTO, K., TANAKA, K., GIBSON, W., MORIISHI, E., MIURA, Y., KURATA, T., IRIE, S. & SATA, T. (2003). Microtubule network facilitates nuclear targeting of human cytomegalovirus capsid. *J Virol,* **77,** 8541-7.

PARI, G. S. (2008). Nuts and bolts of human cytomegalovirus lytic DNA replication. *Human Cytomegalovirus,* **325,** 153-166.

PARK, M. Y., KIM, Y. E., SEO, M. R., LEE, J. R., LEE, C. H. & AHN, J. H. (2006). Interactions among four proteins encoded by the human cytomegalovirus UL112-113 region regulate their intranuclear targeting and the recruitment of UL44 to prereplication foci. *Journal of Virology,* **80,** 2718-2727.

PFEFFER, S., SEWER, A., LAGOS-QUINTANA, M., SHERIDAN, R., SANDER, C., GRASSER, F. A., VAN DYK, L. F., HO, C. K., SHUMAN, S., CHIEN, M. C., RUSSO, J. J., JU, J. Y., RANDALL, G., LINDENBACH, B. D., RICE, C. M., SIMON, V., HO, D. D., ZAVOLAN, M. & TUSCHL, T. (2005). Identification of microRNAs of the herpesvirus family. *Nature Methods,* **2,** 269-276.

PFEFFER, S., ZAVOLAN, M., GRASSER, F. A., CHIEN, M., RUSSO, J. J., JU, J., JOHN, B., ENRIGHT, A. J., MARKS, D., SANDER, C. & TUSCHL, T. (2004). Identification of virus-encoded microRNAs. *Science,* **304,** 734-6.

PLESKOFF, O., CASAROSA, P., VERNEUIL, L., AINOUN, F., BEISSER, P., SMIT, M., LEURS, R., SCHNEIDER, P., MICHELSON, S. & AMEISEN, J. C. (2005). The human cytomegalovirus-encoded chemokine receptor US28 induces caspase-dependent apoptosis. *FEBS J,* **272,** 4163-77.

PRICHARD, M. N., JAIRATH, S., PENFOLD, M. E., ST JEOR, S., BOHLMAN, M. C. & PARI, G. S. (1998). Identification of persistent RNA-DNA hybrid structures within the origin of replication of human cytomegalovirus. *J Virol,* **72,** 6997-7004.

REEVES, M. B., MACARY, P. A., LEHNER, P. J., SISSONS, J. G. & SINCLAIR, J. H. (2005). Latency, chromatin remodeling, and reactivation of human cytomegalovirus in the dendritic cells of healthy carriers. *Proc Natl Acad Sci U S A,* **102,** 4140-5.

RILEY, K. J., RABINOWITZ, G. S., YARIO, T. A., LUNA, J. M., DARNELL, R. B. & STEITZ, J. A. (2012). EBV and human microRNAs co-target oncogenic and apoptotic viral and human genes during latency. *Embo Journal,* **31,** 2207-2221.

ROSSINI, G., CERBONI, C., SANTONI, A., LANDINI, M. P., LANDOLFO, S., GATTI, D., GRIBAUDO, G. & VARANI, S. (2012). Interplay between Human Cytomegalovirus and Intrinsic/Innate Host Responses: A Complex Bidirectional Relationship. *Mediators of Inflammation*.

RYCKMAN, B. J., RAINISH, B. L., CHASE, M. C., BORTON, J. A., NELSON, J. A., JARVIS, M. A. & JOHNSON, D. C. (2008). Characterization of the human cytomegalovirus gH/gL/UL128-131 complex that mediates entry into epithelial and endothelial cells. *Journal of Virology,* **82,** 60-70.

SAMOLS, M. A., SKALSKY, R. L., MALDONADO, A. M., RIVA, A., LOPEZ, M. C., BAKER, H. V. & RENNE, R. (2007). Identification of cellular genes targeted by KSHV-encoded microRNAs. *Plos Pathogens,* **3,** 611-618.

SARISKY, R. T. & HAYWARD, G. S. (1996). Evidence that the UL84 gene product of human cytomegalovirus is essential for promoting oriLyt-dependent DNA replication and formation of replication compartments in cotransfection assays. *Journal of Virology,* **70,** 7398-7413.

SINCLAIR, J. & SISSONS, P. (2006). Latency and reactivation of human cytomegalovirus. *J Gen Virol,* **87,** 1763-79.

SINZGER, C. & JAHN, G. (1996). Human cytomegalovirus cell tropism and pathogenesis. *Intervirology,* **39,** 302-19.

SKALSKY, R. L., CORCORAN, D. L., GOTTWEIN, E., FRANK, C. L., KANG, D., HAFNER, M., NUSBAUM, J. D., FEEDERLE, R., DELECLUSE, H. J., LUFTIG, M. A., TUSCHL, T., OHLER, U. & CULLEN, B. R. (2012). The Viral and Cellular MicroRNA Targetome in Lymphoblastoid Cell Lines. *Plos Pathogens,* **8**.

STARK, T. J., ARNOLD, J. D., SPECTOR, D. H. & YEO, G. W. (2012). High-Resolution Profiling and Analysis of Viral and Host Small RNAs during Human Cytomegalovirus Infection. *Journal of Virology,* **86,** 226-235.

STERN-GINOSSAR, N., ELEFANT, N., ZIMMERMANN, A., WOLF, D. G., SALEH, N., BITON, M., HORWITZ, E., PROKOCIMER, Z., PRICHARD, M., HAHN, G., GOLDMAN-WOHL, D., GREENFIELD, C., YAGEL, S., HENGEL, H., ALTUVIA, Y., MARGALIT, H. & MANDELBOIM, O. (2007). Host immune system gene targeting by a viral miRNA. *Science,* **317,** 376-81.

STERN-GINOSSAR, N., SALEH, N., GOLDBERG, M. D., PRICHARD, M., WOLF, D. G. & MANDELBOIM, O. (2009). Analysis of Human Cytomegalovirus-Encoded MicroRNA Activity during Infection. *Journal of Virology,* **83,** 10684-10693.

STINSKI, M. F. & PETRIK, D. T. (2008). Functional roles of the human cytomegalovirus essential IE86 protein. *Curr Top Microbiol Immunol,* **325,** 133-52.

SUFFERT, G., MALTERER, G., HAUSSER, J., VIILIAINEN, J., FENDER, A., CONTRANT, M., IVACEVIC, T., BENES, V., GROS, F., VOINNET, O., ZAVOLAN, M., OJALA, P. M., HAAS, J. G. & PFEFFER, S. (2011). Kaposi's Sarcoma Herpesvirus microRNAs Target Caspase 3 and Regulate Apoptosis. *Plos Pathogens,* **7**.

TAVALAI, N., PAPIOR, P., RECHTER, S., LEIS, M. & STAMMINGER, T. (2006). Evidence for a role of the cellular ND10 protein PML in mediating intrinsic immunity against human cytomegalovirus infections. *J Virol,* **80,** 8006-18.

TAYLORWIEDEMAN, J., SISSONS, J. G. P., BORYSIEWICZ, L. K. & SINCLAIR, J. H. (1991). Monocytes Are a Major Site of Persistence of Human Cytomegalovirus in Peripheral-Blood Mononuclear-Cells. *Journal of General Virology,* **72,** 2059-2064.

TUDDENHAM, L., JUNG, J. S., CHANE-WOON-MING, B., DOLKEN, L. & PFEFFER, S. (2012). Small RNA Deep Sequencing Identifies MicroRNAs and Other Small Noncoding RNAs from Human Herpesvirus 6B. *Journal of Virology,* **86,** 1638-1649.

TUDDENHAM, L. & PFEFFER, S. (2011). Roles and regulation of microRNAs in cytomegalovirus infection. *Biochim Biophys Acta,* **1809,** 613-22.

ULE, J., JENSEN, K., MELE, A. & DARNELL, R. B. (2005). CLIP: A method for identifying protein-RNA interaction sites in living cells. *Methods,* **37,** 376-386.

WILBERT, M. L. & YEO, G. W. (2011). Genome-wide approaches in the study of microRNA biology. *Wiley Interdisciplinary Reviews-Systems Biology and Medicine,* **3,** 491-512.

WINTER, J., JUNG, S., KELLER, S., GREGORY, R. I. & DIEDERICHS, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat Cell Biol,* **11,** 228-34.

XU, Y. Y., CEI, S. A., HUETE, A. R., COLLETTI, K. S. & PARI, G. S. (2004). Human cytomegalovirus DNA replication requires transcriptional activation via an IE2- and UL84-responsive bidirectional promoter element within oriLyt. *Journal of Virology,* **78,** 11664-11677.

YI, R., QIN, Y., MACARA, I. G. & CULLEN, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev,* **17,** 3011-6.

ZIEGELBAUER, J. M., SULLIVAN, C. S. & GANEM, D. (2009). Tandem array-based expression screens identify host mRNA targets of virus-encoded microRNAs. *Nature Genetics,* **41,** 130-134.

5.2 Appendix

5.2.1 Buffer Recipes

5.2.1.1 Dynabeads-IgG conjugating Buffers

0.1M Sodium Phosphate Buffer ($NaPO_4$) – pH 7.4

NaH_2PO_4 x H_2O (MW 137.99) 2.62g

Na_2HPO_4 x $2H_2O$ (MW 177.99) 14.42g

Dissolve in distilled water, adjust pH if necessary and adjust to 1 litre. This buffer is for pre-washing beads, do not add any protein, sugar, etc.

3M Ammonium Sulphate (stock solution)
$(NH_4)_2SO_4$ (MW 132.1) 39.6g

Dissolve in 0.1M Sodium Phosphate Buffer (pH 7.4) and adjust to 100ml

Phosphate Buffered Saline (PBS) -pH 7.4
NaH_2PO_4 x H_2O (MW 137.99) 0.26g

Na_2HPO_4 x $2H_2O$ (MW 177.99) 1.44g

NaCl (MW 58.5) 8.78g

Dissolve in 900mL distilled water, adjust pH if necessary and adjust to 1 litre.

PBS + 0.5% Triton X-100
Include 0.5% (w/v) Triton X-100 in 100 mL PBS solution

100mM Glycine HCl pH 2.5

10mM Tris pH 8.8

10% Sodium Azide (NaN3)

5.2.1.2 TBST 100ml

10 x TBS	10ml
Tween 20	0.05ml
H_2O	89.95ml

5.2.1.3 SDS PAGE Gel Recipe (10%)

Stacking Gel	Running Gel	
30% Acrylamide/0.8% Bis	3.3 ml	1ml
1.5M Tris-HCL pH 8.7	2.5ml	833µl
H_2O	4.02ml	6.67ml
10% SDS	100µl	10µl
10% Ammonium Persulphate	100µl	50µl
TEMED	10µl	10µl

5.2.1.4 CRAC Assay Buffers

- In bold are the stock concentrations.
- 2-Mercaptoethanol (2ME) is to be added fresh just before using buffer.

Lysis buffer	50ml	
50mM Tris pH 7.5	2.5ml	**1M**
300 mM NaCl	3ml	**5M**
1% NP-40	5ml	**10%**
1mM EDTA	0.5ml	**0.5M**
10% glycerol	10ml	**50%**
5mM 2ME	25µl	**100%**

Ni-EB 200	10ml	
50 mM Tris-HCl pH 7.8	0.5ml	**1M**
50 mM NaCl	0.1ml	**5M**
200mM Imidazole	2ml	**1M**
0.1M NP-40	0.1ml	**10%**
5mM ME	2.5µl	**100%**

5 x PNK buffer	250ml	
250 mM Tris-HCl pH 7.5	62.5ml	**1M**
50 mM MgCl2 (FW= 95.22)	1.19g	
2.5% NP-40	62.5ml	**10%**
250mM NaCl	12.5ml	**5M**

1 x PNK buffer
Contains 5mM 2ME during washes and
10mM 2ME during the reaction

Ni-WB I buffer	50ml	
50 mM Tris-HCl pH 7.8	2.5ml	**1M**
300 mM NaCl	3ml	**5M**
10mM Imidazole	0,5ml	**1M**
6M Guanidine-HCl	28.66g	
0.1M NP-40	0.5ml	**10%**
5mM ME	25µl	**100%**

Ni-WB II buffer
Ni-WB I without Guanidine-HCl

HS IgG-WB buffer	500ml	
50 mM Tris-HCl pH 7.5	25ml	**1M**
0.8 M NaCl	80ml	**5M**
10 mM MgCl2	10ml	**0.5M**
0.5% NP-40	25ml	**10%**
2.5% glycerol	25ml	
Water	335ml	
5 mM 2ME	250µl	**100%**

LS IgG-WB buffer	500ml	
50 mM Tris-HCl pH 7.5	25ml	**1M**
0.3 M NaCl	30ml	**5M**
5 mM MgCl2	5ml	**0.5M**
0.5% NP-40	25ml	**10%**
2.5% glycerol	25ml	
Water	390ml	
5 mM 2ME	250µl	**100%**

5.2.2 Protocol for Conjugation of Dynabeads with Rabbit IgG

Adapted from Rout Lab Protocol, Last Modified by Karen Wei 3/26/07.

Coupling of Dynabeads with Rabbit IgG in order to produce magnetic beads which are capable of pulling out various Protein A tagged complexes.

Note: We are using rabbit IgG -100mg -from Sigma – catalogue number 15006 Dynabeads are from Invitrogen -Dynabeads M270 Epoxy Day ay One – Preparing the Beads and Conjugation

1) Resuspend entire vial of Dynabeads (2 x 10^{10} beads) in 16ml of 0.1M $NaPO_4$ buffer (pH 7.4).

2) Vortex bottle 30 seconds.

3) Divide bead suspension into four 15ml Falcon tubes (4mL of suspension in each tube).

4) Wash any remaining beads in the glass vial with an additional 2ml of 0.1M NaPO4 buffer. Divide equally amongst the four Falcon tubes.

5) Shake bead suspension slowly for 10 minutes on a Nutator or rocking platform.

6) While bead suspension is on Nutator prepare the AB mix.

a. FIRST: Re-suspend the entire bottle of Rabbit IgG (100mg) in 7ml double distilled H_2O – this will result in a concentration ~14mg/ml. Aliquot into 1mL fractions and store any unused IgG at -20°C.

b. Spin down 3525µl, of Rabbit IgG, in a table top centrifuge, for 10 mins, at 14 000 rpm, at 4°C. Save supernatant and discard pellet.

c. Prepare AB mix by adding the solutions to a 50mL falcon tube, in the order listed.

i) 3525µl of IgG (which was previously spun down) If IgG concentration needs to be altered please see end of protocol for instructions on determining appropriate amounts of Sodium Phosphate and Ammonium Sulphate

ii) 9.850mL 0.1M $NaPO_4$ buffer.

iii) 6.650ml 3M Ammonium sulphate. Add slowly shaking the tube a bit.

iv) Filter solution using a 0.22µm Millex GP filter

7) Place the Falcon tubes with the bead suspension onto a magnetic holder and wait until all beads are attached to the magnet (Dynal MPC-6 Magnetic Particle Concentrator Prod. No.120.02). Bead solution will appear clear. Aspirate the buffer off (be careful not to aspirate off the beads too).

8) Wash again with 4mL 0.1M $NaPO_4$ -incubation for 10 minutes is not necessary. Vortex 15 secs. Put on the Magnet -aspirate off the buffer.

9) Add 5ml of AB mix to each tube vortex to completely combine the AB mix and the beads.

10) Wrap tops of Falcon tubes with parafilm and place on rotating wheel at 30^0C overnight (incubation must last at least 18 hours but no more than 24 hours).

Day Two -Washing the Dynabeads after Conjugation

Do all washes as described in the 15ml Falcon tubes. You can aspirate the supernatant by using a Vacuum Aspirator.

1) Wash once with 3mL of 100mM Glycine HCL pH2.5. Put it on and take it off as fast as possible.

2) Wash once with 3ml of 10mM Tris pH 8.8.

3) Wash once with 3mL of 100mM Triethylamine. (Make fresh 100mM Triethylamine by adding 168µl stock to 11.156ml of DDH_20). Put it on and take it off as fast as possible.

4) Wash the coated beads with 1x PBS for 5 minutes – washes should be done on a rocker/nutator – repeat x4.

5) Wash once with PBS + 0.5% Triton X-100 for 5 minutes.

6) Wash again with PBS +0.5% Triton X-100 for 15 minutes on rocker/nutator.

7) Finally, resuspend all beads in a total of 2ml of 1x PBS + 0.02% Sodium azide.

8) Store the coated beads at 4^0C.

5.2.3 CRAC Protocol

Day 1 (full day ~9hrs)

First IP - IgG – Dynabeads

Argonaute-2 protein, part of the RISC complex, is binds to the IgG conjugated dynabeads via the protein A.

Beads freshly coated – 14301 (invitrogen), 60mg (amount for 3 samples, 20mg each), stored at 4C in PBS + 0.02 % sodium azide

- LOADING - always keep beads on ice!

IgG Dynabeads solution (washed directly before use 3 x with 5ml PBS, beads vol. about 70ul) – this is to get rid of IgG that is in solution. Beads mix then divided into 3 (3 falcon tubes) and extract added.

Incubation for 45 mins, temp = 4^{o}C, rotation

Flow through (SAVE FOR WESTERN BLOT) 30ul + 30ul 2 x SB → 5' 95C *(Western to show no protein as the RISC complex is meant to have been bound to the Dynabeads.)*

- WASHES:

1 x LS-IgG WB buffer (vol. = 5ml) fast
2 x HS-IgG WB buffer (vol. = 5ml) 5 mins cold room - rocker
1 x LS-IgG WB buffer (vol. = 5ml) fast
LS for Low Salt and HS for High Salt and these buffers get rid of non-specific binding!!!

Resuspend in 1ml PNK_5 buffer and transfer in eppendorf, put on the small magnet (invitrogen magnorack CS15000) and remove the PNK buffer and proceed to next step.

RNAse digestion

This step aims to digest the hanging target mRNA on both sides of the miRNA bound part.

- Beads re-suspended in pre-warmed/room temp 500µl PNK_5 buffer + 0.5 unit RNAseA+T1 mix (1ul of 1:20 dilution in H_2O of RNase-IT cocktail kept at 4°C; agilent 400720: Stock has 5000 units/ml and this has been diluted in water at a ratio of 1:10 and use 1µl of that to get the 0.5 units required.)

Time = 5 mins, temp = 20°C, water bath (tap once during 5 min so beads don't settle) or use thermoshaker at 20°C and 500 rpm.
Transfer onto ice for 1' (no longer) to slow down the reaction, discard RNase mix and proceed immediately to the next step.
(should not let RNase work longer and the guanidine of the elution buffer will inactivate it)

- ELUTION

3 x 250µl Ni-WB I buffer 10 mins at 20°C (on rotator or wheel, or thermoshaker), pool the fractions together (750µL final volume per sample)

Second IP - Ni-NTA

Binding of the His tag to the Nickel beads for the second part of the RISC IP. ***His-tag:*** *an amino acid motif made up of 6 histidines.*

Equilibrate the Ni-NTA resin with Ni-WB I (take 80 µl slurry per sample, wash 2X with Ni-WB I) spin 1000 g ~ 20 sec, remove supernatant, add new buffer, etc). ** May help to use a lamp to see beads when removing liquid.

Ni-NTA resin is from Qiagen "Ni-NTA superflow, 25 ml"

- LOADING

Add 50ul Ni-NTA slurry (about 50%) to eluate from the IgG resin

Incubated for 2 hours 4°C, rotation

- WASH:

2 x Ni-WB I buffer, 0.75 ml rocking at 4°C for 10 mins.

Beads transferred to the Pierce columns (Spin column snap cap # 69725)
2 x Ni-WB II buffer, 0.75 ml
3 x PNK_5 buffer 0.75 ml
(When using the Ni-WB I buffer before an enzymatic reaction think of washing it away well to not inhibit the enzyme, try to deposit the volume on the bottom of the column and not on the upper wall)

Phosphorylation (on column)

(Put the little cap at the bottom of the column)

- Beads resuspended in 80µl mix:

	1x	3x
ATP (stock 100mM, final 1mM)	0.8µl	2.4ul
RNAsin (20u/µl stock)	2µl	6ul
T4 PNK (5u/µl stock)	4µl	12ul
5 x PNK_{10} buffer	16µl	48ul
water	57.2µl	171.6ul
TOT		240 ul

(Close the lid of the column) Time= 2h 30 mins, temp = 20 ^{o}C

"On-bead intra-molecular ligation" (on column)

Add 80µl of the mix (making a total of 160µl)

	1x	3x
ATP (stock 100mM, final 1mM)	0.8µl	2.4µl
RNAsin (20u/ul stock)	2µl	6µl
T4 RNA ligase I	4µl	12µl
5 x PNK_{10} buffer	16µl	48µl
water	57.2µl	171.6µl
TOT		240 µl

Time = O/N, 16C, rotating (shaking)

Day 2 (short day ~2hrs)
(Open the lid of the column and then remove the cap)

- WASH:

1 x Ni-WB I buffer, 0.4 ml
1 x Ni-WB II buffer, 0.75 ml
3 x PNK buffer, 0.75 ml
(Ni-WB1 buffer contains guanidine which stops the ligation reaction . The imidazole present at low conc. gets rid of the non-specific binding to the Ni-beads.) The following washes will be to remove any residual NI-WB 1.)

Dephosphorylation (on column) (of the 3' end which will allow the ligation of the 3' linker).
(Close the lid of the column to eliminate the maximum of the wash volume, put back the cap at the bottom and open the lid again)

- Beads resuspended in 80ulµof the reaction mix:

	1x	3x
TSAP	8µl	24µl
RNAsin (20u/µl stock)	2µl	6µl
5 x PNK_{10} buffer	16µl	48µl
Water	54µl	162µl
TOT	80µl	240µl

Time = 45 mins, temp. = 20^{o}C

(Open the lid of the column and then remove the cap)

- WASH:

1 x Ni-WB I buffer, 0.4 ml
1 x Ni-WB II buffer, 0.75 ml
3 x PNK buffer, 0.75 ml

3' linker ligation (on column)

(This a linker with a known sequence which will allow amplification and sequencing)

(Close the lid of the column to eliminate the maximum of the wash volume, put back the cap at the bottom and open the lid again)

- Beads resuspended in 80µl of the reaction mix:

	1x	3x
3' linker (miRCat-33, 10µM) *Kept at -80!*	8µl	24µl
T4 RNA Ligase II, truncated K227Q (M0351)	4µl	12µl
RNAsin (20u/ul stock)	2µl	6µl
PEG 8000 (stock 25%, final 10%)	32µl	96µl
5 x PNK_{10} buffer	16µl	48µl
Water	18µl	54µl
	80µl	240µl

Time = 10-24 hrs, temp = 16^{o}C

Day 3 (very long day ~ 12hr +)

(Open the lid of the column and then remove the cap)

WASH:

1 x Ni-WB I buffer, 0.4 ml
1 x Ni-WB II buffer, 0.75 ml
3 x PNK buffer, 0.75 ml

(Close the lid of the column to eliminate the maximum of the wash volume, put back the cap at the bottom and open the lid again)

Radioactive labelling of RNA (on column)

Addition of a radioactive phosphate at the 5' end of the miRNA/mRNAs

- Beads resuspended in 80µl mix:

	1x	3x
^{32}P-ATP	3µl	9µl
RNAsin (20units/µl stock)	2µl	6µl
T4 PNK (5µl stock)	4µl	12µl
5 x PNK buffer	16µl	48µl
Water	55µl	165µl
	80µl	240 µl

Time= 40 mins, temp = 37^{o}C

!!Open the lid of the column and only then remove the cap!!
!! First and second wash in the radioactive liquid waste!!

- WASH:

1 x Ni-WB I buffer, 0.4 ml
1 x Ni-WB II buffer, 0.75 ml
3 x PNK buffer, 0.75 ml
(close the lid of the column to eliminate the maximum of the wash volume, put back the cap at the bottom and open the lid again)

ELUTION

Elution buffer which contains imidazole at high concentration disrupts the Histag-Ni interaction by competitive binding.

Beads resuspended in the elution buffer (Ni-EB 200mM Imidazole), 2 x 200µl, 1x 600µl, 5 mins incubation shaking during each elution step (used thermoshaker). Pull the eluates.

TCA precipitation

Precipitation of proteins and brings the volume down to enable loading onto the gel.
1ml of eluate + 2µg BSA + 200µl 100% TCA 30' on ice**.**
Centrifuged full speed for 20 mins, 4 ^{o}C

!! TCA supernatant in the TCA radioactive waste!!

Washed 2x with 1 ml cold acetone (-20 ^{o}C), vortexed in between, 10 mins spin

!! Acetone washes in the TCA radioactive waste!!

Pellet dried at 37 ^{o}C, re-suspended in 10µl water, then 10µl 2 x SB (NuPage) added, 10 mins at 65 ^{o}C.
(Do not over-dry the pellet after TCA precipitation otherwise it becomes very difficult to re-suspend; do not pipette, do not touch the pellet with the tip)

Gel- using precast:

NuPage 4-12% gel (10 well; product NP0335BOX, Invitrogen)
NuPage MOPS SDS Running buffer, 20X (NP000102)
NuPage Transfer Buffer (NP00061)

Load samples along pre-stained ladder, leave space in between samples!!!!
Run ~ 40 min to 1 hour at 150 V.
A band of approximately 100kd of Ago 2 + Histag +mRNA/miRNA-3' linker is expected.
Transfer to nitrocellulose in transfer buffer (on ice) for 2 hours at 100 V.

Expose to film **with a fluorescent ruler**, 30' and 2h if not strong enough.

Note: At this point it is possible to store the membrane at -80 overnight and continue on Day 4 although this isn't ideal

Put the film over the membrane and with a needle pierce film and membrane in correspondence to the band to cut. Then use the holes on the membrane as a reference to cut the band with a blade, transfer in an eppendorf.

Proteinase K treatment

Proteinase K digests the proteins which will be bound to the membrane and therefore releases the miRNA.

- Bands cut out of the membrane and incubated for 2 hrs, at 55°C in the 400µl proteinase K mix.

Proteinase K mix:

	50ml	
50mM Tris-HCl	2.5 ml	1M
50mM NaCl	0.5ml	5M
10mM imidazole	0.5ml	1M
0.1% NP-40	0.5ml	1M
1% SDS	5ml	10%
5mM EDTA	0,5ml	0,5M
5mM 2ME)	add fresh before use	

+ 100µg proteinase K per sample

PCI (phenol chloroform isoamyl alcohol) extraction and EtOH precipitation

- To the 400µl Proteinase K treated mix, add these separately: +50µl 3M NaAc pH 5.5 + 500µl PCI and centrifuge at max speed and 4°C for 30 minutes and carefully collect the upper phase.
  RNA will stay in the aqueous phase, the upper phase, and proteins will be in the lower/organic phase.

- Added: 20µg glycogen
  1ml 100% EtOH
  Incubate for 30 mins at -80°C

- Spin max speed in cold centrifuge, 30 mins
- 2 washes with 70% EtOH, air-dry

(If needed you can interrupt here leaving the precipitation at -80°C)

Phosphorylation & 5' Linker Ligation

- Pellet resuspended in 12.5ul water
  1.5 ul 10 x RNA ligase I buffer (with 1mM ATP)
  1ul PNK

Incubate - 30', 37C

- 1µl 5' linker - remember to use a different primer for each sample i.e. for distinguishing them during deep sequencing results analysis.
  0.5µl 10 x RNA ligase I buffer
  1µl RNA ligase I
  2.5µl water
  Total vol = 20µl

Incubate – ON, 16^{o}C.

Day 4/5 (full day ~9hrs)

PCI extraction and EtOH precipitation

- Reaction mixture + 360µl water + 40µl 3M NaAc pH 5.5 + 500µl PCI → upper phase collected
- Added: 1µl glycogen (20µg)
  450µl EtOH

  Incubate at -20C, for a few hours.

- Spin max speed in cold centrifuge, 30 mins.
- 2 washes with 70% EtOH, air-dry

Reverse transcription:

- Pellets resuspended in 13µl of the mix:
  4µl dNTPs [2.5mM]
  1µl of 3' miRCat RT (10µM) (DNA primer which is reverse compliment to the 3' linker)
  8µl water

- Incubate – 3 mins, 80^{o}C
- Incubate - 5', on ice

- Add 6µl of the mix: 4µl First strand buffer
  1µl 0.1M DTT
  1µl RNAsin
  Incubate – 3 mins, 50^{o}C

- Add 1µul Superscript III, incubate - 1hr, 50^{o}C
- Inactivate Superscript – 15mins , 65^{o}C
- Degrade RNA – add 2µl RNAse H, incubate 30 mins, 37^{o}C
- Store at -20^{o}C

PCR
Set 4 PCR reactions for each sample with 2.5µl of input per reaction (10µl of cDNA from the previous step, half of the total)

PCR mix
Tot vol = 50µl for each reaction

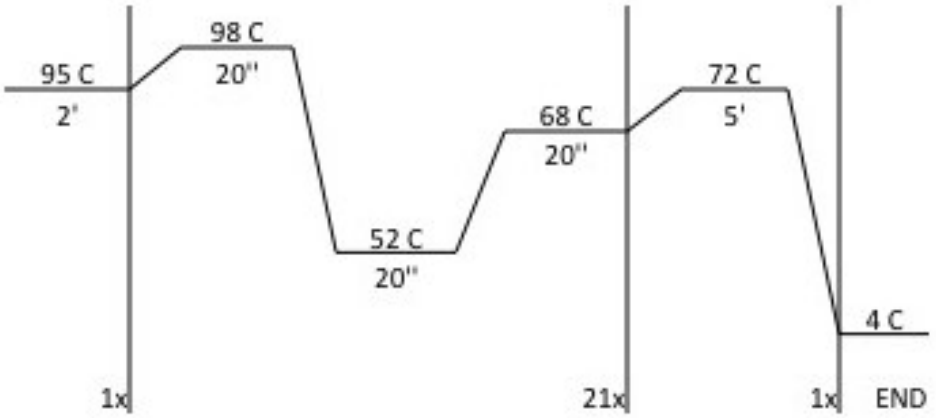
H_2O	35µl
buffer 10X	5µl
dNTP 2.5mM each	5µl
P5 primer 10uM	1µl
PE miRCat primer 10uM	1µl
Takara Taq	0.5µl

Input 2.5µl cDNA.
Master mix for each sample 190µl, add 10µl of cDNA input and distribute in 4 PCR tubes.

Control PCR – miR16 amplification

PCR 1 control		**PCR 2 control**	
Tot vol = 20µl		Tot vol = 20µl	
H_2O	14µl	H_2O	14 µl
Buffer 10X	2µl	Buffer 10X	2 µl
dNTP 2.5mM each	2µl	dNTP 2.5mM each	2 µl
P5 primer 10uM	0.4µl	PE miRCat primer 10uM	0.4 µl
miR-16 Rev 10uM	0.4µl	miR-16 F 10uM	0.4 µl
Takara Taq	0.2µl	Takara Taq	0.2 µl
Input 1µl cDNA		Input 1µl cDNA	

PCR manual



After PCR pool the 4 reactions together and purify with Qiagen QIAquick columns.
USE MINIELUTE COLUMNS! (stored at 4oC)
Elute in 15µl of EB buffer add 3.5µl of loading 6x

Prepare 3% Metaphor agarose gel:
Weigh the agarose, add TBE 1x and leave on ice for 20 min (to avoid foaming). Microwave to dissolve, add syber safe (1:10000, Invitrogen 533102). Let it solidify then transfer to 4oC for 20 mins before running.

Run the gel with the tank on ice. 80 Volts for 1.5-2 hours

Check gel at phosphoimager resolution 100µm. IMPORTANT: Print at 100% the image of the gel including margins and wells.

Draw on printout the square you will cut and put in the order from bottom to top: printout, transparent film (the one for slides) and the gel on the top. Align the margins of the gel and the wells then cut the bands. At this point the bands from different samples can be pooled.

Samples can be purified with Qiaquick gel extraction.
NOTE: dissolve agarose at 37C and use MINIELUTE columns.

Elute in 10µl H_2O per column. This is the final product.
NOTE: Read concentration with the fluorescence method, more accurate.

Cloning for small scale sequencing

Ligation with TA kit Invitrogen

Buffer 1µl
Vector 1µl
H_20 2-1µl
PCR 1-2µl

Leave 30mins at RT and then transform in Top10 one shot
1µl of the ligation with one aliquot of Top10

30 mins on ice, Heat shock competent cells 30 sec 42C and then 2 mins ice.

Resuspend in pre-warmed SOC and incubate at 37^{o}C for between 30 mins and 1h before plating on an ampicillin plate and incubating at 37^{o}C.

Colonies are individually picked and DNA purified for sequencing using the M13 primer.
