CLOSING THE GENOMIC GAPS OF RAT CYTOMEGALOVIRUS ALL-03 IN GENOME SCAFFOLD

KRISHNAN NAIR A/L BALAKRISHNAN

FPV 2016 5

CLOSING THE GENOMIC GAPS OF RAT CYTOMEGALOVIRUS ALL-03 IN GENOME SCAFFOLD

By

KRISHNAN NAIR A/L BALAKRISHNAN

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

January 2016

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the Degree of Master of Science

CLOSING THE GENOMIC GAPS OF RAT CYTOMEGALOVIRUS ALL-03 IN GENOME SCAFFOLD

By

KRISHNAN NAIR A/L BALAKRISHNAN

January 2016

Chairman : Prof. Dato' Mohd Azmi Mohd Lila, PhD
Faculty : Veterinary Medicine

Cytomegalovirus (CMV) has been known to cause acute, latent and persisting infections in animals and humans. Interestingly, Rat CMV ALL-03 strain (RCMV ALL-03) is unique among all previous RCMVs and Murine CMV strain in crossing the placenta naturally which can be used to study the CMV congenital infection. The partial genome sequence for this strain has been produced through Next Generation Sequencing technology using Illumina platform but the incompleteness limits the comprehensive study of genomics and proteomics. Thus, completing the genome sequence by closing all the gaps in draft genome and verify the position and function of all Open Reading Frame (ORF) can aid in future study for antiviral drug and vaccine development. The study was commenced by propagating RCMV ALL-03 extensively to a large amount in rat embryonic fibroblast (REF) cells. Infected cells exhibiting advance cytopathic effect were harvested and concentrated by 8% (w/v) PEG 6000. Concentrated viruses were then purified and proceeded with DNA extraction by conventional method. Four sets of forward and reverse primer for each gap were carefully designed using CLC Genomic workbench software. Four set of primers for each gap were used in Polymerase Chain Reaction (PCR) for gap verification purpose followed by single band of DNA in agarose gel sequenced and 8 templates of sequences were obtained for each gap. All these templates were compared and verified the original missing sequences hence inserting the actual missing sequences taken place in the previous partial sequence. Result shows that, false gap have been identified and all the misassemblies in genome sequence were corrected, thus producing a final completed genome sequence of RCMV ALL-03. From this study, the final number of base pairs of RCMV ALL-03 was 197,958 and has been arranged as single unique sequence flanked by 504 base pair terminal direct repeats. The overall content of G+C content was 46% and total 123 protein coding genes (CDS) were identified. Out of 123 CDS, 46 CDS (36.2%) were grouped into 8 functional classes such as capsid, glycoprotein, tegument, DNA replication, DNA packaging, nuclear egress, immune

evasion and regulatory. Subsequently, all the positions of ORF were corrected and translation of protein has been verified. Next, phylogenetic analysis of RCMV ALL-03 based on conserved genes of herpesvirus revealed that this Malaysian isolate is closest to RCMV-English and RCMV-Berlin strains with 99% and 97% of homology were identified respectively. Moreover, from the tree, it is observed that the evolutionary relationship of RCMV ALL-03 with other strains of herpesviruses from all the three subfamilies. Interestingly, betaherpesvirus subfamily shown to be more closely related with gammaherpesviruses compare to alphaherpesviruses where beta- and gamma- share some of the functional ORF. Results presented have been submitted to genebank and the accession number was obtained. Further exploration of RCMV ALL-03 in future could provide more valuable information in understanding the pathogenesis of the virus as well as congenital infection which could serve as model to mimic the study of HCMV.

Keywords: herpesvirus, cytomegalovirus, RCMV, genomic gaps, open reading frame

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk Ijazah Master Sains

MENUTUP JURANG JUJUKAN GENOM BAGI ALL-03 SITOMEGALOVIRUS TIKUS DALAM STRUKTUR URUTAN GENOM

Oleh

KRISHNAN NAIR A/L BALAKRISHNAN

Januari 2016

Pengerusi : Prof. Dato' Mohd Azmi Mohd Lila, PhD
Fakulti : Perubatan Veterinar

Jangkitan sitomegalovirus (CMV) dikategorikan sebagai kumpulan betaherpesvirus di bawah sistem klasifikasi herpesvirus dimana ianya boleh menyebabkan penyakit berterusan dan penyakit tanpa symptom. Strain tikus CMV (RCMV) ALL-03 merupakan virus yang ditemui dari plasenta dan rahim tikus di Malaysia. Kelebihan virus jenis ini adalah kebolehan menyebabkan jangkitan kongenital di mana RCMV dan MCMV strain lain tidak mempunyai ciri tersebut. Secara keseluruhan, kajian ini menandakan penghasilan urutan RCMV ALL-03 genom secara keseluruhan dengan menutup semua jurang urutan genom dan pencirian gen-gen mengikuti urutan genom yang dikenalpasti serta menganalisis gen-gen tersebut dengan strain lain kumpulan herpesvirus. Pengajian ini dimulakan dengan membiakkan RCMV ALL-03 di sel fibroblast tikus. Sel-sel yang menunjukkan kesan sitopatologi diambil kira untuk penuaian virus dan virus-virus tersebut dipekatkan dengan 8% (w/v) PEG 6000. Virus yang dipekatkan kemudian ditulenkan dengan proses kecerunan sukrosa. Selepas proses penulenan virus, DNA virus diekstrak dengan menggunakan kaedah konvensional. Empat set primer telah direka khas dengan teliti dengan menggunakan sistem CLC Genomic Workbench. Berlainan primer digunakan untuk berlainan jurang genom melalui proses reaksi rangkaian polimerasi (PCR) dan hasil amplifikasinya diperiksa dengan menggunakan gel agarose. Band tunggal yang diperolehi dihantar untuk pengenalpastian susunan genom dan lapan hasil susunan genom diterima untuk setiap jurang genom. Seterusnya, kesemua hasil pengenalpastian susunan genom telah dianalisis dan susunan genom yang tepat telah dimasukkan dalam jurang genom tertentu. Selain daripada itu, jurang genom palsu dan susunan genom yang salah telah dikenalpasti dan diperbetulkan. Akhirnya, susunan genom RCMV ALL-03 yang muktamad dihasilkan dan penyemakan semula posisi mula dan akhir gen-gen dijalankan diikuti dengan perubahan angka keseluruhan susunan genom virus. Selepas itu, susunan genom RCMV ALL-03 yang muktamad telah dihantar ke 'GenBank' untuk memperolehi nombor pengenalan dimana ianya boleh digunakan untuk pelbagai bidang. Setelah nombor pengenalan diperolehi, gen-gen RCMV ALL-03 yang telah

dikenalpasti dianalisis kegunaanya dan perbandingan dibuat antara RCMV lain serta manusia CMV (HCMV). Analisis filogenetik turut dijalankan untuk mengenalpasti cabang perkaitan antara RCMV ALL-03 dengan virus-virus lain daripada semua kumpulan herpesvirus. Konkulisinya, keseluruhan saiz genom RCMV ALL-03 adalah 197,958 bp disusun secara urutan unik tunggal dengan 504 bp diapit pada ulangan langsung terminal. Kandungan G+C adalah 46% dan 123 gen telah dikenaplasti. Daripada 123 gen, 46 gen (36.2%) telah dikategorikan kepada 8 kumpulan seperti capsid, glikoprotein, tegument, replikasi DNA, pembungkusan DNA, pembebasan nuklear dan strategi pengelakan sistem keimunan. Analisis filogenetik menunjukkan RCMV ALL-03 bercabang rapat dengan RCMV-E dan RCMV-B dengan 99% dan 97% masing-masing.

Sepertimana yang dijangkakan, kumpulan betaherpesvirus berkaitan rapat dengan gammaherpesvirus dari segi evolusi berbanding dengan alfaherpesvius di mana kedua dua beta dan gamma berkongsi fungsi gen-gen tertentu. Pada masa akan datang, RCMV ALL-03 boleh dieksploitasi untuk kajian yang berkaitan dengan pemahaman kepatogenan virus serta jangkitan kongenital di mana RCMV ALL-03 boleh digunakan sebagai model untuk memahami lebih mendalam tentang HCMV.

Kata kata kunci: Herpesvirus, sitomegalovirus, RCMV, urutan genom, jurang genom

ACKNOWLEDGEMENTS

I would never been able to complete my dissertation without the assistance, encouragement, valuable advice and positive comments from many respectable individuals.

First and foremost, I offer my personal gratitude to my supervisor Prof. Dato' Dr. Mohd Azmi Mohd Lila, who supported and guided me throughout my research without any hesitation. I would also like to thank my co-supervisors, Prof Dr. Noordin bin Mohamed Mustapha, Datin Paduka Prof Dr. Aini Ideris and Dr. Farina Mustapha Kamal for their valuable insights and sharing their knowledge and expertise with me.

Completing my studies on time would be remain as dream without the guidance and motivation from my mentor and sister Miss. Ashwaq Ahmed Abdullah where her continuous support and kindness were my inspiration to complete this study. In addition, I am indebted to Dr. Siti Aimi Sarah who helped me a lot in correcting my thesis as well as provide suggestions in presenting my thesis in more order form. My sincere thanks to Dr. Faez Firdaus Jesse Abdullah who motivates and guided me to finish my thesis on time.

For my virology lab members and officers, thank you for making my research more enjoyable.

My thanks and love goes to my family, especially my parents whom provides me the best childhood as well as continue supporting me till date and my brothers, sister and loved ones to encourage me throughout many years in my masters journey.

Finally, I would like to acknowledge the financial support from MyBrain (MyMaster), Graduate Research Fellowship (GRF) and 'RUGS' from Kementerian Pengajian Tinggi and Universiti Putra Malaysia.

I certify that a Thesis Examination Committee has met on 11 January 2016 to conduct the final examination of Krishnan Nair a/l Balakrishnan on his thesis entitled "Closing the Genomic Gaps of Rat Cytomegalovirus ALL-03 in Genome Scaffold" in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of Science.

Members of the Thesis Examination Committee were as follows:

Abdul Aziz bin Saharee, PhD
Professor
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Chairman)

Siti Suri binti Arshad, PhD
Associate Professor
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Internal Examiner)

Imad Ibrahim Al-Sultan, PhD
Professor
Universiti Malaysia Kelantan
Malaysia
(External Examiner)

ZULKARNAIN ZAINAL, PhD
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date: 24 March 2016

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

Mohd Azmi bin Mohd Lila, PhD, MBA, LLM
Professor
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Chairman)

Noordin b Mohamed Mustapha, PhD
Professor
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Member)

Aini Ideris, PhD
Professor
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Member)

Farina Mustapha Kamal, PhD
Lecturer
Faculty of Veterinary Medicine
Universiti Putra Malaysia
(Member)

BUJANG BIN KIM HUAT, PhD
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

Declaration by graduate student

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in theUniversiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: _____ Date: _____

Name and Matric No.: Krishnan Nair A/L Balakrishnan GS 37335

Declaration by Members of Supervisory Committee

This is to confirm that:

- the research conducted and the writing of this thesis was under our supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013)are adhered to.

Signature: _____
Name of Chairman of Supervisory Committee: Mohd Azmi bin Mohd Lila, PhD, MBA, LLM

Signature: _____
Name of Member of Supervisory Committee: Noordin b Mohamed Mustapha, PhD

Signature: _____
Name of Member of Supervisory Committee: Aini Ideris, PhD

Signature: _____
Name of Member of Supervisory Committee: Farina Mustapha Kamal, PhD

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A	Adenine
α	Alpha
ATV	Antibiotic-trypsin-versene
BCMV	Baboon cytomegalovirus
bp	Base pair
BSA	Bovine serum albumin
β	Beta
CO_2	Carbon dioxide
CCMV	Chimpanzee cytomegalovirus
CDS	Coding DNA sequence
CRFK	Crandell Rees feline kidney
CPE	Cytopathic effect
CID	Cytomegalic inclusion disease
CMV	Cytomegalovirus
C	Cytosine
˚C	Degree Celsius
DAB	3, 3'-Diaminobenzidine
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Dinucleotide triphosphate
E	Early
EDTA	Ethylenediaminetetraacetic acid
FBS	Fetal bovine serum
FW	Formula weight
γ	Gamma
gB	Glycoprotein B
g	Gram
G	Guanine
GPCMV	Guinea pig cytomegalovirus
h	Hour
IE	Immediate early
IgG	Immunoglobulin G
IIP	Indirect immunoperoxidase staining
IU/mL	International unit per milliliter
HCMV	Human cytomeglaovirus
kb	Kilo bases
kbp	Kilo base pair
L	Late
L	Liter
pp65	Lower matrix phosphoprotein
$MgCl_2$	Magnesium chloride
µg	Microgram
µg/µL	Microgram per microliter
µg/mL	Microgram per mililiter
µL	Microliter

µm	Micrometer
mM	Millimolar
min	Minute
mL	Milliliter
mAb	Monoclonal antibody
MCMV	Mouse cytomegalovirus
MuHV-1	Muridherpesvirus 1
MuHV-2	Muridherpesvirus 2
MuHV-8	Muridherpesvirus 8
NIEPs	Noninfectious envelope particles
N	Normal
nm	Nanometer
NGS	Next generation sequencing
ORF	Open reading frames
%	Percent
PBS	Phosphate buffered saline
PBST	PBS containing 1% Triton-X
PCMV	Porcine cytomegalovirus
PEG	Polyethylene glycol
PCR	Polymerase chain reaction
p.i.	Post infection
RCMV	Rat cytomegaloviurs
RCMV ALL-03	Rat cytomegalovirus strain ALL-03
RCMV-E	Rat cytomegalovirus, strain English
RCMV-M	Rat cytomegalovirus, strain Maastricht
RCMV-Sg	RCMV strain Sg
REF	Rat embryonic fibroblast
RE	Restriction enzyme
rpm	Revolutions per minute
RhCMV	Rhesus cytomegalovirus
SPCR	Simulated PCR
NaCl	Sodium chloride
SEM	Scanning Electron Microscope
SDS	Sodium dodecyl sulfate
NaOH	Sodium hydroxide
cm^2	Square centimeter
$TCID_{50}$	Median tissue culture infective dose
T	Thymine
TEM	Transmission Electron Microscope
TFF	Tangential flow filtration
UL	Unique long
US	Unique short
V	volt
v/v	volume per volume
w/v	weight per volume
X	Times
xg	x gravity, measure of centrifugal force

CHAPTER 1

INTRODUCTION

Cytomegalovirus (CMV) is a double stranded DNA virus which has the largest genomic size compare to other members in the family. Since the virus-infected cells become large and swollen, it was referred as "*cytomegalia*" and the name cytomegalovirus was derived from this character. Contagious CMV can be found in tears, saliva, blood transfusion, urine, breast milk and semen. CMV can exhibit silent without showing any symptoms in healthy individuals but can cause morbidity and mortality in patients whom undergo dsyregulation of the immune system. CMV is very good in molding host cell functions to support its replication and establish latency in fibroblasts, epithelial cells, endothelial cells, glial cells, smooth muscle cells of blood vessels, parenchymal cells in various organs and others (Liu et al., 2013). In addition, CMV is usually related with congenital viral infection where the virus infects maternal endothelial cells followed by placental cytotrophoblasts, eventually lead to transmission of the virus to foetus (Fisher et al., 2000). Other than that, congenital CMV infection may not show any symptoms or signs for newborn child but they can acquire hearing damage, mental retardation and cerebral palsy when growing old (Arvin et al., 2004). As a member of betaherpesvririnae subfamily, CMV infection is very specific to respective host. Since studying HCMV using human as host is ethically restricted, a suitable model to study CMV infection is needed. CMV infecting primates like Chimpanzee CMV (CCMV) and Rhesus CMV (RhCMV) became as first choice as model to mimic HCMV infection (Powers & Früh, 2008). Although primate CMV is closely related to Human CMV (HCMV), these strains are not frequently used as a model for HCMV infection due to impracticalities and high cost (Li et al., 2012). On the other hand, Murine CMV (MCMV) and Rat CMV (RCMV) have became well known models for studying HCMV because of low cost, high reproductive rates and simplicity of handling (Mocarski et al., 2007). The major drawback of these strains is that they do not cross the placental barrier and cause *in utero* infection, hence it is inapplicable and complicated to use for congenital infections. To overcome this, Loh et al (2003) acquired a new strain of RCMV strain ALL-03, from the uterus and placenta of the *Rattus rattus diaardi* (house rat). This study has demonstrated the ability of this strains' vertical transmission in pregnant rats therefore making it an appropriate model of choice to study the congenital infection of CMV in humans. In addition, the similar pathogenicity between HCMV and RCMV ALL-03 justifies its suitability as a good model to study HCMV (Loh et al., 2006).

To date, many CMVs have been sequenced including primates and non-primates CMV. As per our interest, many rodent CMVs also have been sequenced in different time frame. The first available Rat cytomegalovirus (RCMV) sequence was Maastricht strain in the year 2000 (Vink et al., 2000). Second, RCMV of England strain has been sequenced in year 2012 (Ettinger et al., 2012) and recently in year 2015, the complete genome sequence of the third strain of RCMV known as RCMV-Berlin have been at National Centre of Biotechnology Information (NCBI) (Geyer et al., 2015). From these data, two points can be highlighted where RCMV gain more attention from researchers to use as model to study the pathogenesis of Human Cytomegalovirus (HCMV) and producing a

complete genome sequence is vital to master the pathogenecity of HCMV at molecular level. Hence any biochemistry event related to the virus can be studied in deeper level.

To further elucidate the pathogenesis of our local Malaysian strain of RCMV ALL-03, genome sequencing of this virus is much crucial. Complete genome sequence of this virus is very important in terms of examine the coding sequence, understand the genomic organization, identify the protein coding genes and possible protein function and to carry out genome comparison with other strains of herpesviruses. Previously, RCMV ALL-03 strain was sequenced using Next Generation Sequencing with Illumina platform followed by the genome assembly which was carried out using CLC Genomic Workbench (Quah, 2013). However, complete genome sequence could not be produced due to the presence of gaps in the draft genome. Finishing phase known as final stage is needed to correcting the misassemblies and close all the gaps and come up with a complete genome sequence (Tang et al., 2013) .

The draft genome with gaps known as incomplete, even one of high coverage, represents a collection of contigs of various sizes, with unknown order and orientation, which contain sequencing errors and possible misassemblies. Until a genome has been closed, it is often difficult to identify contaminating sequences and these can confound subsequent comparative and functional genomics studies. There are many reasons to completing a draft genome. The draft genome of RCMV all-03 covered at least 90% of genome, however, further analysis should be made to exclude contaminating sequences, sequence errors and misassemblies. Identifying and correcting the low coverage regions and misassemblies can provide more accurate information to the community (Nagarajan et al., 2010b). The complete genome sequence is known as high quality reference compared to others and very suitable for all types of detailed analysis of genomic, proteomic as well as studying gene regulation.

However, currently the draft sequence of RCMV ALL-03 is still incomplete, with few gaps produced. Approximately, 198,895 total number of bases have been identified by previous researcher and arranged as single unique sequence flanked by 504 bp terminal direct repeats (Quah, 2013). It is anticipated that, by closing the gaps, complete sequence can be obtained and more accurate position and length of Open Reading Frame (ORF) can be identified. Once the gene coding region identified, the protein products and the functions can be predicted through bioinformatics means. Thus, in order to achieve the complete sequence, present study was conducted to identify the missing sequences in gap position thus replacing it to produce a complete genome sequence. Once complete genome sequence has been produced, all the open reading frame (ORF) were verified and subjected for blast comparison with other herpesviruses. In addition, all the ORF were also grouped into conserved genes as well categorized under functional gene families.

Complete CMV genome is important for researchers in developing recombinant DNA vaccine identification and manipulation of the virulent genes can be conducted in order to develop safer vaccines. Moreover, this research will enlighten our understanding on substantial genetic variation among other strains of herpesviruses, as well as develop more sensitive methods to detect, distinguish and identify different virus strains.

1.1 Research Hypothesis

It is hypothesized that by closing all the gaps present in RCMV ALL-03 draft sequence and resolved the misassemblies, a complete genome sequence of RCMV ALL-03 could be generated. Combination of conventional PCR assay and bioinformatics were used to achieve the objectives.

Ho: It is not possible to close all the sequencing gaps in draft sequence of RCMV ALL-03

Hi: It is possible to close all the sequencing gaps hence producing a complete genome sequence of Rat Cytomegalovirus ALL-03 (Malaysian strain).

Aim and Objectives of the study

This research aims to identify gap present in Rat Cytomegalovirus strain ALL-03 (Malaysian strain) and closing all the gaps in draft sequence. Specific objectives are:

1. To identify and resolve the gap regions located in the RCMV ALL-03 genome scaffold by conventional PCR and Sanger sequencing.

2. To verify the final assemble of RCMV ALL-03 and its functions of gene coding region (CDS) by CLC Genomic Workbench software.

3. To compare and study the Open Reading Frame (ORF) of RCMV ALL-03 with other strain of herpesviruses in terms of homology and function by online tool such as NCBI, ExPasy and related publications.

REFERENCES

Ablashi, D., Agut, H., Alvarez-Lafuente, R., Clark, D. A., Dewhurst, S., DiLuca, D., . . . Yoshikawa, T. (2014). Classification of HHV-6A and HHV-6B as distinct viruses. *Archives of virology, 159*(5), 863-870.

Adair, R., Douglas, E. R., Maclean, J. B., Graham, S. Y., Aitken, J. D., Jamieson, F. E., & Dargan, D. J. (2002). The products of human cytomegalovirus genes UL23, UL24, UL43 and US22 are tegument components. *The journal of general virology, 83*(Pt 6), 1315-1324.

Adler, S. P. (1989). Cytomegalovirus and child day care. Evidence for an increased infection rate among day-care workers. *The new england journal of medicine, 321*(19), 1290-1296.

Ahlqvist, J., & Mocarski, E. (2011). Cytomegalovirus UL103 controls virion and dense body egress. *Journal of virology, 85*(10), 5125-5135.

Ahn, J. H., & Hayward, G. S. (1997). The major immediate-early proteins IE1 and IE2 of human cytomegalovirus colocalize with and disrupt PML-associated nuclear bodies at very early times in infected permissive cells. *Journal of virology, 71*(6), 4599-4613.

Ahn, J. H., Jang, W. J., & Hayward, G. S. (1999). The human cytomegalovirus IE2 and UL112-113 proteins accumulate in viral DNA replication compartments that initiate from the periphery of promyelocytic leukemia protein-associated nuclear bodies (PODs or ND10). *Journal of virology, 73*(12), 10458-10471.

Albrechtand, Jens-Christian, & Fleckenstein, Bernhard. (1990). Structural organization of the conserved gene block of Herpesvirus saimiri coding for DNA polymerase, glycoprotein B, and major DNA binding protein. *Virology, 174*(2), 533-542.

Alford, C. A., Stagno, S., Pass, R. F., & Britt, W. J. (1990). Congenital and perinatal cytomegalovirus infections. *Reviews of infectious diseases, 12(7)*, 745-753.

Arvin, A. M., Fast, P., Myers, M., Plotkin, S., & Rabinovich, R. (2004). Vaccine development to prevent cytomegalovirus disease: report from the National Vaccine Advisory Committee. *Clinical infectious diseases, 39*(2), 233-239.

Asodi, Vera. (2008). Multiplex pcr for gap closing (whole-genome assembly). In M.-Y. Kao (Education.), *Encyclopedia of algorithms*,1-99.

Bao, Y., Chetvernin, V., & Tatusova, T. (2014). Improvements to pairwise sequence comparison (PASC): a genome-based web tool for virus classification. *Archives of virology, 159*(12), 3293-3304.

Battista, M. C., Bergamini, G., Boccuni, M. C., Campanini, F., Ripalti, A., & Landini, M. P. (1999). Expression and characterization of a novel structural protein of human cytomegalovirus, pUL25. *Journal of virology, 73*(5), 3800-3809.

Bechtel, J. T., & Shenk, T. (2002). Human cytomegalovirus UL47 tegument protein functions after entry and before immediate-early gene expression. *Journal of virology, 76*(3), 1043-1050.

Beisser, P. S., Kaptein, S. J., Beuken, E., Bruggeman, C. A., & Vink, C. (1998). The Maastricht strain and England strain of rat cytomegalovirus represent different betaherpesvirus species rather than strains. *Virology, 246*(2), 341-351.

Bello, C., & Whittle, H. (1991). Cytomegalovirus infection in Gambian mothers and their babies. *Journal of clinical pathology, 44*(5), 366-369.

Beuken, E., Grauls, G., Bruggeman, C. A., & Vink, C. (1999). The rat cytomegalovirus R32 gene encodes a virion-associated protein that elicits a strong humoral immune response in infected rats. *The journal of general virology, 80 (Pt 10)*, 2719-2728.

Biegalke, B. J., Lester, E., Branda, A., & Rana, R. (2004). Characterization of the Human Cytomegalovirus UL34 Gene. *Journal of virology, 78*(17), 9579-9583.

Bilofsky, H. S., Burks, C., Fickett, J. W., Goad, W. B., Lewitter, F. I., Rindone, W. P., . . . Tung, C. S. (1986). The GenBank genetic sequence databank. *Nucleic acids research, 14*(1), 1-4.

Blakesley, RobertW, Hansen, NancyF, Gupta, Jyoti, McDowell, JenniferC, Maskeri, Baishali, Barnabas, BeatriceB, . . . Green, EricD. (2010). Effort required to finish shotgun-generated genome sequences differs significantly among vertebrates. *BMC genomics, 11*(1), 1-15.

Blewett, E. L., White, G., Saliki, J. T., & Eberle, R. (2001). Isolation and characterization of an endogenous cytomegalovirus (BaCMV) from baboons. *Archives of virology, 146*(9), 1723-1738.

Boetzer, Marten, & Pirovano, Walter. (2012). Toward almost closed genomes with GapFiller. *Genome biology, 13*(6), R56.

Bogner, Elke, Radsak, Klaus, & Stinski, Mark F. (1998). The Gene Product of Human Cytomegalovirus Open Reading Frame UL56 Binds the pac Motif and Has Specific Nuclease Activity. *Journal of virology, 72*(3), 2259-2264.

Boivin, G., Goyette, N., Gilbert, C., & Covington, E. (2005). Analysis of cytomegalovirus DNA polymerase (UL54) mutations in solid organ transplant patients receiving valganciclovir or ganciclovir prophylaxis. *Journal of medical virology, 77*(3), 425-429.

Bolovan-Fritts, C. A., Trout, R. N., & Spector, S. A. (2004). Human cytomegalovirus-specific CD4+-T-cell cytokine response induces fractalkine in endothelial cells. *Journal of virology*, *78*(23), 13173-13181.

Borst, E. M., Kleine-Albers, J., Gabaev, I., Babic, M., Wagner, K., Binz, A., . . . Messerle, M. (2013). The human cytomegalovirus UL51 protein is essential for viral genome cleavage-packaging and interacts with the terminase subunits pUL56 and pUL89. *Journal of virology*, *87*(3), 1720-1732.

Borst, E. M., Wagner, K., Binz, A., Sodeik, B., & Messerle, M. (2008). The essential human cytomegalovirus gene UL52 is required for cleavage-packaging of the viral genome. *Journal of virology*, *82*(5), 2065-2078.

Bozidis, Petros, Williamson, Chad D., Wong, Daniel S., & Colberg-Poley, Anamaris M. (2010). Trafficking of UL37 Proteins into Mitochondrion-Associated Membranes during Permissive Human Cytomegalovirus Infection. *Journal of virology*, *84*(15), 7898-7903.

Britt, W. J., & Mach, M. (1996). Human cytomegalovirus glycoproteins. *Intervirology*, *39*(5-6), 401-412.

Brok, Herbert P. M., Boven, Leonie, van Meurs, Marjan, Kerlero de Rosbo, Nicole, Celebi-Paul, Liesbeth, Kap, Yolanda S., . . . t Hart, Bert A. (2007). The human CMV-UL86 peptide 981–1003 shares a crossreactive T-cell epitope with the encephalitogenic MOG peptide 34–56, but lacks the capacity to induce EAE in rhesus monkeys. *Journal of neuroimmunology*, *182*(1–2), 135-152.

Brown, TA (2002). Understanding a Genome Sequence. *Genomes 2nd edition. Oxford: wiley-liss, Chapter 7*.

Bruggeman, C. A., Meijer, H., Dormans, P. H. J., Debie, W. M. H., Grauls, G. E. L. M., & van Boven, C. P. A. (1982). Isolation of a cytomegalovirus-like agent from wild rats. *Archives of virology*, *73*(3-4), 231-241.

Buerger, I., Reefschlaeger, J., Bender, W., Eckenberg, P., Popp, A., Weber, O., . . . Hallenberger, S. (2001). A novel nonnucleoside inhibitor specifically targets cytomegalovirus DNA maturation via the UL89 and UL56 gene products. *Journal of virology*, *75*(19), 9077-9086.

Camalxaman, Siti Nazrina. (2012). Pathogenesis of RCMV ALL-03. *PhD Thesis, Universiti Putra Malaysia*.

Cantrell, Stacy R., & Bresnahan, Wade A. (2006). Human Cytomegalovirus (HCMV) UL82 Gene Product (pp71) relieves hdaxx-mediated repression of hcmv replication. *Journal of virology*, *80*(12), 6188-6191. Cao, Y., Wang, L., Xu, K., Kou, C., Zhang, Y., Wei, G., . . . Zhao, L. (2005). Information theory-based algorithm for in silico prediction of PCR products with whole genomic sequences as templates. *BMC bioinformatics*, *6*, 190.

Caposio, P., Riera, L., Hahn, G., Landolfo, S., & Gribaudo, G. (2004). Evidence that the human cytomegalovirus 46-kDa UL72 protein is not an active dUTPase but a late protein dispensable for replication in fibroblasts. *Virology, 325*(2), 264-276.

Castillo, J. P., & Kowalik, T. F. (2002). Human cytomegalovirus immediate early proteins and cell growth control. *Gene, 290*(1-2), 19-34.

Chambers, J., Angulo, A., Amaratunga, D., Guo, H., Jiang, Y., Wan, J. S., . . . Ghazal, P. (1999). DNA microarrays of the complex human cytomegalovirus genome: profiling kinetic class with drug sensitivity of viral gene expression. *Journal of virology, 73*(7), 5757-5766.

Chapa, T. J., Johnson, L. S., Affolter, C., Valentine, M. C., Fehr, A. R., Yokoyama, W. M., & Yu, D. (2013). Murine cytomegalovirus protein pM79 is a key regulator for viral late transcription. *Journal of virology, 87*(16), 9135-9147.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., . . . et al. (1990a). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Current topics in microbiology and immunology, 154*, 125-169.

Chee, M. S., Bankier, A. T., Beck, S., Bohni, R., Brown, C. M., Cerny, R., . . . et al. (1990b). Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169. *Current topics in microbiologyand immunology, 154*, 125-169.

Cole, Rufus, & Kuttner, Ann G. (1926). A filterable virus present in the submaxillary glands of guinea pigs. *The journal of experimental medicine, 44*(6), 855-873.

Crough, T., & Khanna, R. (2009). Immunobiology of human cytomegalovirus: from bench to bedside. *Clinical microbiology reviews, 22*(1), 76-98.

Davison, A. J. (2002). Evolution of the herpesviruses. *Veterinary microbiology, 86*(1-2), 69-88.

Davison, A. J., Dolan, A., Akter, P., Addison, C., Dargan, D. J., Alcendor, D. J., . . . Hayward, G. S. (2003). The human cytomegalovirus genome revisited: comparison with the chimpanzee cytomegalovirus genome. *The journal of general virology, 84*(Pt 1), 17-28.

Davison, A. J., Eberle, R., Ehlers, B., Hayward, G. S., McGeoch, D. J., Minson, A. C., . . . Thiry, E. (2009). The order Herpesvirales. *Archives of virology, 154*(1), 171-177.

Davison, AJ. , Arvin, A , Campadelli-Fiume, G, & E, Mocarski. (2007). Comparative analysis of the genomes. *Human herpesviruses: biology, therapy, and immunoprophylaxis, Cambridge university press*(Chapter 2.).

Davison, Andrew J. (2010). Herpesvirus systematics. *Veterinary microbiology, 143*(1-2), 52-69.

DeRussy, B. M., & Tandon, R. (2015). Human cytomegalovirus pUL93 is required for viral genome cleavage and packaging. *Journal of virology, 89*(23), 12221-12225.

Dolan, A., Cunningham, C., Hector, R. D., Hassan-Walker, A. F., Lee, L., Addison, C., . . . Davison, A. J. (2004). Genetic content of wild-type human cytomegalovirus. *The journal of general virology, 85*(Pt 5), 1301-1312.

Dunn, W., Chou, C., Li, H., Hai, R., Patterson, D., Stolc, V., . . . Liu, F. (2003b). Functional profiling of a human cytomegalovirus genome. *Proceedings of the national academy of sciences of the United States of America, 100*(24), 14223-14228.

Ettinger, J., Geyer, H., Nitsche, A., Zimmermann, A., Brune, W., Sandford, G. R., . . . Voigt, S. (2012). Complete genome sequence of the english isolate of rat cytomegalovirus (Murid herpesvirus 8). *Journal of virology, 86*(24), 13838.

Evans, Alfred S., Niederman, James C., & McCollum, Robert W. (1968). Seroepidemiologic studies of infectious mononucleosis with eb virus. *New england journal of medicine, 279*(21), 1121-1127.

Farber, Sidney, & Wolbach, S. Burt. (1932). Intranuclear and cytoplasmic inclusions ("protozoan-like bodies") in the salivary glands and other organs of infants. *The american journal of pathology, 8*(2), 123-136.123.

Fisher, S., Genbacev, O., Maidji, E., & Pereira, L. (2000). Human cytomegalovirus infection of placental cytotrophoblasts in vitro and in utero: implications for transmission and pathogenesis. *Journal of virology, 74*(15), 6808-6820.

Fitch, W. M., & Margoliash, E. (1967). Construction of phylogenetic trees. *Science, 155*(3760), 279-284.

Fraser, C. M., Eisen, J. A., Nelson, K. E., Paulsen, I. T., & Salzberg, S. L. (2002). The value of complete microbial genome sequencing (you get what you pay for). *Journal of bacteriology, 184*(23), 6403-6405.

Gandhi, M. K., & Khanna, R. (2004). Human cytomegalovirus: clinical aspects, immune regulation, and emerging treatments. *Lancet infectious diseases, 4*(12), 725-738.

Gao, Y., Colletti, K., & Pari, G. S. (2008). Identification of human cytomegalovirus UL84 virus- and cell-encoded binding partners by using proteomics analysis. *Journal of virology, 82*(1), 96-104.

Geyer, H., Ettinger, J., Moller, L., Schmolz, E., Nitsche, A., Brune, W., . . . Voigt, S. (2015). The English isolate and a newly identified Berlin isolate of Rat Cytomegalovirus (RCMV) share similarities with but separate as an anciently diverged clade from Mouse CMV and the Maastricht isolate of RCMV. *Journal of general virology*, 96(7), 1873-1881.

Gibson, W. (1996). Structure and assembly of the virion. *Intervirology, 39*(5-6), 389-400.

Gibson, W., Baxter, M. K., & Clopper, K. S. (1996). Cytomegalovirus "missing" capsid protein identified as heat-aggregable product of human cytomegalovirus UL46. *Journal of virology, 70*(11), 7454-7461.

Gilbert, Walter, & Maxam, Allan. (1973). The nucleotide sequence of the lac operator. *Proceedings of the national academy of sciences of the United States of America, 70*(12), 3581-3584.

Gocayne, J., Robinson, D. A., FitzGerald, M. G., Chung, F. Z., Kerlavage, A. R., Lentes, K. U., . . . Venter, J. C. (1987). Primary structure of rat cardiac beta-adrenergic and muscarinic cholinergic receptors obtained by automated DNA sequence analysis: further evidence for a multigene family. *Proceedings of the national academy of sciences of the United States of America, 84*(23), 8296-8300.

Goodrum, Felicia, Caviness, Katie, & Zagallo, Patricia. (2012). Human cytomegalovirus persistence. *Cellular microbiology, 14*(5), 644-655.

Greis, K. D., Gibson, W., & Hart, G. W. (1994). Site-specific glycosylation of the human cytomegalovirus tegument basic phosphoprotein (UL32) at serine 921 and serine 952. *Journal of virology, 68*(12), 8339-8349.

Hahn, Gabriele, Khan, Hanna, Baldanti, Fausto, Koszinowski, Ulrich H., Revello, M. Grazia, & Gerna, Giuseppe. (2002). The human cytomegalovirus ribonucleotide reductase homolog ul45 is dispensable for growth in endothelial cells, as determined by a bac-cloned clinical isolate of human cytomegalovirus with preserved wild-type characteristics. *Journal of virology, 76*(18), 9551-9555.

Halwachs-Baumann, G. (2006). The congenital cytomegalovirus infection: virus-host interaction for defense and transmission. *Current pharmaceutical biotechnology, 7*(4), 303-312.

Halwachs-Baumann, Gabriele. (2011). Long known, long ignored — a brief history of cytomegalovirus research. In G. Halwachs-Baumann (Ed.), *Congenital cytomegalovirus infection*, 1-10.

Hirst, M., & Marra, M. A. (2010). Next generation sequencing based approaches to epigenomics. *Briefings in functional genomics, 9*(5-6), 455-465.

Ho, M. (2008). The history of cytomegalovirus and its diseases. *Medical microbiology and immunology, 197*(2), 65-73.

Huang, E. S., Kilpatrick, B. A., Huang, Y. T., & Pagano, J. S. (1976). Detection of human cytomegalovirus and analysis of strain variation. *The yale journal of biology and medicine, 49*(1), 29-43.

Huber, Mary T., & Compton, Teresa. (1998). The human cytomegalovirus UL74 gene encodes the third component of the glycoprotein h-glycoprotein l-containing envelope complex. *Journal of virology, 72*(10), 8191-8197.

Husemann, Peter. (2011). Bioinformatic approaches for genome finishing. *Ph.D. Thesis*.

Irmiere, Alice, & Gibson, Wade. (1983). Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. *Virology, 130*(1), 118-133.

Isomura, Hiroki, Stinski, Mark F., Murata, Takayuki, Nakayama, Sanae, Chiba, Shigeki, Akatsuka, Yoshiki, . . . Tsurumi, Tatsuya. (2010). The human cytomegalovirus UL76 gene regulates the level of expression of the UL77 gene. *Plos one, 5*(7), 11901.

Isomura, Hiroki, Stinski, Mark F., Murata, Takayuki, Yamashita, Yoriko, Kanda, Teru, Toyokuni, Shinya, & Tsurumi, Tatsuya. (2011). The human cytomegalovirus gene products essential for late viral gene expression assemble into prereplication complexes before viral dna replication. *Journal of virology, 85*(13), 6629-6644.

Jesionek , A , & Kiolemenoglou. (1904). Ueber einen befund von protozoenartigen gebilden in den organen eines hereditar-luetischen foetus. *Muenchner med wochenschr, 51*, 1905-1907.

Jones, T. R., Sun, L., Bebernitz, G. A., Muzithras, V. P., Kim, H. J., Johnston, S. H., & Baum, E. Z. (1994). Proteolytic activity of human cytomegalovirus UL80 protease cleavage site mutants. *Journal of virology, 68*(6), 3742-3752.

Kaye, J. F., Gompels, U. A., & Minson, A. C. (1992). Glycoprotein H of human cytomegalovirus (HCMV) forms a stable complex with the HCMV UL115 gene product. *The journal of general virology, 73 (Pt 10)*, 2693-2698.

Kern, E. R. (2006). Pivotal role of animal models in the development of new therapies for cytomegalovirus infections. *Antiviral research, 71*(2-3), 164-171.

Kim, C. S. (2010). Congenital and perinatal cytomegalovirus infection. *Korean journal of pediatric, 53*(1), 14-20.

Kimura, Motoo. (1969). The rate of molecular evolution considered from the standpoint of population genetics. *Proceedings of the national academy of sciences of the United States of America, 63*(4), 1181-1188.

Kinzler, E. R., & Compton, T. (2005). Characterization of human cytomegalovirus glycoprotein-induced cell-cell fusion. *Journal of virology, 79*(12), 7827-7837.

Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing--concepts and limitations. *Bioessays: news and reviews in molecular, cellular and developmental biology, 32*(6), 524-536.

Komazin, G., Townsend, L. B., & Drach, J. C. (2004). Role of a mutation in human cytomegalovirus gene UL104 in resistance to benzimidazole ribonucleosides. *Journal of virology, 78*(2), 710-715.

Krzyzaniak, Magdalena A., Mach, Michael, & Britt, William J. (2009). HCMV Encoded Glycoprotein M (UL100) Interacts with Rab11 Effector Protein FIP4. *Traffic, 10*(10), 1439-1457.

Lai, Mohd-Azmi, M.L , Khoo, K.K, & Sheikh Omar, A.R. (1999b). Random amplified polymorphic DNA (RAPD) for identification of new rat cytomegalovirus isolated from rice-field rats. *Jurnal veterinar Malaysia, 11*, 23-26.

Landolfo, S., Gariglio, M., Gribaudo, G., & Lembo, D. (2003). The human cytomegalovirus. *Pharmacology & therapeutics, 98*(3), 269-297.

Lang, D. J., & Kummer, J. F. (1975). Cytomegalovirus in semen: observations in selected populations. *The journal of infectious diseases, 132*(4), 472-473.

Lapidus, Alla L. (2009). Genome sequence databases (overview): sequencing and assembly.

Lawrence, R. M. (2006). Cytomegalovirus in human breast milk: risk to the premature infant. *Breastfeeding medicine : the official journal of the academy of breastfeeding medicine, 1*(2), 99-107.

Li, Z., Wang, X., Yan, S., Zhang, Z., Jie, C., Sustento-Reodica, N., . . . Abecassis, M. (2012). A mouse model of CMV transmission following kidney transplantation. *American journal of transplantation 12*(4), 1024-1028.

Liao, John B. (2006). Viruses and human cancer. *The yale journal of biology and medicine, 79*(3-4), 115-122.

Lindquester, G. J., & Pellett, P. E. (1991). Properties of the human herpesvirus 6 strain Z29 genome: G + C content, length, and presence of variable-length directly repeated terminal sequence elements. *Virology, 182*(1), 102-110.

Lipschutz, B. (1921). Untersuchungen uber die atiologie der krankheiten der herpesgruppe (herpes zoster, herpes genitalis, herpes febrilis). *Archieves of dermatology and syphilis*, *136*, 428-482.

Liu, B., & Stinski, M. F. (1992). Human cytomegalovirus contains a tegument protein that enhances transcription from promoters with upstream ATF and AP-1 cis-acting elements. *Journal of virology*, *66*(7), 4434-4444.

Liu, Lin, Li, Yinhu, Li, Siliang, Hu, Ni, He, Yimin, Pong, Ray, . . . Law, Maggie. (2012a). Comparison of next-generation sequencing systems. *Journal of biomedicine and biotechnology*, *2012*, 11.

Liu, Lin, Li, Yinhu, Li, Siliang, Hu, Ni, He, Yimin, Pong, Ray, . . . Law, Maggie. (2012b). Comparison of next-generation sequencing systems. *Journal of biomedicine and biotechnology*, *2012*, 251364.

Liu, X. F., Wang, X., Yan, S., Zhang, Z., Abecassis, M., & Hummel, M. (2013). Epigenetic control of cytomegalovirus latency and reactivation. *Viruses*, *5*(5), 1325-1345.

Liu, Yingguang, & Biegalke, Bonita J. (2002). The human cytomegalovirus UL35 gene encodes two proteins with different functions. *Journal of virology*, *76*(5), 2460-2468.

Loh, H. S., Mohd-Azmi, M. L., Lai, K. Y., Sheikh-Omar, A. R., & Zamri-Saad, M. (2003). Characterization of a novel rat cytomegalovirus (RCMV) infecting placenta-uterus of Rattus rattus diardii. *Archives of virology*, *148*(12), 2353-2367.

Loh, Hwei-San, Mohd-Lila, Mohd-Azmi, Abdul-Rahman, Sheikh-Omar, & Kiew, Lik-Jun. (2006). Pathogenesis and vertical transmission of a transplacental rat cytomegalovirus. *Virology journal*, *3*, 42-42.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., . . . Wang, J. (2012). SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience*, *1*(1), 18.

Ma, Y., Wang, N., Li, M., Gao, S., Wang, L., Ji, Y., . . . Ruan, Q. (2011). An antisense transcript in the human cytomegalovirus UL87 gene region. *Virology journal*, *8*, 515.

Mach, M., Kropff, B., Dal Monte, P., & Britt, W. (2000). Complex formation by human cytomegalovirus glycoproteins M (gpUL100) and N (gpUL73). *Journal of virology*, *74*(24), 11881-11892.

Margulies, M., Egholm, M., Altman, W. E., Attiya, S., Bader, J. S., Bemben, L. A., . . . Rothberg, J. M. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature*, *437*(7057), 376-380.

Martin, M., Gilbert, C., Covington, E., & Boivin, G. (2006). Characterization of human cytomegalovirus (HCMV) UL97 mutations found in a valganciclovir/oral ganciclovir prophylactic trial by use of a bacterial artificial chromosome containing the HCMV genome. *The journal of infectious diseases, 194*(5), 579-583.

Maussang, D., Vischer, H. F., Leurs, R., & Smit, M. J. (2009). Herpesvirus-encoded G protein-coupled receptors as modulators of cellular function. *Molecular pharmacology, 76*(4), 692-701.

McCormick, A. Louise, Roback, Linda, Livingston-Rosanoff, Devon, & St. Clair, Courtney. (2010). The Human cytomegalovirus UL36 gene controls caspase-dependent and -independent cell death programs activated by infection of monocytes differentiating to macrophages. *Journal of virology, 84*(10), 5108-5123.

McGeoch, D. J., Rixon, F. J., & Davison, A. J. (2006a). Topics in herpesvirus genomics and evolution. *Virus research, 117*(1), 90-104.

McGeoch, D. J., Rixon, F. J., & Davison, A. J. (2006b). Topics in herpesvirus genomics and evolution. *Virus research, 117*(1), 90-104.

McGregor, A. (2010). Current and new cytomegalovirus antivirals and novel animal model strategies. *Inflammatory allergy drug targets, 9*(4), 286-299.

McWatters, B. J., Stenberg, R. M., & Kerry, J. A. (2002). Characterization of the human cytomegalovirus UL75 (glycoprotein H) late gene promoter. *Virology, 303*(2), 309-316.

Melnick, J. L., Petrie, B. L., Dreesman, G. R., Burek, J., McCollum, C. H., & DeBakey, M. E. (1983). Cytomegalovirus antigen within human arterial smooth muscle cells. *Lancet, 2*(8351), 644-647.

Mertes, Florian, ElSharawy, Abdou, Sauer, Sascha, van Helvoort, Joop M.L.M., van der Zaag, P.J., Franke, Andre, . . . Brookes, Anthony J. (2011). Targeted enrichment of genomic DNA regions for next-generation sequencing. *Briefings in functional genomics, 10*(6), 374-386.

Mettenleiter, T. C. (2004). Budding events in herpesvirus morphogenesis. *Virus research, 106*(2), 167-180.

Minder, W. H. (1953). Etiology of cytomegaly in infants. *Schweiz med wochenschr, 83*(49), 1180-1182.

Mocarski, Shenk, T, & Pass, R.F (2007). Cytomegaloviruses. *Fields virology*, 22701-22772.

Mocarski, E.S. (2001). Cytomegaloviruses and their replication. *Fields virology*, 2629-2673.

Mocarski, Jr , Arvin, A , & Campadelli-Fiume, G (2007). Comparative analysis of herpesvirus-common proteins. *Human herpesviruses: biology, therapy, and immunoprophylaxis. Cambridge university press*(chapter 4).

Motamedifar, Mohammad, & Noorafshan, Ali. (2008). Cytopathic effect of the herpes simplex virus type 1 appears stereologically as early as 48 hours after infection of Vero cells. *Micron, 39*(8), 1331-1334.

Mulyukov, Z., & Pevzner, P. A. (2002). EULER-PCR: finishing experiments for repeat resolution. *Pacific symposium and biocomputer*, 199-210.

Nagarajan, Cook, C., Di Bonaventura, M., Ge, H., Richards, A., Bishop-Lilly, K. A., . . . Pop, M. (2010a). Finishing genomes with limited resources: lessons from an ensemble of microbial genomes. *BMC genomics, 11*, 242.

Nagarajan, Cook, Christopher, Di Bonaventura, MariaPia, Ge, Hong, Richards, Allen, Bishop-Lilly, Kimberly, . . . Pop, Mihai. (2010b). Finishing genomes with limited resources: lessons from an ensemble of microbial genomes. *BMC genomics, 11*(1), 242.

Neubauer, Susan R. (2011). Characterization of the genome of baboon cytomegalovirus strain (OCOM4-37) isolated from the olive baboon, Papio cynocephalus anubis. (3498740 Ph.D.), Oklahoma state university, Ann Arbor.

Omoto, Shinya, & Mocarski, Edward S. (2014). Transcription of true late (γ2) cytomegalovirus genes requires UL92 function that is conserved among beta- and gammaherpesviruses. *Journal of virology, 88*(1), 120-130.

Pandey, J. P., Namboodiri, A. M., Radwan, F. F., & Nietert, P. J. (2015). The decoy Fcgamma receptor encoded by the cytomegalovirus UL119-UL118 gene has differential affinity to IgG proteins expressing different GM allotypes. *Human immunology, 76*(8), 591-594.

Pass, Robert F., Hutto, Cecelia, Ricks, Rebecca, & Cloud, Gretchen A. (1986). Increased rate of cytomegalovirus infection among parents of children attending day-care centers. *New england journal of medicine, 314*(22), 1414-1418.

Pellet, PE, & Roizman, B (2007). The Family Herpesviridae: A brief introduction. *In: David M, Knipe PMH, editors. Fields virology*, 2479–2499.

Perdue, M. L., Cohen, J. C., Randall, C. C., & O'Callaghan, D. J. (1976). Biochemical studies of the maturation of herpesvirus nucleocapsid species. *Virology, 74*(1), 194-208.

Pereira, L., Stagno, S., Hoffman, M., & Volanakis, J. E. (1983). Cytomegalovirus-infected cell polypeptides immune-precipitated by sera from children with congenital and perinatal infections. *Infection and immunity, 39*(1), 100-108.

Perng, Y. C., Qian, Z., Fehr, A. R., Xuan, B., & Yu, D. (2011a). The human cytomegalovirus gene UL79 is required for the accumulation of late viral transcripts. *Journal of virology, 85*(10), 4841-4852.

Perng, Y. C., Qian, Z., Fehr, A. R., Xuan, B., & Yu, D. (2011b). The human cytomegalovirus gene UL79 is required for the accumulation of late viral transcripts. *Journal of virology, 85*(10), 4841-4852.

Phillips, S. L., & Bresnahan, W. A. (2012). The human cytomegalovirus (HCMV) tegument protein UL94 is essential for secondary envelopment of HCMV virions. *Journal of virology, 86*(5), 2523-2532.

Phillips, S. L., & Bresnahan, W. A. (2011). Identification of binary interactions between human cytomegalovirus virion proteins. *Journal of virology, 85*(1), 440-447.

Piro, Faoro, Helisson, Weiss, Vinicius, Steffens, Maria, Pedrosa, Fabio, Souza, Emanuel, & Raittz, Roberto. (2014a). FGAP: an automated gap closing tool. *BMC research notes, 7*(1), 371.

Piro, Faoro, Helisson, Weiss, ViniciusA, Steffens, MariaBR, Pedrosa, FabioO, Souza, EmanuelM, & Raittz, RobertoT. (2014b). FGAP: an automated gap closing tool. *BMC research notes, 7*(1), 1-5.

Ploegh, H. L. (1998). Viral strategies of immune evasion. *Science, 280*(5361), 248-253.

Plosa, E. J., Esbenshade, J. C., Fuller, M. P., & Weitkamp, J. H. (2012). Cytomegalovirus infection. *American academy of pediatrics, 33*(4), 156-163.

Poncet, Delphine, Pauleau, Anne-Laure, Szabadkai, Gyorgy, Vozza, Angelo, Scholz, Sebastian R., Le Bras, Morgane, . . . Kroemer, Guido. (2006). Cytopathic effects of the cytomegalovirus-encoded apoptosis inhibitory protein vMIA. *The journal of cell biology, 174*(7), 985-996.

Powers, Colin, & Früh, Klaus. (2008). Rhesus CMV: an emerging animal model for human CMV. *Medical microbiology and immunology, 197*(2), 109-115.

Pramod, N. P., Gopalakrishnan, V., Mohan, R., Chandriga, S., Anandakannan, K., Menon, T., & Thyagarajan, S. P. (1998). Enhanced detection of herpes simplex virus from ocular specimens of herpetic keratitis patients. *Indian journal of pathology and microbiology, 41*(1), 49-53.

Priscott, P. K., & Tyrrell, D. A. (1982). The isolation and partial characterisation of a cytomegalovirus from the brown rat, Rattus norvegicus. *Archives of virology, 73*(2), 145-160.

Puranik, Rutika, Quan, Guangri, Werner, Jacob, Zhou, Rong, & Xu, Zhaohui. (2015). A pipeline for completing bacterial genomes using in silico and wet lab approaches. *BMC genomics, 16*.

Qu, W., Shen, Z., Zhao, D., Yang, Y., & Zhang, C. (2009). MFEprimer: multiple factor evaluation of the specificity of PCR primers. *Bioinformatics, 25*(2), 276-278.

Quah, Yi Wan. (2013). Molecular characterisation of Glycoprotein B and Lower matrix phosphoprotein of Rat Cytomegalovirus ALL-03 strain. *Unpublished Master dissertation. Universiti Putra Malayisa*.

Quail, M. A., Smith, M., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., . . . Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics, 13*, 341.

Rajcani, J., & Durmanova, V. (2001). Mechanisms of replication of alpha- and betaherpesviruses and their pathogenesis. *Bratislavske lekarske listy, 102*(11), 505-514.

Ranneberg-Nilsen, Toril, Dale, Hege Avsnes, Luna, Luisa, Slettebakk, Ragnhild, Sundheim, Ottar, Rollag, Halvor, & Bjørås, Magnar. (2008). Characterization of Human Cytomegalovirus Uracil DNA Glycosylase (UL114) and Its Interaction with Polymerase Processivity Factor (UL44). *Journal of molecular biology, 381*(2), 276-288.

Rasmussen, L., Geissler, A., Cowan, C., Chase, A., & Winters, M. (2002). The genes encoding the gCIII complex of human cytomegalovirus exist in highly diverse combinations in clinical isolates. *Journal of virology, 76*(21), 10841-10848.

Rawlinson, W. D., Farrell, H. E., & Barrell, B. G. (1996). Analysis of the complete DNA sequence of murine cytomegalovirus. *Journal of virology, 70*(12), 8833-8849.

Reddehase, M. J. (2002). Antigens and immunoevasins: opponents in cytomegalovirus immune surveillance. *Nature reviews of immunology, 2*(11), 831-844.

Reed, L.J., & Muench, H. (1938). A simple method of estimating fifty per cent endpoints. *American journal of epidemiology, 27*(3), 493-497.

Ribbert, H. (1904). Ueber protozoenartige Zellen in der Niere eines syphilitischen Neugeborenen und in der Parotis von Kindern. *Pathology,15*, 945-948.

Rice, C. M., Fuchs, R., Higgins, D. G., Stoehr, P. J., & Cameron, G. N. (1993). The EMBL data library. *Nucleic acids research, 21*(13), 2967-2971.

Riddell, S. R., & Greenberg, P. D. (1995). Principles for adoptive T cell therapy of human viral diseases. *Annual review of immunology, 13*, 545-586.

Rixon, F. J. (1993). Structure and assembly of herpesviruses. *Seminars in virology, 4*(3), 135-144.

Roizman, B., Carmichael, L. E., Deinhardt, F., de-The, G., Nahmias, A. J., Plowright, W., . . . Wolf, K. (1981). Herpesviridae. Definition, provisional nomenclature, and taxonomy. The Herpesvirus Study Group, the International Committee on Taxonomy of Viruses. *Intervirology, 16*(4), 201-217.

Russo, J. J., Bohenzky, R. A., Chien, M. C., Chen, J., Yan, M., Maddalena, D., . . . Moore, P. S. (1996). Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus (HHV8). *Proceedings of the national academy of sciences of the United States of America, 93*(25), 14862-14867.

Ruvolo, V. R., Berneman, Z., Secchiero, P., & Nicholas, J. (1996). Cloning, restriction endonuclease mapping and partial sequence analysis of the genome of human herpesvirus 7 strain JI. *The journal of general virology, 77 (Pt 8)*, 1901-1912.

Ryner, Martin, Stromberg, Jan-Olov, Soderberg-Naucler, Cecilia, & Homman-Loudiyi, Mohammed. (2006). Identification and classification of human cytomegalovirus capsids in textured electron micrographs using deformed template matching. *Virology journal, 3*(1), 57.

Sandford, Gordon R., Brock, Lorine E., Voigt, Sebastian, Forester, Craig M., & Burns, William H. (2001). Rat Cytomegalovirus Major Immediate-Early Enhancer Switching Results in Altered Growth Characteristics. *Journal of virology, 75*(11), 5076-5083.

Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, C. A., . . . Smith, M. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature, 265*(5596), 687-695.

Schauflinger, Martin, Fischer, Daniela, Schreiber, Andreas, Chevillotte, Meike, Walther, Paul, Mertens, Thomas, & von Einem, Jens. (2011). The Tegument Protein UL71 of Human Cytomegalovirus Is Involved in Late Envelopment and Affects Multivesicular Bodies. *Journal of virology, 85*(8), 3821-3832.

Schleiss, M. R. (2006). Nonprimate models of congenital cytomegalovirus (CMV) infection: gaining insight into pathogenesis and prevention of disease in newborns. *Institute of laboratory animal resources, 47*(1), 65-72.

Schmolke, S., Drescher, P., Jahn, G., & Plachter, B. (1995). Nuclear targeting of the tegument protein pp65 (UL83) of human cytomegalovirus: an unusual

bipartite nuclear localization signal functions with other portions of the protein to mediate its efficient nuclear transport. *Journal of virology*, *69*(2), 1071-1078.

Schmolke, S., Kern, H. F., Drescher, P., Jahn, G., & Plachter, B. (1995). The dominant phosphoprotein pp65 (UL83) of human cytomegalovirus is dispensable for growth in cell culture. *Journal of virology*, *69*(10), 5959-5968.

Schnee, Margit, Ruzsics, Zsolt, Bubeck, Anja, & Koszinowski, Ulrich H. (2006). Common and specific properties of herpesvirus UL34/UL31 Protein family members revealed by protein complementation assay. *Journal of virology*, *80*(23), 11658-11666.

Seckert, C. K., Renzaho, A., Reddehase, M. J., & Grzimek, N. K. (2008). Hematopoietic stem cell transplantation with latently infected donors does not transmit virus to immunocompromised recipients in the murine model of cytomegalovirus infection. *Medical microbiology and immunology*, *197*(2), 251-259.

Severi, B., Landini, M. P., Cenacchi, G., Zini, N., & Maraldi, N. M. (1992). Human cytomegalovirus nuclear and cytoplasmic dense bodies. *Archives of virology*, *123*(1-2), 193-207.

Sharma, M., Kamil, J. P., Coughlin, M., Reim, N. I., & Coen, D. M. (2014). Human cytomegalovirus UL50 and UL53 recruit viral protein kinase UL97, not protein kinase C, for disruption of nuclear lamina and nuclear egress in infected cells. *Journal of virology*, *88*(1), 249-262.

Sheaffer, A. K., Weinheimer, S. P., & Tenney, D. J. (1997). The human cytomegalovirus UL98 gene encodes the conserved herpesvirus alkaline nuclease. *The journal of general virology*, *78 (Pt 11)*, 2953-2961.

Shen, A., Lei, J., Yang, E., Pei, Y., Chen, Y. C., Gong, H., . . . Liu, F. (2011). Human cytomegalovirus primase UL70 specifically interacts with cellular factor Snapin. *Journal of virology*, *85*(22), 11732-11741.

Shendure, Jay, & Aiden, Erez Lieberman. (2012). The expanding scope of DNA sequencing. *Nature biotechnology*, *30*(11), 1084-1094.

Shendure, Jay, & Ji, Hanlee. (2008). Next-generation DNA sequencing. *Nature biotechnology*, *26*(10), 1135-1145.

Singh, Kanwar. (2010). Gap Closing/Finishing by Targeted Genomic Region Enrichment and Sequencing.

Sinigalia, Elisa, Alvisi, Gualtiero, Segré, Chiara V., Mercorelli, Beatrice, Muratore, Giulia, Winkler, Michael, . . . Loregian, Arianna. (2012). The Human

Cytomegalovirus DNA Polymerase Processivity Factor UL44 Is Modified by SUMO in a DNA-Dependent Manner. *Plos one*, *7*(11), 49630.

Smith, David Roy. (2014). Buying in to bioinformatics: an introduction to commercial sequence analysis software. *Briefings in bioinformatics*.

Smith, J. A., & Pari, G. S. (1995). Human cytomegalovirus UL102 gene. *Journal of virology*, *69*(3), 1734-1740.

Smith, Jean A., Jairath, Sanju, Crute, James J., & Pari, Gregory S. (1996). Characterization of the human cytomegalovirus UL105 gene and identification of the putative helicase protein. *Virology*, *220*(1), 251-255.

Smith, L. M., Sanders, J. Z., Kaiser, R. J., Hughes, P., Dodd, C., Connell, C. R., . . . Hood, L. E. (1986). Fluorescence detection in automated DNA sequence analysis. *Nature*, *321*(6071), 674-679.

Smith, M. G. (1954). Propagation of salivary gland virus of the mouse in tissue cultures. *Society for experimental biology and medicine 86*(3), 435-440.

Smuda, C., Bogner, E., & Radsak, K. (1997). The human cytomegalovirus glycoprotein B gene (ORF UL55) is expressed early in the infectious cycle. *The journal of general virology, 78 (Pt 8)*, 1981-1992.

Sohn, Y. M., Oh, M. K., Balcarek, K. B., Cloud, G. A., & Pass, R. F. (1991). Cytomegalovirus infection in sexually active adolescents. *The journal of infectious diseases*, *163*(3), 460-463.

Speir, E., Modali, R., Huang, E. S., Leon, M. B., Shawl, F., Finkel, T., & Epstein, S. E. (1994). Potential role of human cytomegalovirus and p53 interaction in coronary restenosis. *Science*, *265*(5170), 391-394.

Stals, F. S., Bosman, F., van Boven, C. P., & Bruggeman, C. A. (1990a). An animal model for therapeutic intervention studies of CMV infection in the immunocompromised host. *Archieves of virology*, *114*(1-2), 91-107.

Stals, F. S., Bosman, F., van Boven, C. P., & Bruggeman, C. A. (1990b). An animal model for therapeutic intervention studies of CMV infection in the immunocompromised host. *Archives of virology*, *114*(1-2), 91-107.

Stamminger, T., Gstaiger, M., Weinzierl, K., Lorz, K., Winkler, M., & Schaffner, W. (2002). Open reading frame UL26 of human cytomegalovirus encodes a novel tegument protein that contains a strong transcriptional activation domain. *Journal of virology*, *76*(10), 4836-4847.

Steven, A. C , & Spear, P. G. . (1997). Herpes capsid assembly and envelopment. *Structural biology of viruses*, 312–351.

Strauss, E. C., Kobori, J. A., Siu, G., & Hood, L. E. (1986). Specific-primer-directed DNA sequencing. *Analytical biochemistry, 154*(1), 353-360.

Tadagaki, K., Tudor, D., Gbahou, F., Tschische, P., Waldhoer, M., Bomsel, M., . . . Kamal, M. (2012). Human cytomegalovirus-encoded UL33 and UL78 heteromerize with host CCR5 and CXCR4 impairing their HIV coreceptor activity. *Blood, 119*(21), 4908-4918.

Tandon, Ritesh, Mocarski, Edward, & Conway, James. (2015). The A, B, Cs of Herpesvirus Capsids. *Viruses, 7*(3), 899.

Tandon, Ritesh, & Mocarski, Edward S. (2011). Cytomegalovirus pUL96 Is Critical for the Stability of pp150-Associated Nucleocapsids. *Journal of virology, 85*(14), 7129-7141.

Tang, Biao, Wang, Qi, Yang, Minjun, Xie, Feng, Zhu, Yongqiang, Zhuo, Ying, . . . Zheng, Huajun. (2013). ContigScape: a Cytoscape plugin facilitating microbial genome gap closing. *BMC genomics, 14*(1), 289.

Terhune, S., Torigoi, E., Moorman, N., Silva, M., Qian, Z., Shenk, T., & Yu, D. (2007). Human cytomegalovirus UL38 protein blocks apoptosis. *Journal of virology, 81*(7), 3109-3123.

Tettelin, H., Radune, D., Kasif, S., Khouri, H., & Salzberg, S. L. (1999). Optimized multiplex PCR: efficiently closing a whole-genome shotgun sequencing project. *Genomics, 62*(3), 500-507.

Tortorella, D., Gewurz, B. E., Furman, M. H., Schust, D. J., & Ploegh, H. L. (2000). Viral subversion of the immune system. *Annual review of immunology, 18*, 861-926.

Tsai, Isheng, Otto, Thomas, & Berriman, Matthew. (2010). Improving draft assemblies by iterative mapping and assembly of short reads to eliminate gaps. *Genome biology, 11*(4), R41.

Vajda, B. P. (1978). Concentration and purification of viruses and bacteriophages with polyethylene glycol. *Folia microbiology, 23*(1), 88-96.

Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., . . . Johnson, S. M. (2008). A high-resolution, nucleosome position map of C. elegans reveals a lack of universal sequence-dictated positioning. *Genome research, 18*(7), 1051-1063.

van Hijum, Sacha A. F. T., Zomer, Aldert L., Kuipers, Oscar P., & Kok, Jan. (2005). Projector 2: contig mapping for efficient gap-closure of prokaryotic genome sequence assemblies. *Nucleic acids research, 33*(2), W560-W566.

Varnum, S. M., Streblow, D. N., Monroe, M. E., Smith, P., Auberry, K. J., Pasa-Tolic, L., . . . Nelson, J. A. (2004). Identification of proteins in human cytomegalovirus (HCMV) particles: the HCMV proteome. *Journal of virology, 78*(20), 10960-10966.

Vink, Cornelis, Beuken, Erik, & Bruggeman, Cathrien A. (2000). Complete DNA sequence of the rat cytomegalovirus genome. *Journal of virology, 74*(16), 7656-7665.

Voigt, S., Sandford, G. R., Hayward, G. S., & Burns, W. H. (2005). The English strain of rat cytomegalovirus (CMV) contains a novel captured CD200 (vOX2) gene and a spliced CC chemokine upstream from the major immediate-early region: further evidence for a separate evolutionary lineage from that of rat CMV Maastricht. *Journal of general virology, 86*(Pt 2), 263-274.

Vonglahn, W. C., & Pappenheimer, A. M. (1925). Intranuclear Inclusions in Visceral Disease. *The american journal of pathology, 1*(5), 445-466 443.

Wakeling, M. N., Roy, D. J., Nash, A. A., & Stewart, J. P. (2001). Characterization of the murine gammaherpesvirus 68 ORF74 product: a novel oncogenic G protein-coupled receptor. *The journal of general virology, 82*(Pt 5), 1187-1197.

Wan, Yi. (2013). Molecular characterisation of Glycoprotein B and Lower matrix phosphoprotein of Rat cytomegalovirus ALL-03 strain. *Thesis*

Wang, Jianlei, Loveland, Amy N., Kattenhorn, Lisa M., Ploegh, Hidde L., & Gibson, Wade. (2006). High-Molecular-Weight Protein (pUL48) of Human Cytomegalovirus Is a Competent Deubiquitinating Protease: Mutant Viruses Altered in Its Active-Site Cysteine or Histidine Are Viable. *Journal of virology, 80*(12), 6003-6012.

Wang, S. K., Duh, C. Y., & Wu, C. W. (2004). Human cytomegalovirus UL76 encodes a novel virion-associated protein that is able to inhibit viral replication. *Journal of virology, 78*(18), 9750-9762.

Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature, 171*(4356), 737-738.

Weir, J. P. (1998). Genomic organization and evolution of the human herpesviruses. *Virus genes, 16*(1), 85-93.

Winkler, M., Rice, S. A., & Stamminger, T. (1994). UL69 of human cytomegalovirus, an open reading frame with homology to ICP27 of herpes simplex virus, encodes a transactivator of gene expression. *Journal of virology, 68*(6), 3943-3954.

Wright, H. T., Jr., Goodheart, C. R., & Lielausis, A. (1964). Human cytomegalovirus. morphology by negative staining. *Virology, 23*, 419-424.

Wu, J., O'Neill, J., & Barbosa, M. S. (2001). Late temporal gene expression from the human cytomegalovirus pp28US (UL99) promoter when integrated into the host cell chromosome. *The journal of general virology, 82*(Pt 5), 1147-1155.

Wynn, Katherine K., & Khanna, Rajiv. (2006). Models of CMV infection. *Drug discovery today: disease models, 3*(1), 91-96.

Xing, Yang, Medvin, D., Narasimhan, G., Yoder-Himes, D., & Lory, S. (2011, 3-5 Feb. 2011). CloG: A pipeline for closing gaps in a draft assembly using short reads. Paper presented at the computational advances in biology and medical sciences.

Yow, M. D., Williamson, D. W., Leeds, L. J., Thompson, P., Woodward, R. M., Walmus, B. F., . . . Griffiths, P. D. (1988). Epidemiologic characteristics of cytomegalovirus infection in mothers and their infants. *American journal of obstetrics and gynecology, 158*(5), 1189-1195.

Yu, Dong, Silva, Maria C., & Shenk, Thomas. (2003). Functional map of human cytomegalovirus AD169 defined by global mutational analysis. *Proceedings of the national academy of sciences, 100*(21), 12396-12401.

Yu, X., Trang, P., Shah, S., Atanasov, I., Kim, Y. H., Bai, Y., . . . Liu, F. (2005). Dissecting human cytomegalovirus gene function and capsid maturation by ribozyme targeting and electron cryomicroscopy. *Proceedings of the national academy of sciences, 102*(20), 7103-7108.

Zerbino, Daniel R., & Birney, Ewan. (2008). Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome research, 18*(5), 821-829.

Zhang, W., Li, H., Li, Y., Zeng, Z., Li, S., Zhang, X., . . . Zhou, T. (2010). Effective inhibition of HCMV UL49 gene expression and viral replication by oligonucleotide external guide sequences and RNase P. *Virology journal, 7*, 100.

Zhang, X. W., Li, F., Yu, X. W., Shi, X. W., Shi, J., & Zhang, J. P. (2007). Physical and intellectual development in children with asymptomatic congenital cytomegalovirus infection: a longitudinal cohort study in Qinba mountain area, China. *Journal of clinical virology 40*(3), 180-185.

APPENDIX A

Buffers and Solutions

1. Dulbecco's Modified Eagle's Medium (DMEM) (1L)

(With phenol red, L-glutamine, 4.5 g/L glucose, but without sodium pyruvate)

DMEM (GibcoBRL, USA)	1 packet
Sodium bicarbonate, NaHCO3 (FW 84.01; Sigma, USA)	2 g
100 IU/ml Penicillin (Gibco, USA)	0.1 g
100 IU/ml Streptomycin (Gibco, USA)	0.1 g
Hepes	10 mM

All chemicals were dissolved in 800 ml of sterile deionized water and made up to 1 L with sterile deionized water. The solution was stirred for 4 h and sterilized by filtration (0.22 µm) and kept at 4˚C.

2. Phosphate Buffered Saline (PBS) (1L)

Sodium Chloride, NaCl (FW58.44; Merck, Germany)	8 g
Disodium hydrogen phosphate, Na_2HPO_4 (FW141.96; Merck, Germany)	1.15 g
Potassium chloride, KCl (FW74.55; Merck, Germany)	0.2 g
Potassium dihydrogen phosphate, KH2PO4 (FW136.08; Merck, Germany)	0.2 g

All chemicals were dissolved in 800 ml of sterile deionized water. The solution pH was adjusted to 7.2 with 1 M HCl or 1 M NaOH. The solution is then made up to 1 L with sterile deionized water. The mixture was autoclaved at 121˚C, 15 min. After sterilization, PBS was stored in 4˚C.

3. Antibiotic-Trypsin-Versene (ATV) (1L)

Sodium Chloride, NaCl (FW58.44; Merck, Germany)	8 g
Disodium hydrogen phosphate, Na_2HPO_4 (FW141.96; Merck, Germany)	1.15 g
Potassium chloride, KCl (FW74.55; Merck, Germany)	0.2 g
Potassium dihydrogen phosphate, KH_2PO_4 (FW136.08; Merck, Germany)	0.2 g
Trypsin 1:250 (Amresco, USA)	0.5 g
Glucose (Sigma, USA)	1 g
Ethylenediaminetetraacetic acid (EDTA), $C_{10}H_{16}N_2O_8$	0.2 g
Phenol red (Merck, Germany)	0.02 g
Sodium bicarbonate, NaHCO3 (FW 84.01; Sigma, USA)	0.58 g
100 IU/ml Penicillin	0.1 g
100 IU/ml Streptomycin (Gibco, USA)	0.1 g

All chemicals were dissolved in 800 mL of sterile deionized water and made up to 1 L with sterile deionized water. The solution was stirred for 4 h. It is then sterilized by filtration (0.22 µm) and kept at 4˚C.

4. Sucrose Gradient

Surcrose solution at different concentration is prepared with PBS buffer pH7.4 as follows:-

Solution A: 20%(W/V) Sucrose solution
10g Sucrose (Merck, Germany) in 50ml PBS

Solution B: 30%(W/V) Sucrose solution
15g Sucrose (Merck, Germany) in 50ml PBS

Solution C: 40%(W/V) Sucrose solution
20g Sucrose (Merck, Germany) in 50ml PBS

Solution D: 50%(W/V) Sucrose solution
25g Sucrose (Merck, Germany) in 50ml PBS

Solution E: 60% (W/V) Sucrose solution
30g Sucrose (Merck, Germany) in 50ml PBS

Three quarter of the PBS buffer volume added to the sucrose.
Stir with a sterile magnetic stirrer continuously until all sugar dissolve.
Transfer into a measuring cylinder and top up to 50 ml. Label and keep refrigerate.

APPENDIX B

Virus titer determination by $TCID_{50}$

Calculation of infectivity titer of RCMV ALL-03 by $TCID_{50}$ according to the method of Reed &Muench (1938).

Dilution	Infected	Non Infected	Cumulative effect		Ratio	Percentage
			CPE	No CPE		
10^{-1}	4	0	17	0	17/17	100%
10^{-2}	4	0	13	0	13/13	100%
10^{-3}	4	0	9	0	9/9	100%
10^{-4}	3	1	5	1	5/6	83.3%
10^{-5}	2	2	2	3	2/5	40%
10^{-6}	0	4	0	7	0/7	0
10^{-7}	0	4	0	11	0/11	0
10^{-8}	0	4	0	15	0/15	0

End point = $\frac{\text{Infectivity above 50\%} - 50\%}{\text{Infectivity above 50\%} - \text{below 50\%}}$

$= \frac{83.3-50}{83.3-40}$

$=0.8$

$TCID_{50}$=Dilution where CPE >50% +0.8 =4.8µl ⟶ The titer of RCMV ALL-03 suspension is $10^{4.8}$ $TCID_{50}$ /0.1 ml

APPENDIX C

Sequencing data obtained for gap 1, gap 2, gap 3, gap 4 and gap 5 for gap closure purpose

Gap 1	First Trial	Forward	ANNNNNNNGCGNATACGCGTTNTATGGAAACGAATCGACCCGAGCATACATCGTCACCGATGAGGCG CAGTGCATCGATGGATATACAACGATGTCGTTCTCGTTCTCGTTCACGTTCACGTTCTCGTTCTCGCTC TCGTTCTCCTTCTCGTTCTCGTTCTTATTCCTTTATTCGTTGGCGCTCGCATTCTCATATTCGGTTGCCAT ATGACGTTTCAAAATTGACACGATAAACCATTTTCATGATGTAATACTTTTTATTATGTATATTTGTAC AATGGTAGGTGTAACACAACAATTTTTACAGAAAACTCACAGAGACCCATAAAGAGGGTGACTAATA AGGTTATAGCACTCAATACAGGAAAAAAAAGGTTGATATTCACTAAGAATAGGTTAGCAAGAATACC ATTGATATAATGCAGTTTGTTTAGATTAACATACGATAAGAATTCACGTACGCCTTTCCGACGATGATT ATGTGTCTCTCGTTCTCGTTCTCGTTCA(Total=490)
		Reverse	GGNNNGNAGNGTAAGAGACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAGA ACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAG AACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGA GAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGAGAACGATTTCG AGAACGAGTACGAGTACGATATCTAGATCTAGATCTAGATCGATATCGAGTTCGAGTTTGATTTGGGT TTGGGTTTTGGTTTCGTTTTGATTTTGATTTTGTTCTTGTTTCTGTTCTTCTTTATTATCTTTATTGTTATT GTTAGAGTGGTCCTGATTCTGATAGTCTGTCTCTGGGTGGTGTACACCTAATCGTATTCTTATTTTTATC TTTATATTACTACTCGTTTGGTTATTTCTGTTTTTGTCTTTGTCTTTGTCTTTGTCTTTGCCTTTGTCTTTG TTATTGCGNATACTGTTGCGGTTGGGGCTGGGTTTATGGGTGTGTTTGTCGTTGTGGTT (Total=490)
	Second Trial (Middle Up)	Forward	NANTNTTCATTAACTTATAATCAGCACAAGATCCGCCCACACAAAGCGATAATAAAACTCTTAATAAA TCACAATGAACGACATTCGTCGTCAAATCATTTTTTACACACGACTTTCGTAGTTGCCGGTAGTTTATT TCTCTTTACAAACGCCTTTACTTTTGCCCTTTCCGCGTGTAGACGGTTTAGACGACTTATCTTCTGTGGA TACTTTATCTTCGTCTGATGACGATTTACCAAGTATTTGTTTATCTAAAGCCTCATTCATCATTTTATTC ATGCGGTCTATGTCATCAAATGTGGTGATGTGTTGTTCTTCATCGTCTCCATCGTCCGAATCAGAGTCG

			TCGCTATCTTCACCCATAACATAGTGAAAACGTCTCTTACGTTCGGCAGCGACTTCTTCGTCAAAACGG TTTTGCTCTTCTGGAGTGAGCCCATCATAGAATTTAACGACCTTTTTCATATCTTTTACAGCCATACGA CCGGCATACGATGCCATAATCTCAAAACAGACTTCCCAAAGTTCCCCTTTATCATCCATTATCACTAAC CCGAATTACACACGCCCACACAGCCCGGCAAAAAACGAGCGTATATATAGACAACCGAAAACGTTGT TAGATATTATTCTGCTTCTCAGTCAACTCAGACACCGCTTTATATAGAACTTGCATGTACCCGGAAACC TTTCTCGATCAACGACAGAACTAACCAACCCCCAACACATTAACCGCGTATACGCGTTAATATGGAAA CGAATCGACCCGAGCATACATCGTCACCGATGAGGCGCAGTGCATCGATGGATATACAACGATGTCG TCCCCGCTTCTCGAGACTCGAGAA(Total=900)
		Reverse	CNNNTTGNATCATCGATGCACTGCGCCTCATCGGTGACGATGTATGCTCGGGTCGATTCGTTTCCATAT TAACGCGTATACGCGGTTAATGTGTTGGGGGTTGGTTAGTTCTGTCGTTGATCGAGAAAGGTTTCCGG GTACATGCAAGTTCTATATAAAGCGGTGTCTGAGTTGACTGAGAAGCAGAATAATATCTAACAACGTT TTCGGTTGTCTATATATACGCTCGTTTTTTGCCGGGCTGTGTGGGCGTGTGTAATTCGGGTTAGTGATA ATGGATGATAAAGGGGAACTTTGGGAAGTCTGTTTTGAGATTATGGCATCGTATGCCGGTCGTATGGC TGTAAAAGATATGAAAAAGGTCGTTAAATTCTATGATGGGCTCACTCCAGAAGAGCAAAACCGTTTTG ACGAAGAAGTCGCTGCCGAACGTAAGAGACGTTTTCACTATGTTATGGGTGAAGATAGCGACGACTCT GATTCGGACGATGGAGACGATGAAGAACAACACATCACCACATTTGATGACATAGACCGCATGAATA AAATGATGAATGAGGCTTTAGATAAACAAATACTTGGTAAATCGTCATCAGACGAAGATAAAGTATC CACAGAAGATAAGTCGTCTAAACCGTCTACACGCGGAAAGGGCAAAAGTAAAGGCGTTTGTAAAGAG AAATAAACTACCGGCAACTACGAAAGTCGTGTGTAAAAAATGATTTGACGACGAATGTCGTTCATTGT GATTTATTAAGAG GAACCACTACCGACGCCGCGAAAAAGACACATTCTCTGTGCGTCTCCAGCTATCTGACGATCGAGCGT TCTCATACTTTTATTATCGCTTTGTGTGGGCGGATCTTGTGCTGATTATAAGTATAAAAGAAAAAATTT TCTTTTGCTTTAAGNTGGTGGTGTCACA (Total=900)
	(Middle Down)	Forward	GGGNNANATCTTGTAATGTCTCGTTCTGGTACTGGTTCTAGTTTTGGTTATGGTTAGGGGCGGGAGAT AGTTTTTATTATAGTTCTAGGTCTTGGGTGAGTTCTGGTTCTCGGTCGGGTTCTTGATCTCGTTCTCGTT CTCGTTCTTGTTTTGGTTCTCGTTCTGGTTTTGGTTTTCGTTGTCGCCCACTCTCCCGATCCCCACAGTTC CCATTACGATATCTACAGTCCCCGCAGCCGATCTTGGTCGAATTACCGGATACGTCCTCGGTCCTTTGC

			ATGTACTCGACGGGGAATGAGGGATCGGGTGTGCCTGTGTTGGTCATTCCCCGTTATCATTCGTTTTGG TTGTACAGAGACGGGGTGAATTAGAAAATAAAGCGATGCCCGTATTAGTACTGGAATCGGGATAACG TCATCGCTTCTACAGAAAGGGAACTTACACAAAACGAAACGTATCGAAAAATAGAGATATCGAATCG TTACGTGTTTTTGTTTATTTCACATTAGAGGGCAACTAGGTCAACTCGTTTATATGCCCGTAGATATGT CCGAATATTTCTCCAGATCACATCTCAATCCCTGTATGTTTCACGAATAACGGTACACGTTTCACGGTG CACATCACCTATCTTCAGTGTCTCCATCACCGTCTCTCTCTCATCCTCGTCCACTTGACCCCAAACCTAT CATCCCGATACGACGTTCGTCGAATTATCAGATAGGTGTAAATGTAATACAATACAAACGAAGTGCCA CGAGAACGATCGCTGTAAACTCTTCTCGTCGATTCGGGCCGGACCAAGCTCGGACGCGTTCATGTCAG TACGGCGAACGGCTGGAACACCGAACCTCCCGTCCGCAGTACAGCGGCCGGAACAGCTAACTTCCCA CCCACGGACAGAACTTTGCGAGAACGAA (Total=900)
		Reverse	CNNNTTGNATCATCGATGCACTGCGCCTCATCGGTGACGATGTATGCTCGGGTCGATTCGTTTCCATAT TAACGCGTATACGCGGTTAATGTGTTGGGGGTTGGTTAGTTCTGTCGTTGATCGAGAAAGGTTTCCGG GTACATGCAAGTTCTATATAAAGCGGTGTCTGAGTTGACTGAGAAGCAGAATAATATCTAACAACGTT TTCGGTTGTCTATATATACGCTCGTTTTTTGCCGGGCTGTGTGGGCGTGTGTAATTCGGGTTAGTGATA ATGGATGATAAAGGGGAACTTTGGGAAGTCTGTTTTGAGATTATGGCATCGTATGCCGGTCGTATGGC TGTAAAAGATATGAAAAAGGTCGTTAAATTCTATGATGGGCTCACTCCAGAAGAGCAAAACCGTTTTG ACGAAGAAGTCGCTGCCGAACGTAAGAGACGTTTTCACTATGTTATGGGTGAAGATAGCGACGACTCT GATTCGGACGATGGAGACGATGAAGAACAACACATCACCACATTTGATGACATAGACCGCATGAATA AAATGATGAATGAGGCTTTAGATAAACAAATACTTGGTAAATCGTCATCAGACGAAGATAAAGTATC CACAGAAGATAAGTCGTCTAAACCGTCTACACGCGGAAAGGGCAAAAGTAAAGGCGTTTGTAAAGAG AAATAAACTACCGGCAACTACGAAAGTCGTGTGTAAAAAATGATTTGACGACGAATGTCGTTCATTGT GATTTATTAAGAGTTTTATTATCGCTTTGTGTGGGCGGATCTTGTGCTGATTATAAGTATAAAAGAAAA AATTTTCTTTTGCTTTAAGNTGGTGGTGTCACA (Total=900)
	Third Trial	Forward	NGGTTGAGTACGACCTTTTTCTATCTTTTACAGCCATACGACCGGCATACGATGCCATAATCTCAAAAC AGACTTCCCAAAGTTCCCCTTTATCATCCATTATCACTAACCCGAATTACACACGCCCACACAGCCCG GCAAAAAACGAGCGTATATATAGACAACCGAAAACGTTGTTAGATATTATTCTGCTTCTCAGTCAACT

			CAGACACCGCTTTATATAGAACTTGCATGTACCCGGAAACCTTTCTCGATCAACGACAGAACTAACCA ACCCCCAACACATTAACCGCGTATACGCGTTAATATGGAAACGAATCGACCCGAGCATACATCGTCAC CGATGAGGCGCAGTGCATCGATGGATATACAACGATGTCGTTCTCGTTCTCGTTCACCATCGCGATCC AGATACCGGNAAAAAACCAAA (Total=500)
		Reverse	NNGGTAGTAGAGACGAGACGACATCGTTGTATATCCATCGATGCACTGCGCCTCATCGGTGACGATGT ATGCTCGGGTCGATTCGTTTCCATATTAACGCGTATACGCGGTTAATGTGTTGGGGGTTGGTTAGTTCT GTCGTTGATCGAGAAAGGTTTCCGGGTACATGCAAGTTCTATATAAAGCGGTGTCTGAGTTGACTGAG AAGCAGAATAATATCTAACAACGTTTTCGGTTGTCTATATATACGCTCGTTTTTTGCCGGGCTGTGTGG GCGTGTGTAATTCGGGTTAGTGATAATGGATGATAAAGGGGAACTTTGGGAAGTCTGTTTTGAGATTA TGGCATCGTATGCCGGTCGTATGGCTGTAAAAGATATGAAAAAGGTCGTTAAATTCTATGATGGGCTC ACTCCAGAAGACAAAACCGTAA (Total=480)
Gap 2	First Trial	Forward	NNAANAAATCAACATATGACTTACATGTTGGCAGGTATTACTTGATGTGGGCGCTTGTATGTTACTAA AAAATTTGAGACTTATAGTGTTTTATATATTTATTGTCACCTTTATCCCATAGAGATGGTGGTACGGTT TGTTCCAAAAGAATAAGGTCTTTGGATGTACATAACGTGAGCACGGATATATCGTAGGCACAACTACA ATTCACGTATGATGTGCTGAGGGAATATATCAATCAGGCTTTGAGAAACATTTTGGATGCTTGGTATT CGGGTTTGTTGAATGTCGTATCTCGGCTAAAGTTCAATAAATTACGTAAGTGTCTGGATGTGTAACCA CATCAAACGCATATGGAAATTCCGTGTCCTGTACTCCCGTATCCCGTACCATGTCTAGATGTATGACCG CATCATCCACAAATGGAAATTCCGTTTGTTGACTCGTTTCCTGTACTGAAGTTGCTCGGTAAATAATTA TTCAGTAAACGATATACACGGACATAAAAAAACGATATGATACGTGACAGCGATTATTACTCGTTATC TAATTGATATTGCATAATAAATTTATGACTAGTATAGAATTATTATGCGTCTAATTCTTTCGCATATGT GGATGAAAAAATCGGCCATCGGCAGACATATCCTTTCTCTCAATGTAAATATAACCAATATAAATATA ATCAAAAATAACATTAAGAAACTGATAAGTATTCCATAATTTCCTCGACCCGATACGTACTTATTGTG TGTAAAATGAGATGTTGTGCTAATAGTACTCGGCATAATTGTATTCACCTTTCTGACGATGGTATTCGG GTTTGTTGG (Total=800)

		Reverse	NNNNAAAAACGTCCTAGATATATGTATCCACCATACGTTTCATTTTTTAATTTAAAATAGTGGCGCCCT CTATTGTTGATATTGTTAATTTATAAATAGGAAAGGTTCTTGGCACCGGATTGGGTTTCGTTAATTCGG CTTCGCAGCGGATGATTGTTCGAAAACCTTCCAATGCCTGAAGAGATTTTGGACTGATACGCTGCCAA AAAACAAGAAGCCCAGCATTCGTGATGTACATCTCAGGATATCCAATCTTATCGTGGGTTTCGTTAAT TAGGCTCATATCGTTTTTTTATGTCCGTGTATATCGTTTACTGAATAATTATTTACCGAGCAACTTCAGT ACAGGAAACGAGTCAACAAACGGAATTTCCATTTGTGGATGATGCGGTCATACATCTAGACATGGTAC GGGATACGGGAGTACAGGACACGGAATTTCCATATGCGTCTGATGTGGTTACACATCCAGACACCTAC GTAATTTATTGAACTTTAGCCGAGATACGACATTATCCACAATACGTTTCGCGTATTGCGGCGAGTAG AAATGGGAACATCGAATCTCGGTCCTTAAATTTTAATGTTTGCTCGATTACCCAGACCCGTTTTTGAAT GTGTTTGCTTTGTGTAGTAAATCTGTATACACGTAATCGGGAACTCAGATAAAAAGGACTAACAATTA TATTAGCACAGACTCTTGTTTTTGGTGTCTATCTCTGCGCATCGTTAATAGCTACATATTCCCACGGAC ATGAAGACGGCTTTCCTTTNTGTGGCGACAATTGCGTTACTATCTCCGTCGTACGGGTTTCGTTAATTA GGCTAAAA (Total=800)
	Second Trial (Middle Up)	Forward	NCTCAGCATACGTGTTGTATATTCTGTTATATTCTATTATATGTGCTCTGAACTTCGTTGTTCTTTTATG TGATGCAGTTTCAAAACCGTCTTATGTATAGATCTTCTCGGGACGACGATTAGCACATCTACCAGCCT GCTATATGGATAATTACATAAAGATCTCTCCCCGGTCTAAGATTTCAAAGGGTCCCTGTAGTCGTTCCC TTCGGTGCCGTGCCGCCCCGTAGATCGAATTGTATTCACTGCGTATGCTCACTTCTTGACAAACTCCTG TTTGTACTGATTTTCGTCAATAATCGCGCCAGACGAGGTCACCGCTATTCGACATGTTCTCGAGAGTGC CTCTGTACACGTCTTCCATCTCGCGAGCCTGTATGATACGTCATCGTCGTTGTCCATTTGCTGTTACTGT TGGGAACCACTCGGGAGATCCTCTGCACGTGACCTGGATCCGTTCCGCAGTCTTCGTATACAGGTCGA ACAACGGGTTCGAGAGTCACAGCTTCGCGCCGAAGCCTAACGGGACTCTGCCGGAAGGACGGAACGC TCCCCTTTCTATAGACCCGGTTCTCCATAAGGCTGTTCATTTCGATGGGACGGAGGCTAACGTTGTTGT TAATTTATACCC (Total=600)
		Reverse	NNNNNNNNNNGGAANGTTNGNNANNCTAGATGCTGCGGATGGTGATGAACGAAAGATACGTATCA CTACCGTGTTAAAAACTGTACATAATTGTAAATAAAGAAAATGCGAGTCGTGAAATTTTTTATTCATC GTTCACTTTATTGTGCCGAGGACCTGGTTTTGTGAAGTGTACTTGGGATTTGGGGATGTGTAGGGAGG

			TTATGATGCTGGTGAGTGCCGGCGAGAACTGGGTTATACCGGCAAACGTAAATNGGAAAATCCTCCTA GATCACATATTCCCGGGTGTTCACACGTTGACGGTGTCGAAATTTGATCCTGTTTTTAACTGATAGGGG TCCTGGGGTATATGTTAGTTATCATAAACAANACACTGCAGCTCCTCACCTACCATCGTGGGGTGGAG GTGTCCATATATATCAATCGANTACCCTCACCGACAAAGGTACCCTTTCACCCCTACGCCATGCAGGG TTCCTACCACACCGGGCCACGTACCCCNTTGTTTAGCAGNNGACGACCCTCCGCTCGGCCAGTACACC AGTTGTTGGGCGTATGGCATTTCTCTCATTCTGGGTATCCTTGAGCTNTTGAAATAAAAAGGATTGTTT NNGGTNAATTNTTNN (Total=580)
	(Middle Down)	Forward	TNNNNNNNNNNNNATNTNNANTNNNNGNTTAGGAAGGTCTTTGTAGTGNACCTGCGATCGGACGGTN TACCGGGTGGGGTGATGAGATACGACAGTACCTTTGTTCTGAGAATGAGATCCGCTACCTGGATAGAA CGTATAGGGGATCGTGTGATGTCGCTAGCCATGAAGCGATTTTCTCCGGAATCGTTACCTAGCGATTG GTTCCGCCACATGATCGATCCGTGTCTGACAGGGGAGGATATCTCCTCTCTTGATGGGGAGTGTATCG TAGGCGATCGTCAGAAAGCGCGGCCTCCGCCGTTTCTCCCTTTCACGGTCTTACTTATCACGGGCACCG CCGGCGCCGGTAAGACTTCCAGCGTTCAGGTACTAGCCGCTAATCTGGATTGTGTGGTCACCGGGAGC ACGGTTATCTCCTCCCAGGCTTTGAGTTCTGCGTTGAATAGGACGAAATCTGCACAAATTAAGACTAT ATTCCGTACTTTCGGATTTAATAGTAAACACGTTGCCTTGGCCGATTGCGTTCATCTGCGTAAAAAGGA TGATGTCTCGTTCGACGGCGTCGAACCCATATGCGAACAACAATGGCGAGATTTGTCGGTGTATTGGT CCGTTATATCGGATATCGCCAATAATGCCCTCAAGGCCGGAAAGGGTGCGCGGGATACCATGGATAT GTGTCAGAGTAATATCATAGTCGTTGATGAATCCGGGACTATACTCAGACACATGCTACACGTCATTG TGTTCTTCTACTACTTCTACAATGCCCTGAATAACAGTGATCTGTATAAGAAATGTGCCGTGCCGTGTA TAGTTTGCGTGGGGTCCCCTACGCAGTCGGAGGCTCTTGNAATGTCGATTCGATCACCATTCTCAGAA CAGAAACATNCAGAGAGGTGTAGACGTGTTATCTGCGCTCTTAGCGATCCGGTTTGTCGGAATTTTGT TATGGTTNGGAAAATTTAAAAAAANNNANANTTN (Total=950)
		Reverse	NNNNNNNNNNCNCATNNNANCTTCANAGGTTGTCAGTTNACTCACTCAGNNGGAGGGGCGGCCTNC CATTCCTAACCTACCATATTGGCAGTTAGATTAGACTGTGCTCTAGTCATAGTATGGCTTTGCGCTGTC TGTCATTGAGACTGGAGCGGAACATATTTGATCCCGGCCGCCTGGCACTCATCCTCAAAAGAGCCCNC TTGAAACAATGTCCGAGGAAATAACCCATGATTACTAGCCTGACTTGTACGTGGTGTTATTTCTTTAAT

			TGTTGCTTTAAATTTTTCAAAAACAAGTTTTCTCAAGTCTTCCAAATTAAATGCTTTATCTAATTCATTA TTCAGATCTGGAGGGAAGAATTAGGATATTCTTACACGTTGTTCTTATGGATTTTTCTACTCTCCAATC CTTGCTATGACACTACGTGTCATTCTCCCAGAATCATAGTTTATTGATCGGTCCAGCTCCTGGAAACAT GTTCCCTTCCTTATATGCTATTTTTTCTTGGTTCCTGTTATACTTGAGATTCAGGGANTCATCGTGAAGG NACGCACTTTTGGAGAATCGATGAATGGGGCGGGGTTTTATTCTAACGGAAAGTCAACAGTTCAGTGN AGCTATTACGGATCATCGTCTGCATTTAGTTGTTCACCGGTGGAGTNGTCGTACTGCCAAAATCGAAT AGACGGGTCGACTGCGTATTGATACATGCAATAGGNTGCCGTTTCTGCTCAGTCCGTCTGCTCTGAAC GCAAGCCGACACCAAACTNCCCNAGTANGACGATCAGGGCCACGTTTGTATNNGCTGTACCAGGAAN CGCNCNTNNNTTAGNACGANAGNNTCCNNAGNCCTATNNGACTCTCCTGCNAGCGNGAGNCACGNNN (Total=950)
	Third Trial	Forward	AAATCGCATTATGTAAAATGCTGATTCTACCTCTGTTCATAAACGAGTAAGGTAGTTCCGCTACTGAG TATGCTTCTTTTAATTGATTTAATTATCTAAACCTACATAGTACTTAAAAATATACATAAGCAAATGAA GTATGAGGAACTAAGAAGCCTAATTAACGAATCCCTTTTGAAACAGTTAAACGTTATGGAATAAACCC TATTAATATCCTATATCGTAATTCGAGCACGTGTATATCCCTTTCCAAGTGTACCCCTCCCTAACCCCC CCCTAACCACGATCTTTTCCGCTGTGCGCACGGAAACCCCGTTCCCGGAGAATGTATGGTCGTGGATT ATATGGGCTTTTTCAATTTTCCCGCACAGTACCAGGCTTGTTGGTTCGGTCAGTCAGACCAGAGATTTT CTTCACAAAGAATTCCATCCTCAAAGTAACTGACATCCTAGGTATTAATCTCACCAATTTCACTAACAT ATCTGGTGATCGCCTTATGTCACAGCTTAAGTTATTCATTGGTTCTGGTTATTAAGGGATGCTTGGCAA CTCGACTTACAACTAAAATACCAACGCTTTCAGTTCCCCACCTCAGTCGCNGTGTCCGGTCTGAGAAA CCTGAAACAGTCCTGAGTCCTGTCCAACATCAACATCAATACAGCC (Total=400)
		Reverse	GCCGGGGTCTTCATGTACATCTTTTAGCCCCTTCTCCATGCCCCCCGCATGTCTAAGTTGTCTGTGCTTT TATGTGAAACTCAACGTCCTTTTTTTTTTCGCATATTAAATCAACCTTATTGCTTGATTTTTTGGGTGGC CGGACCTTGGACCGGGGCACGGGACCGGGGAACGGGGCCGGGGAACCGGATGAGCCGAGGAGGCAA GGGTCCGGGAAAGGGACCGGTCAAGGGGACGGGGAGGGGACCGGGAAAGGGACCGGGAAAGGGAC CGGGAAAGGGACCGGGAAAGGGACCGGGAAAGGGACCGGGAAAGGGACCGGGAAAGGGACCGGGA AAGGGACCGGGAAAGGGACCGGGGAATGGACCGGGAAAGGTACCGGGAAATGGATGGGGCAATGGA CCGGGAAAGGCACCGGGAAAGGCACCGGGAAAGGGACCGGGAAACGGACCGGGGAGCCGATTTTTG

			CGGGGATCCATTTTGGAAATAAACCCAGGTTCCCACAGCACCGGCGACACCTGATGAGCATACTACCA AACAAAATCGTTTAGTAGAATTGATGGATGGGATATTGCTACGTACCGGAAATTTTTTCGAATAATTT GACAATCCAATGGACCGTTTAAGGGACTTATTGCCTGATTGTCCGAGCATTCCGATGGCTATTCAATG GATGAGGTAAATTGACGGGCCAGTTTGTGGACGAGCGAGCCTATCACACCGGGTAGCGGAGCAGCTG CGGCTGCCGATGCGCCTGGTGATCCGAGAATTCCGGTGAGCCGAGCACACACTGTTTGCGTTTGTCCA GGTATAGCTTGTATTCCCAGGCTCCAAGTACCAATTTTTTTGAAAATTCGTTGTCATCCGATTGCTCTT GGGATGTTAGCATATGGTGGTATCCAAGCTTTCGGAGAAGTAATTACCTTCAGAGCCGACAGCAGGG GAGCTGAGCACCTGGGGATCTATGAGCAGTGTAGCGAGATACGAGGTAGCGATGAGTCGAGGTAGCC GAGCAGCCGGTGAGTCTAACGTGACACCTCTGATCGCTGACAGGCTTGACCGAATGGCTCCCACGTCT TACCGTACTTATACCCGAATCTAGCGCTGGACTTCTCGCAGTAATACGTTCATGTAACCTGGCGGTTAG GTCGGTAAGGGCAACGGAGGGGCTAGAAAGCTCGAGCTGACCCTTATGTACT (Total=550)
Gap 3	First Trial	Forward	NNNGNACNNNNTTTGTCGAANATCTTAAATTTGTAGATACCCCTGGCATTCTGGATCATAACAACGTT AGCGATCCGGAGACCTTGTTATGGTTACTTTTTTGTGGGCCCCAGAGTCTCTGTCAGAATCCGACCTGT TTCGGTCGTGATCGTGAGTGCGAGTTGTCTTTTCCCGCTTTGTTGCCGCCAGTTTTCTACGATACCGTC ACCGTCTTGGCCGCATACCCTCGGTCCCCGGGTATGCAGCAAAGACGGTGGCGGATTGGTTGTAACAT GGAGGAATCCAAGGGGGAACTGACACCACGCCTCCCCAAGGACAACAGAAACAATGCCGATTAAGA GAGAAAAGGGGGGGTGTGGAAAAGTAACAATTACAAGGTCCCCGGGTCGCTAACGTTGTTATGATCC AGAATGCCAGGGGTATCTACAAATTTTAAGATCTTTTCGACAATAGCCCTCTTGTCCGTGGAGAGGGA TCGTTTAAGTTGAGATCCTTGAATACGAGTTGTCGTATTCCCGCTGTTTTGCCGCCCNCTTNCTTGGAA ACCCTCACCCCCCTACCNCATAC (Total=490)
		Reverse	NNNNNNNNNCGTTTAGTTGCTCTTTAGTATAAACGACGTACGACCCACTCAGCATACGTCCGTCGTAC GACAAAGCCGTCACCGTCTTGGCCGCATACCCGGGGACCGAGGGTATGAAGCCAAGACGGTGACGGC CTTGTAGAACAACGTCACGTCTACTATGTGGATCGAACGTCGTCTAATTCTAACAGAAACTATGCGGC CCACAAACAACGTAACCAGAAGGCGGTTGCCGAATAACTAACGTTGTTATGATCCACAATGCCATGA GTATCTACAAATTTTAAAATCTTTTCGACAATACCCCTCTTGTCATTAGATAGACATCGTTTAACTTGT CCCTCGAGAGGGATCGTTGAAGATGAAAAATTAAAAATAAATTAAACCCCTGGCATTCTAGATCATAC

			AAACGTAGTTAACCTCAAAGCTTCGTAAAGATACTTATTTGGGGGCGACATAGTATCTGTAAGAATCA GACCAGATTCTCCCTTCATCGACATTGCAAATTGTCTTTTCTCGCTTTGTTGCCACATGTTTCTACAATA TCGTCTACGTCCTTCTACNAGAGC (Total=500)
	Second Trial (Middle Up)	Forward	GNCCGCTACGGCCTACTTGGGGATTCTGTAGTTATCACGGCCGACGGCTTATCTCTCTTGTTGAGAATC GTTTCATAGCGCTGTTTTACTATATCGGCGACTGCGTCCGTCAGCATGGCTGCGGAATCGAGTGTGTCT ACGACGCTTTGTGATATAGAGGCCCTTGTGGCCGTGGATGAGGGGAGAGTTCCGGACGCCGATATTAA GAAATACAGGGAGGCTGTGGATGCGGCTCTCGTGGCTTGTGAGGCGTCTTCTCCGCGTGATCGGTTCA GATTAGTTGAGACGGCCGGTGGAAACTTTTTGTTGGTCACGAACGCTTTGCCGAAGGAAAGGTCTGAG CAGACTCAATGTGGCGATACGAGCTTGGTAGGTAGTGAGCGAAACGAGGGTGTCTTCGACGGTCTTTT GTCCTTGAGTGATGATCGTGCTAGCGGGGCCGGTCTTATCGCCTCCATACCCTCGGTCCCCGGGTATGC GGCCAAGACGGTGACGGCTTTGTCGTACGACGGACGTATGCTGAGTGGGTCGTACGTCGTTTATACTA AAGAGCAACTTAAACGATCCCTCTCCACGGACAAGAGGGCTATTGTCGAAAAGATCTTAAAATTTGTA GATACCCCTGGCATTCTGGATCATAACAACGTTAGCGATCCGGAGACCTTGTTATGGTTACTTTTTTGT GGGCCCCAGAGTCTCTGTCAGAATCCGACCTGTTTCGGTCGTGATCGTGAGTGCGAGTTGTCTTTTCCG TTTTTTTTTGCCAAA(Total=700)
		Reverse	ACNTCNTCGATACGACCGAACAGGTCGGATTCTGACAGAGACTCTGGGGCCCACAAAAAAGTAACCA TAACAAGGTCTCCGGATCGCTAACGTTGTTATGATCCAGAATGCCAGGGGTATCTACAAATTTTAAGA TCTTTTCGACAATAGCCCTCTTGTCCGTGGAGAGGGATCGTTTAAGTTGCTCTTTAGTATAAACGACGT ACGACCCACTCAGCATACGTCCGTCGTACGACAAAGCCGTCACCGTCTTGGCCGCATACCCGGGGACC GAGGGTATGGAGGCGATAAGACCGGCCCCGCTAGCACGATCATCACTCAAGGACAAAAGACCGTCGA AGACACCCTCGTTTCGCTCACTACCTACCAAGCTCGTATCGCCACATTGAGTCTGCTCAGACCTTTCCT TCGGCAAAGCGTTCGTGACCAACAAAAAGTTTCCACCGGCCGTCTCAACTAATCTGAACCGATCACGC GGAGAAGACGCCTCACAAGCCACGAGAGCCGCATCCACAGCCTCCCTGTATTTCTTAATATCGGCGTC CGGAACTCTCCCCTCATCCACGGCCACAAGGGCCTCTATATCACAAAGCGTCGTAGACACACTCGATT CCGCAGCCATGCTGACGGACGCAGTCGCCGATATAGTAAAACAGCGCTATGAAACGATTCTCAACAA GAGAGATAAGCCGTCGGCCGTGATAACTACAGAATTCCCCAATGTAGGCGTATTCTGTAAAAAGAGA TATCACACAATGGCCACCCGAAGA

			(Total=750)
	(Middle Down)	Forward	AGACGTTCAAACGTCCCGATTATTCTTCCTACATCAACCTGGCCGAGTTGTATGTGTATGTCTGGTATA AGGATTACGACTTCTCCTCGGAGTCGGCGGGGTGTTACGATTTAGGTGAGGTGGCCATGGACAGGGTC AAGAAGACGTTGGCGTCGGTTTGTGATAGGTTTGGCGATAAGAACGTACCCGTTTGGCCAATATCGTC TCGAATATGCATATTTTGTGCTTTATATAATCAAAACAGGGTATGTCTGGACTTGGCGAAGAACGATA TTAATTTCACCGCGTATAGTCCGATAATCGTAAAAGATTGTCGTGACGCTGCGACAAACGTTACCCTG AGCCACGTTCTGCCTGACAATCGTGCCGCTTCTTTGTTTCCTGTCTATGACATCGGAATTCTATCACGC GTTTTGTGTGATTCTTCCGATGGAGAGGAGCGCAGGAAACGTGTGCGGGAAAACATAGAGTCGGCGA TCAGTTGTTTGGATGACTGATGATGTCGAATCGCACACAGAAAAAGTGGAAGGGGGGA(Total=600)
		Reverse	GGNNNCTAACTCGGTCATCCAAACACTGATCGCCGACTCTATGTTTTCCCGCACACGTTTCCTGCGCTC CTCTCCATCGGAAGAATCACACAAAACGCGTGATAGAATTCCGATGTCATAGACAGGAAACAAAGAA GCGGCACGATTGTCAGGCAGAACGTGGCTCAGGGTAACGTTTGTCGCAGCGTCACGACAATCTTTTAC GATTATCGGACTATACGCGGTGAAATTAATATCGTTCTTCGCCAAGTCCAGACATACCCTGTTTTGATT ATATAAAGCACAAAATATGCATATTCGAGACGATATTGGCCAAACGGGTACGTTCTTATCGCCAAACC TATCACAAACCGACGCCAACGTCTTCTTGACCCTGTCCATGGCCACCTCACCTAAATCGTAACACCCC GCCGACTCCGAGGAGAAGTCGTAATCCTTATACCAGACATACACATACAACTCGGCCAGGTTGATGTA GGAAGAATAATCGGTGACGGTATCGTAGAAAACTGGCGGCAACAAAGCGGGAAAAGACAAAA(Total= 600)
	Third Trial	Forward	NNNGACTCAGTCTTTGTCTTGAGTGATGATCGTGCTAGCGGGGCCGGTCTTATCGCCTCCATACCCTCG GTCCCCGGGTATGCGGCCAAGACGGTGACGGCTTTGTCGTACGACGGACGTATGCTGAGTGGGTCGTA CGTCGTTTATACTAAAGAGCAACTTAAACGATCCCTCTCCACGGACAAGAGGGCTATTGTCGAAAAGA TCTTAAAATTTGTAGATACCCCTGGCATTCTGGATCATAACAACGTTAGCGATCCGGAGACCTTGTTAT GGTTACTTTTTTGTGGGCCCCAGAGTCTCTGTCAGAATCCGACCTGTTTCGGTCGTGATCGTGAGTGCG AGTTGTCTTTTCCCGCTTTGTTGCCGCCAGTTTTCTACGATACCGTCACCGATTATTCTTCCTAAATCAA CCTGAA(Total=600)
		Reverse	GNNGATCGGACGTAGAAACTGGCGGCACAAAGCGGGAAAAGACAACTCGCACTCACGATCACGACC GAAACAGGTCGGATTCTGACAGAGACTCTGGGGCCCACAAAAAAGTAACCATAACAAGGTCTCCGGA

			TCGCTAACGTTGTTATGATCCAGAATGCCAGGGGTATCTACAAATTTTAAGATCTTTTCGACAATAGCC CTCTTGTCCGTGGAGAGGGATCGTTTAAGTTGCTCTTTAGTATAAACGACGTACGACCCACTCAGCAT ACGTCCGTCGTACGACAAAGCCGTCACCGTCTTGGCCGCATACCCGGGGACCGAGGGTATGGAGGCG ATAAGACCGGCCCCGCTAGCACGATCATCACTCAAGGACAAAAGACCGTCGAAGACACCCTCGTTTC GCTCCCTACCTACCAAA(Total=600)
Gap 4	First Trial	Forward	NNNNCATGACTTCGGAGTTATTTGGGCCAACATACGATGTAGTAGATGATCACGATGGCGAAGATGA TTTCGGTGAAAATCACGTAGATCACTAGCTGAACGGGGTCTAAATATAGTTGCGGCGTAAGGTGATGT ATCGCGTTATATGTAGACAGGTTCATGTTGCCGAAATCTATAATTTTTGGATAATAGCACGGAAAACC GGCACCGGGAAAGTGTGCCGCTATGGAAAATATAACGATGTTTACGAACGAGATTAACGATAACACT ATAGACATAGTCCATGTTCTCACGTTTACACGATCTACGTGAGAGAGCGTCATTGTGTTTATCGCCATC GTACCTGCTATCGCGTCGATGGCATCCGTCTATGGGGTCAAATAGTGAAAACTCAGTATGGCGGCGTG TGGTCTTGTGACGTGGTGTGTTACATATCTCATTGGTAGATCATAAT (Total=520)
		Reverse	NNNNANTTTCCACCACGTCACAGACCACACGCCGCCATACTGAGTTTTCACTATTTGACCCCATAGAC GGATGCCATCGACGCGATAGCAGGTACGATGGCGATAAACACAATGACGCTCTCTCACGTAGATCGT GTAAACGTGAGAACATGGACTATGTCTATAGTGTTATCGTTAATCTCGTTCGTAAACATCGTTATATTT TCCATAGCGGCACACTTTCCCGGTGCCGGTTTTCCGTGCTATTATCCAAAAATTATAGATTTCGGCAAC ATGAACCTGTCTACATATAACGCGATACATCACCTTACGCCGCAACTATATTTAGACCCCGTTCAGCT AGTGATCTACGTGATTTTCACCGAAATCATCTTCGCCATCGTGATCATCTACTACATCGTATGTTGGGC CCAAATATACTTCCGAAAAGACAACGGTAATCAGGTAAATCATCTTACGAGAGAA(Total=520)
	Second Trial (Middle Up)	Forward	GGNNNNNNNNCNATTGATTTATCTCATCGTTATCGACCAAGGGTCGGTAACGACGGGCGCGACGTAG AATAAATCGCACCACGCGTATCGTTGTGAATATCACACATAACAACAGTAGGACGGCTAGTGTTATGT TAATATCTTTGGTGTACTCGGAAGCGTTCAGGGCCTCGTATTGTATGATAGGGTACATGGCTCCGCAC ACTCCTAATATGCTTCCGATGTGGTATCCGAACTGGACTTTCATGTACCGCACTAGTATGGACTCGATT ATGGAGAAATAAATTGTCGATACCACCGCGAAGGCGACCATGGCGCCGAATACCACGTGTGCGGTCT TAACGAAAAAGCTGTTCCCGAACCCGAGGGCGAGCGACATTGCCAGCACCATGGTGCTGATTCCGAG CAAGGCTTGGCTCAGATTGACCACCGCGGTCCGGTATCGAACGGTGCCTTGTAGCTTTGGGTGGATTT TAGCGAGATTGAAGGCGCTGCGTTCTGTTGACTCATAGTGGGTGACGAACGTGACCACAAAAGCCGTT

			AAGCATATGAAGTACATGCATTTAGCGAAGGCGACCATGGACGGGAATCTGAACGAGAGGCTGAGGA CGTAAATCTGGAACGCGTCCATAGTTAGGGTAAAGGCAAAGCATGACGCAGTATCTCCTATGCAACTG ATGTCTCTCGTAGACTGATTTACCTGATTACCGTTGTCTTTTCGGAAGTATATTTGGGCCCAACATACG ATGTAGTAGATGATCACGATGGCGAAGATGATTTCGGTGAAAATCACGTAGATCACTAGCTGAACGG GGTCTAAATATAGTTGCGGCGTAAGGTGATGTATCGCGTATTGTTTGAACACGGAANNAAA (Total=700)
		Reverse	ANNANTGNCGNACTATATTTAGACCCCGTTCAGCTAGTGATCTACGTGATTTTCACCGAAATCATCTTC GCCATCGTGATCATCTACTACATCGTATGTTGGGCCCAAATATACTTCCGAAAAGACAACGGTAATCA GGTAAATCAGTCTACGAGAGACATCAGTTGCATAGGAGATACTGCGTCATGCTTTGCCTTTACCCTAA CTATGGACGCGTTCCAGATTTACGTCCTCAGCCTCTCGTTCAGATTCCCGTCCATGGTCGCCTTCGCTA AATGCATGTACTTCATATGCTTAACGGCTTTTGTGGTCACGTTCGTCACCCACTATGAGTCAACAGAAC GCAGCGCCTTCAATCTCGCTAAAATCCACCCAAAGCTACAAGGCACCGTTCGATACCGGACCGCGGTG GTCAATCTGAGCCAAGCCTTGCTCGGAATCAGCACCATGGTGCTGGCAATGTCGCTCGCCCTCGGGTT CGGGAACAGCTTTTTCGTTAAGACCGCACACGTGGTATTCGGCGCCATGGTCGCCTTCGCGGTGGTAT CGACAATTTATTTCTCCATAATCGAGTCCATACTAGTGCGGTACATGAAAGTCCAGTTCGGATACCAC ATCGGAAGCATATTAGGAGTGTGCGGAGCCATGTACCCTATCATACAATACGAGGCCCTGAACGCTTC CGAGTACACCAAAGATATTAACATAACACTAGCCGTCCTACTGTTGTTATGTGTGATATTCACAACGA TACGCGTGGTGCGATTTATTCTACGTCGCGCCCGTCGTTACCGACCCTTGGTCGATAACGATGAGATA AAATCACTACGCGGCGACACGGAATAAAATTCCCCCCTCCCCTAAAA(Total=700)
	(Middle Down)	Forward	NNNNGGGGTGTTCGCGTTATATGTAGACAGGTTCATGTTGCCGAAATCTATAATTTTTGGATAATAGC ACGGAAAACCGGCACCGGGAAAGTGTGCCGCTATGGAAAATATAACGATGTTTACGAACGAGATTAA CGATAACACTATAGACATAGTCCATGTTCTCACGTTTACACGATCTACGTGAGAGAGCGTCATTGTGT TTATCGCCATCGTACCTGCTATCGCGTCGATGGCATCCGTCTATGGGGTCAAATAGTGAAAACTCAGT ATGGCGGCGTGTGGTCTTGTGACGTGGTGTGTTACATATCTCATTGGTAGATCATAAATAGCGGAGGA CGGTGTCTATGAAGTATCCACTCGGGTTTTGAAAACGCGATGGACAGACAGTGGCTGTCCGGTGCGAT CGCTCAGGTCTCGGTCTACAGTTTTCTCCCCGAAGTCAACGAAGCGATCCTGCAGTGTCTCTTCCTCGA GGCCGAGGAAAACGATCGGGTGGCTTCTCGTGTTCTGGTGTTTTCGGTTCAGGAA (Total=650)

		Reverse	GNNNACCGAATCGTTTTCTCGGCCTCGAGGAGAGACACTGCAGGATCGCTTCGTTGACTTCGGGGAGA AAACTGTAGACCGAGACCTGAGCGATCGCACCGGACAGCCACTGTCTGTCCATCGCGTTTTCAAAACC CGAGTGGATACTTCATAGACACCGTCCTCCGCTATTTATGATCTACCAATGAGATATGTAACACACCA CGTCACAAGACCACACGCCGCCATACTGAGTTTTCACTATTTGACCCCATAGACGGATGCCATCGACG CGATAGCAGGTACGATGGCGATAAACACAATGACGCTCTCTCACGTAGATCGTGTAAACGTGAGAAC ATGGACTATGTCTATAGTGTTATCGTTAATCTCGTTCGTAAACATCGTTATATTTTCCATAGCGGCACA CTTTCCCGGTGCCGGTTTTCCGTGCTATTATCCAAAAATTATAGATTTCGGCAACATGAACCTGTCTAC ATATAACGCGATACATCACCTTACGCCGCAACTATATTTAGACCCCGAANGA (Total=600)
	Third Trial	Forward	NNCCTCATGCACTGATGTCTCTCGTAGACTGATTTACCTGATTACCGTTGTCTTTTCGGAAGTATATTT GGGCCCAACATACGATGTAGTAGATGATCACGATGGCGAAGATGATTTCGGTGAAAATCACGTAGAT CACTAGCTGAACGGGGTCTAAATATAGTTGCGGCGTAAGGTGATGTATCGCGTTATATGTAGACAGGT TCATGTTGCCGAAATCTATAATTTTTGGATAATAGCACGGAAAACCGGCACCGGGAAAGTGTGCCGCT ATGGAAAATATAACGATGTTTACGAACGAGATTAACGATAACACTATAGACATAGTCCATGTTCTCAC GTTTACAA (Total=400)
		Reverse	GGGGNNANNNGCACGTTATCTCGTTCGTAACATCGTTATATTTTCCATAGCGGCACACTTTCCCGGTG CCGGTTTTCCGTGCTATTATCCAAAAATTATAGATTTCGGCAACATGAACCTGTCTACATATAACGCGA TACATCACCTTACGCCGCAACTATATTTAGACCCCGTTCAGCTAGTGATCTACGTGATTTTCACCGAAA TCATCTTCGCCATCGTGATCATCTACTACATCGTATGTTGGGCCCAAATATACTTCCGAAAAGACAAC GGTAATCAGGTAAATCAGTCTACGAGAGACATCAGTTGCATAGGAGATACTGCGTCATGCTTTGCCAT TTACCCAAAAN (Total 500)
Gap 5	First Trial	Forward	GGGNATTNNACCGAGGAATCCCGCTTTGGGGAATCTCCGGATCGACGGAGCCTTCGGAGCGGAGGTA TCTCCGCATCGGTGAAGACTCCGCCTCGGAGAGGTCTCCGAAGCGAGGCAATCCTCGGATCGGTACGG TCTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGT CTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGTCGCCGTATCGGATCCGCAGTATCCAGCACGGGAA AACTCCTGCGCGGGGAATCACCATCATAGGTGCAATTCCTCGGAGCGGAGTTTTTTCCGTACAAGAGC

			AGGTCTGAAACTGCTTCGGTACTCTATACCGCGCCGTTCTTCCTTGGTATTCATCTTTTACGGTCAGTC ACACCGTGTAATTTGTTTGTTATTTGTAATTTTTTACAGGTCTCGGTGACGCCCATCTACCTGTAACAG GCTTCCCATCCTTCCTTCATCCAA (Total= 500)
		Reverse	CGNTACNATAAATGAGAGTCCTGAGATCTGTAAAATTACAAATAACAAACAAATTACACGGTGTGAC TGACCGTAAAAGATGAATACCAAGGAAGAACGGCGCGGTATAGAGTACCGAAGCAGTTTCAGACCTG CTCTTGTACGGAAAAAACTCCGCTCCGAGGAATTGCACCTATGATGGTGATTCCCCGCGCAGGAGTTT TCCCGTGCTGGATACTGCGGATCCGATACGGCGACCGCTCCGAGACGGAGACCGCTCCGAGACGGAG ACCGCTCCGAGACGGAGACCGCTCCGAGACGGAGACCGCTCCGAGACGGAGACCGCTCCGAGACGGA GACCGTACCGATCCGAGGATTGCCTCGCTTCGGAGACCTCTCCGAGGCGGAGTCTTCACCGATGCGGA GATACCTCCGCTCCGAAGGCTCCGTCGATCCGGAGATGTCCCCAAAGCGGGGATATCCCCGGGTCGGA GACTACATCGATCACATAGTTTTTACCGCC (Total=550)
	Second Trial (Middle Up)	Forward	AGGCCTACGGGTGTTCGCTCGTTGTCGTCTGACGGCGTTCTTAATCGAGAGCGCCTTTTATACACGGCG GCGTCGGCGGGGCGATGATGATTTTACATACATTTGATACTTTAATATTAACATTTATCACATAATGAC GTTTTGTGACAAACGCTGTTTTTGTAGAATCACGTGATGTACGTACCGAGGATGCGGACACATCGCAA TAGAAAATGGTCTTTTATTCGGAGTGGTCTCCGATACGGGGCGATCTTCATATTACTCTCCCCGCGTCG GTGGGCACTCTGAATCAAAGTCGCCATAATCTATTCGTTATCGCCGGTCTCGTCCACGTTCCTACTCGG ACTCTTATCGCTGTTCTGTATTGCGCTGTTGTCCGACGTGGTCCCTTTGGGGGATGAAGGTGCTTTGTA CACTCACCTCGTGGCGGTAAATCTATGTGATCGATGTAGTCTCCGACCCGGGGATATCCCCGCTTTGG GGACATCTCCGGATCGACGGAGCCTTCGGAGCGGAGGTATCTCCGCATCGGTGAAGACTCCGCCTCGG AGAGGTCTCCGAAGCGAGGCAATCCTCGGATCGGTACGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGG AGCGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGGAGCGGTCTCCGTCTCGG AGCGGTCGCCGTATCGGATCCGCAGTATCCAGCACGGGAAAACTCCTGCGCGGGGAATCACCATCAA GTGGCAATTTCCAA(Total=500)

		Reverse	NNCNNCCAGTACTTTCCGTGCTGGATACTGCGGATCCGATACGGCGACCGCTCCGAGACGGAGACCG CTCCGAGACGGAGACCGCTCCGAGACGGAGACCGCTCCGAGACGGAGACCGCTCCGAGACGGAGACC GCTCCGAGACGGAGACCGTACCGATCCGAGGATTGCCTCGCTTCGGAGACCTCTCCGAGGCGGAGTCT TCACCGATGCGGAGATACCTCCGCTCCGAAGGCTCCGTCGATCCGGAGATGTCCCCAAAGCGGGGAT ATCCCCGGGTCGGAGACTACATCGATCACATAGATTTACCGCCACGAGGTGAGTGTACAAAGCACCTT CATCCCCCAAAGGGACCACGTCGGACAACAGCGCAATACAGAACAGCGATAAGAGTCCGAGTAGGA ACGTGGACGAGACCGGCGATAACGAATAGATTATGGCGACTTTGATTCAGAGTGCCCACCGACGCGG GGAGAGTAATATGAAGATCGCCCCGTATCGGAGACCACTCCGAATAAAAGACCATTTTCTATTGCGAT GTGTCCGCATCCTCGGTACGTACATCACGTGATTCTACAAAAACAGCGTTTGTCACAAAACGTCATTA TGTGATAAATGTTAATATTAAAGTATCAAATGTATGTAAAATCATCATCGCCCCGCCGACGCCGCCGT GTATAAAAGGCGCTCTCGATTAAGAACGCCGTCAGACGGACAACGAGCGAACACCGGTGATCGCGAC AAAAAATTCAGGAGGGGGGTGATTAAAA (Total= 600)
	(Middle Down)	Forward	TNCGTTTACGGGCTTTTCGTACAGAGCAGGTCTGAACTGCTTCGGTACTCTATACCGCGCCGTTCTTCC TTGGTATTCATCTTTTACGGTCAGTCACACCGTGTAATTTGTTTGTTATTTGTAATTTTTTACAGGTCTC GGTGACGCCCATCTACCTGTAACAGGCTTCCCATCCTATCTTCATCCTTTTTATAGGCGGTGGTATATG TGGGCGGAGGTGACCTTAACTATGCAATCCCCACGGTCCACTTACGGTAAGTGGTTAGATGTTCGTGA TTTATATTGTAATTATCGATAGAGTTCGTATATTTTCTTTAGCAAACATGGGGTTGCATCTTCTTTGTGA CTAGGTGTCAAAATATATAATTTTTTTTCGTCGACACATCAGGCCTATATGTGTCGACGAAAAAAGTA AACATCTCGCATTGACGAAACGGATTCATAGCAGAATCGGGTGAGTATACGTTGTTTCTCGCTAGGAA TCGGAGCCGTCTCCGAACCACGTCGACCGTTACAAGAAACAACGTCATGCATAAATTACAAAATGGG CGTGCTGAAAGTTTATAAATGAAGAGTTACTCAGAAACTGGTTTAGTCTTCTGCGCTCTCCGCTGCTAA AAGCATAAGATAGGGTTCCGTGGGTGGTAAAAGTTATATTTAAGCTTATTAACCCCGGGAACTAAA (Total=800)
		Reverse	NNNNGNTTCACCACGGNACCCTATCTTATGCTTTTAGCAGCGGAGAGCGCAGAAGACTAAACCAGTTT CTGAGTAACTCTTCATTTATAAACTTTCAGCACGCCCATTTTGTAATTTATGCATGACGTTGTTTCTTGT AACGGTCGACGTGGTTCGGAGACGGCTCCAATTCCTAGCGAGAAACAACGTATACTCACCCGATTCTG CTATGAATCCGTTTCGTCAATGCGAGATGTTTACTTTTTTCGTCGACACATATAGGCCTGATGTGTCGA

			CGAAAAAAAATTATATATTTTGACACCTAGTCACAAAGAAGATGCAACCCCATGTTTGCTAAAGAAA ATATACGAACTCTATCGATAATTACAATATAAATCACGAACATCTAACCACTTACCGTAAGTGGACCG TGGGGATTGCATAGTTAAGGTCACCTCCGCCCACATATACCACCGCCTATAAAAAGGATGAAGATAG GATGGGAAGCCTGTTACAGGTAGATGGGCGTCACCGAGACCTGTAAAAAATTACAAATAACAAACAA ATTACACGGTGTGACTGACCGTAAAAGATGAATACCAAGGAAGAACGGCGCGGTATAGAGTACCGAA GCAGTTTCAGACCTGCTCTTGTACGGAAAAAACTCCGCTCCGAGGAATTGCACCTATGAGGNGGAATT CCCAAA(Total= 750)
	Third Trial	Forward	NNNNNGGAATGGTCGCAAGTGATTTCACCGGACCGGGTGTGTATGTAAATACAGAAACAGTAAAATT TGGCATGAAACATAGAGACGACAAAATGATTTTCGTTCTGCGAAGATGGGTCGTCGAGGCAGAGAAA ACTTACCAATGCAGCTTTAGACCGTTTTCATGTACACAGACCGGACTAAGTCGCAAATTCCTGCTTTCT CATAGCGCTCTCAACGTAGCGACCGCTAATATGGATGCTATACGCATGTACCCATACGGAGTTGCAAT AGAAAACGCTAAAATCTGCTGCTCGTTAGTAGGTCCCATCGTACAGCAAATCTCATGGTACAGATCTG GAAACTGTGGATCCCCCTCATTGATATATCAGTATAAAACTAATTCTAGTATAGATATTGGAAGCAAG CCGACAACTTTGGACAACGTGAAGGCTTTTTATGTAAGTGGAATTCCAAAACCTAACGAAGGGTATGT CATCTTGTGTTTACAATCATTTAAGTCAAATATGTCCCTTGTTCCGGATGTACCTGTACCCCCCGGTGA ATATCAGTGCCAAGTAGATTATGATAGTCCAATTGAACAACAAACTTTTGAAAAATCTATAACCATTC AGGATGATGTTGATGTAAAAGTAATACATGGTCAGAAAAGTTACGTTGTAATTTGCACTTTTCCGTAA AAGAGCAGGGAN(Total=750)

		Reverse	NNNNNNNACTTTTCTGACATGTATTACTTTTACATCAACATCATCCTGAATGGTTATAGATTTTTCAAA AGTTTGTTGTTCAATTGGACTATCATAATCTACTTGGCACTGATATTCACCGGGGGGTACAGGTACATC CGGAACAAGGGACATATTTGACTTAAATGATTGTAAACACAAGATGACATACCCTTCGTTAGGTTTTG GAATTCCACTTACATAAAAAGCCTTCACGTTGTCCAAAGTTGTCGGCTTGCTTCCAATATCTATACTAG AATTAGTTTTATACTGATATATCAATGAGGGGGATCCACAGTTTCCAGATCTGTACCATGAGATTTGCT GTACGATGGGACCTACTAACGAGCAGCAGATTTTAGCGTTTTCTATTGCAACTCCGTATGGGTACATG CGTATAGCATCCATATTAGCGGTCGCTACGTTGAGAGCGCTATGAGAAAGCAGGAATTTGCGACTTAG TCCGGTCTGTGTACATGAAAACGGTCTAAAGCTGCATTGGTAAGTTTTCTCTGCCTCGACGACCCATCT TCGCAGAACGAAAATCATTTTGTCGTCTCTATGTTTCATGCCAAATTTTACTGTTTCTGTATTTACATAC ACACCCGGTCCGGTGAAATCACTTTTTGCGACCATCTGTCCATCGGCGGCAATATGAAAATCGCCCCC GAAAN (Total=750)

☆ Primers for each gaps have been listed down in Table 3.1

APPENDIX D

Complete genome sequence as appear online in GenBank database

Rat cytomegalovirus ALL-03 isolate Malaysian, complete genome

GenBank: KP967684.1

LOCUS KP967684 197958 bp DNA linear VRL 23-JUN-2015
DEFINITION Rat cytomegalovirus ALL-03 isolate Malaysian, complete genome.
ACCESSION KP967684
VERSION KP967684.1 GI:814970590
KEYWORDS .
SOURCE Rat cytomegalovirus ALL-03
ORGANISM Rat cytomegalovirus ALL-03
Viruses; dsDNA viruses, no RNA stage; Herpesvirales; Herpesviridae; Betaherpesvirinae; Muromegalovirus; unclassified Muromegalovirus.
REFERENCE 1 (bases 1 to 197958)
AUTHORS Balakrishnan,K.N., Abdullah,A.A., Camalxaman,S.N., Quah,Y.W., Abba,Y., Hani,H., Loh,H.S., Kamal,F.M., Zeenathul,N.A., Aini,I., Omar,A.R., Noordin,M.M. and MohdAzmi,M.L.
TITLE Complete Genome Sequence of Rat Cytomegalovirus Strain ALL-03 (Malaysian Strain)
JOURNAL Genome Announc 3 (3), e00451-15 (2015)
PUBMED 26044413
REMARK Publication Status: Online-Only
REFERENCE 2 (bases 1 to 197958)
AUTHORS Balakrishnan,K.N.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-2015) Veterinary Pathology and Microbiology, Faculty of Veterinary Medicine, University Putra Malaysia, Serdang, Selangor 00601, Malaysia
COMMENT ##Assembly-Data-START##
Assembly Method :: CLC Genomic Workbench v. 4.7.2
Sequencing Technology :: Illumina
##Assembly-Data-END##

BIODATA OF STUDENT

The student, Krishnan Nair S/O Balakrishnan was born on 24th December 1989 at Kuala Lumpur. He received his primary education at Sekolah Rendah Jenis Kebangsaan Tamil (SJKT) Kajang, Selangor and his secondary education at Sekolah Menengah Kebangsaan Tinggi Kajang (Form 3) and Sekolah Menegah Kebangsaan Seri Titiwangsa Kuala Lumpur (Form 5). He continued his sixth form of secondary education in Sekolah Menengah Kebangsaan Maxwell, Kuala Lumpur. He obtained his first degree, Bsc. (Hons) Microbiology in 2012 from the Faculty of Science and Technology, University Kebangsaan Malaysia. Upon completion of his studies, he worked as social research officer at a private NGO. In year 2013, he pursued his Master's degree at the Faculty of Veterinary Medicine, Universiti Putra Malaysia, under the supervision of Prof. Dato'. Dr. MohdAzmi Lila.

LIST OF PUBLICATIONS

Krishnan Nair Balakrishnan, Ashwaq Ahmed Abdullah, Siti Nazrina Camalxaman, Quah Yi Wan, Yusuf Abba, Homayoun Hani , Hwei San Loh, Farina Mustaffa Kamal, Zeenathul Allaudin Nazariah, Ideris Aini, Noordin Mohamed Mustapha and Mohd Azmi Mohd Lila 2015. Complete genome sequence of rat cytomegalovirus strain all-03 (Malaysian strain).Genome Announcements.(Published)

Krishnan Nair Balakrishnan, Ashwaq Ahmed Abdullah, Yusuf Abba, Jamilu Abubakar Bala, Faez Firdaus Jesse Abdullah, Farina Mustaffa Kamal, Zeenathul Allaudin Nazariah, Ideris Aini, Noordin Mohamed Mustapha and Mohd Azmi Mohd Lila 2015.Closing the Gaps in Rat Cytomegalovirus ALL-03 (Malaysian Strain) Genomic Scaffold.American Journal of Animal and Veterinary Sciences.(Published)

Ashwaq Ahmed Abdullah, Krishnan Nair Balakrishnan, Yusuf Abba, Faez Firdaus Jesse Abdullah, Zeenathul Allaudin Nazariah, Rasedee Abdullah, Noordin Mohamed Mustapha and Mohd Azmi Mohd Lila 2015. RCMV ALL-03 Model and Study of CMV Pathogenesis in Congenital Infection.American Journal of Animal and Veterinary Sciences. (Published)

UNIVERSITI PUTRA MALAYSIA

STATUS CONFIRMATION FOR THESIS / PROJECT REPORT AND COPYRIGHT

ACADEMIC SESSION : _____

TITLE OF THESIS / PROJECT REPORT :

CLOSING THE GENOMIC GAPS OF RAT CYTOMEGALOVIRUS ALL-03 IN GENOME SCAFFOLD

NAME OF STUDENT : KRISHNAN NAIR A/L BALAKRISHNAN

I acknowledge that the copyright and other intellectual property in the thesis/project report belonged to Universiti Putra Malaysia and I agree to allow this thesis/project report to be placed at the library under the following terms:

1. This thesis/project report is the property of Universiti Putra Malaysia.
2. The library of Universiti Putra Malaysia has the right to make copies for educational purposes only.
3. The library of Universiti Putra Malaysia is allowed to make copies of this thesis for academic exchange.

I declare that this thesis is classified as :

*Please tick (√)

☐	**CONFIDENTIAL**	(Contain confidential information under Official Secret Act 1972).
☐	**RESTRICTED**	(Contains restricted information as specified by the organization/institution where research was done).
☐	**OPEN ACCESS**	I agree that my thesis/project report to be published as hard copy or online open access.

This thesis is submitted for :

☐	**PATENT**	Embargo from_____ until _____ (date) (date)

Approved by:

_____	_____
(Signature of Student)	(Signature of Chairman of Supervisory Committee)
New IC No/ Passport No.:	Name:
Date :	Date :

[Note : If the thesis is CONFIDENTIAL or RESTRICTED, please attach with the letter from the organization/institution with period and reasons for confidentially or restricted.]