CHARACTERIZATION AND IMMUNOGENETIC PROFILING OF HUMAN CYTOMEGALOVIRUS INFECTION IN A STUDY COHORT FROM NORTHERN MALAYSIA

SURUTHIMITRA OKPOLUAEFE A/P OKPOLUAEFE GODSPOWER UDEZI

UNIVERSITI SAINS MALAYSIA

2025

CHARACTERIZATION AND IMMUNOGENETIC PROFILING OF HUMAN CYTOMEGALOVIRUS INFECTION IN A STUDY COHORT FROM NORTHERN MALAYSIA

by

SURUTHIMITRA OKPOLUAEFE A/P OKPOLUAEFE GODSPOWER UDEZI

Thesis submitted in fulfilment of the requirements
for the degree of
Master of Science

August 2025

ACKNOWLEDGEMENT

First and foremost, I thank God for His unfailing protection, and for granting me the opportunity. I extend my deepest gratitude to my beloved mother, Thavamani Ramasamy, for her unwavering support throughout this journey. A heartfelt thanks goes to my late uncle, Mr. R. Rajamani Ramasamy, whose presence and influence continue to guide me in spirit. I hope I have made him proud. I am immensely grateful to my supervisor, Dr. Norfarazieda Hassan, for her steadfast guidance and mentorship, which helped me build confidence in working independently in the lab and pursuing knowledge at a higher level. I truly feel fortunate to have come across her advertisement on LinkedIn.

My deep appreciation goes to everyone at the Advanced Medical and Dental Institute, especially Ms. Hanis Nabilah Mohd Nazman, Ms. Nurin Hanisah Mohd Yatim, Ms. Nurulanis Farhah Jamil Razif, Ms. Mo Yaxin, Ms. Mugguna Balasubramaniam, the Emerging Infectious Disease research group, and the helpful CRL 2 and 4 staff. Special thanks also go to my badminton buddies and everyone who supported me along the way. I deeply thank Associate Professor Dr. Wan Zuhainis Saad from Universiti Putra Malaysia, whose passion for teaching and research inspired my curiosity and love for the exploration of the sciences.

Finally, I thank myself. Thank you, Sue, for persevering through long nights in the lab, overcoming countless obstacles, and never giving up. You did it!

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

>	More than
<	Less than
≥	More than and equal to
≤	Less than and equal to
°C	Degree celsius
%	Percentage
™	Trademark

LIST OF ABBREVIATIONS

HCMV	Human cytomegalovirus
HHV-5	Human herpes virus-5
kbp	Kilobase pairs
bp	Base pairs
CDC	Centre for Disease Control and Prevention
NK cells	Natural killer cells
KIRs	Killer immunoglobulin-like receptors
SNPs	Single nucleotide polymorphism
ELISA	Enzyme-linked immunosorbent assay
SSP-PCR	Sequence specific primer-polymerase chain reaction
IL-10	Interleukin-10
POI	Population of interest
USA	United States of America
HIV	Human immunodeficiency virus
AIDS	Acquired immune deficiency syndrome
IgM	Immunoglobulin M
IgG	Immunoglobulin G
QNAQT	Quantitative nucleic acid quantification tests
IFN	Interferon
IFN-γ	Interferon-gamma
IL-12	Interleukin-12
IL-15	Interleukin-15
IL-18	Interleukin-18
IL-7	Interleukin-7
IL-21	Interleukin-21
TNF-α	Tumor necrosis factor-alpha

WT/WT	Homozygous NKG2C wildtype genotype
WT/DEL	Heterozygous NKG2C wildtype/deletion genotype
DEL/DEL	Homozygous NKG2C deletion genotype
rsID	Reference SNP cluster ID
NPC	Nasopharyngeal carcinoma
MHC	Major histocompatibility complex
VL9	9 amino acid peptide with valine at position 1, and leucine at position 9
HLA	Human leukocyte antigen
E*01:01	HLA-E*01:01 allele
E*01:03	HLA-E*01:03 allele
E*01:03:01	HLA-E*01:03:01 allele
E*01:03:02	HLA-E*01:03:02 allele
E*01:04	HLA-E*01:04 allele
E*01:02	HLA-E*01:02 allele
EBV	Epstein-Barr virus
PBMC	Peripheral blood mononuclear cells
WT	NKG2C wildtype allele
DEL	NKG2C deletion allele
AMDI	Advanced Medical and Dental Institute
EDTA	Ethylenediaminetetraacetic acid
RU/mL	Relative units per milliliter
RU	Relative units
µL	Microliter
mL	Milliliter
mM	Millimolar
ng	Nanograms
nm	Nanometer
HD	Healthy donor
POD	Peroxidase
TMB	3,3’,5,5’-tetramethylbenzidine

R^2	Coefficient of determination
v/v	Volume by volume
Catalogue ID	Catalogue identification number
g	Gravitational force
rpm	Revolutions per minute
PK	Proteinase K
PBS	Phosphate buffered saline
PBS-PK	Phosphate buffered saline-proteinase K solution
DNA	Deoxyribonucleic acid
PCR	Polymerase chain reaction
T_A	Annealing temperature
dNTP	Deoxynucleotide triphosphate
$MgCl_2$	Magnesium chloride
TAE	Tris-Acetate-EDTA
ATCC	American Type Culture Collection
DMEM	Dulbecco's Modified Eagle's Medium
RPMI-1640	Roswell Parks Memorial Institute media-1640
FBS	Fetal bovine serum
DMSO	Dimethyl sulfoxide
V	Cramer's V coefficient
p	P-value
2x2	Two-by-two
HAC	Haldane-Anscombe correction
IgG POS	Anti-HCMV Immunoglobulin G positive
IgG NEG	Anti-HCMV Immunoglobulin G negative
µM	Micromolar
mM	Millimolar
95% CI	95% confidence interval
HM	Homozygous for alternate allele
HT	Heterozygous for alternate allele
X	Homozygous for reference allele

LIST OF APPENDICES

PENCIRIAN DAN PEMPROFILAN IMUNOGENETIK JANGKITAN SITOMEGALOVIRUS MANUSIA DALAM KUMPULAN KAJIAN DARI MALAYSIA UTARA

ABSTRAK

Sitomegalovirus manusia (HCMV) merupakan virus β-herpes yang lazim dan biasanya tidak mendatangkan bahaya kepada individu yang sihat, tetapi boleh menyebabkan komplikasi kesihatan yang serius dalam populasi yang mempunyai sistem imun terjejas. Meskipun sel pembunuh semula jadi (NK) NKG2C+ yang dipandu oleh HCMV merupakan komponen tindak balas imun yang dikenali terhadap jangkitan HCMV, masih sedikit yang diketahui tentang variasi genetik yang mempengaruhi proses ini dalam populasi Asia Tenggara, khususnya di Malaysia. Jurang pengetahuan ini menghadkan pemahaman kita tentang faktor imunogenetik khusus populasi yang mempengaruhi kebolehjangkitan jangkitan HCMV di Malaysia, terutamanya memandangkan bukti terkini bahawa HCMV berkemungkinan memberi kesan yang bermanfaat dalam konteks klinik tertentu seperti dalam kalangan pesakit leukemia. Oleh itu, kajian ini bertujuan untuk mencirikan profil imunogenetik yang berkaitan dengan jangkitan HCMV dalam sebuah kumpulan sihat dari Malaysia Utara, dengan memberi tumpuan kepada reseptor NKG2C (*KLRC2*) dan NKG2A (*KLRC1*) pada sel NK serta HLA-E pada sel yang dijangkiti. Pengesanan tahap Imunoglobulin G (lgG) terhadap HCMV dilakukan menggunakan ujian serapan imuno terpaut enzim (ELISA). Tindak balas rantai polymerase primer jujukan spesifik (SSP-PCR) digunakan untuk menentukan genotip NKG2C (*KLRC2*). Varian wildtype dan delesi NKG2C (*KLRC2*), ekson 2 dan 5 NKG2A (*KLRC1*), serta ekson 3

HLA-E dianalisis melalui penjujukan Sanger. Kajian kami menunjukkan kadar seropositiviti HCMV sebanyak 78.3%, iaitu lebih rendah berbanding laporan terdahulu iaitu sebanyak 92% pada tahun 2012. Agihan genotip NKG2C menunjukkan bahawa WT/WT merangkumi 54.4%, WT/DEL sebanyak 44.1%, dan DEL/DEL sebanyak 1.5%. Agihan genotip HLA-E pula adalah 50.0% untuk E*01:01/01:03, 45.5% untuk E*01:03/01:03, dan 4.6% untuk E*01:01/01:01. Individu seropositif HCMV tidak menunjukkan genotip NKG2C yang spesifik atau polimorfisme nukleotida tunggal (SNP) tertentu bagi varian wildtype dan delesi NKG2C (*KLRC2*), ekson 2 dan 5 NKG2A (*KLRC1*), dan ekson 3 HLA-E. Peranan factor-faktor lain yang mungkin mempengaruhi keputusan perlu diuji dalam kajian masa depan dengan sampel saiz yang lebih besar yang mewakili rakyat Malaysia. Analisis SNP mendedahkan sepuluh SNP, yang mana dua SNP adalah baru; NM002260.4.c*142 C>A dalam varian wildtype NKG2C (*KLRC2*), dan NC_000012.12:g.10449502 T>A (GRCh38.p14) dalam ekson 5 NKG2A (*KLRC1*). Kesimpulannya, kajian ini memberi laporan pertama taburan genotip NKG2C (*KLRC2*) dan HLA-E, bersama taburan profil SNP dalam ekson 2 dan 5 NKG2A, ekson 3 HLA-E, serta varian wildtype dan delesi NKG2C dalam sebuah kumpulan berwarganegara Malaysia. Penemuan ini menawarkan data asas berguna untuk penyelidikan imunogenetik masa hadapan dan menyokong usaha pembangunan imunoterapi yang disasar dan diperipadikan di Malaysia dan serantau.

CHARACTERIZATION AND IMMUNOGENETIC PROFILING OF HUMAN CYTOMEGALOVIRUS INFECTION IN A STUDY COHORT FROM NORTHERN MALAYSIA

ABSTRACT

Human cytomegalovirus (HCMV) is a prevalent β-herpesvirus that typically does not harm healthy individuals, but can lead to severe health complications in immunocompromised populations. While HCMV-driven NKG2C+ NK natural killer (NK) cells expansion is a well-established component of the immune response towards HCMV infection, little is known about the genetic variations that influence this process in Southeast Asian populations, particularly in Malaysia. This gap limits our understanding of population-specific immunogenetics that influence susceptibility of HCMV infection in Malaysia, especially considering emerging evidence that HCMV may have beneficial effects in certain clinical contexts, such as in patients with leukemia. Therefore, this study aimed to characterize the immunogenetic profiles related to HCMV infection in a healthy cohort from Northern Malaysia, focusing on the NKG2C (*KLRC2*) and NKG2A (*KLRC1*) receptors on NK cells, as well as HLA-E on infected cells. The detection of anti-HCMV Immunoglobulin G (IgG) levels was performed using enzyme-linked immunosorbent assay (ELISA). Sequence specific-primer polymerase chain reaction (SSP-PCR) was employed for the determination of NKG2C (*KLRC2*) genotypes. NKG2C (*KLRC2*) wildtype and deletion variants, NKG2A (*KLRC1*) exons 2 and 5, and HLA-E exon 3 sequences were sequenced via Sanger sequencing. Our findings indicate an HCMV seropositivity rate of 78.3%, which is lower than the previously reported 92% in 2012. The NKG2C genotype

distribution showed that WT/WT accounts for 54.4%, WT/DEL for 44.1% and DEL/DEL for 1.5%. HLA-E genotype distribution was 50.0% for E*01:01/01:03, 45.5% for E*01:03/01:03, and 4.6% for E*01:01/01:01. HCMV seropositive individuals did not exhibit distinct NKG2C genotypes, or specific SNPs for NKG2C (*KLRC2*) wildtype and deletion variants, NKG2A (*KLRC1*) exons 2 and 5, and HLA-E exon 3. The role of compounding factors should be investigated in future studies, using a larger sample size that represents Malaysians. SNP analysis revealed ten SNPs, where two SNPs showed novelty; NM002260.4.c*142 C>A within NKG2C (*KLRC2*) wildtype variant, and NC_00012.12:g.10449502 T>A (GRCh38.p14) within NKG2A (*KLRC1*) exon 5. In conclusion, this study provides the first report of NKG2C (*KLRC2*) and HLA-E genotype distributions along with SNP profiles in NKG2A exons 2 and 5, HLA-E exon 3, and both wild-type and deletion variants of NKG2C in a Malaysian cohort. The findings offer valuable baseline data for future immunogenetic research and support efforts in the development of personalized and targeted immunotherapies in Malaysia and the region.

CHAPTER 1
INTRODUCTION

Human cytomegalovirus (HCMV) is a ubiquitous virus, belonging to the family Herpesviridae, and has co-evolved with humans for millions of years. It is also referred to as the human herpes virus-5 (HHV-5). HCMV or HHV-5 has the largest known viral genome at approximately 235 kilobase pairs (kbp) (Dolan et al., 2004). Comparable to other viruses in its family, the immune system is unable to remove the virus in its entirety once primary infection has ensued, establishing lifelong cycles of infection and latency. HCMV is transmitted through the transfer of body fluids, thereby enabling the transfer of the virions through the close association of infected parties with non-infected individuals (**Figure 1.1**).

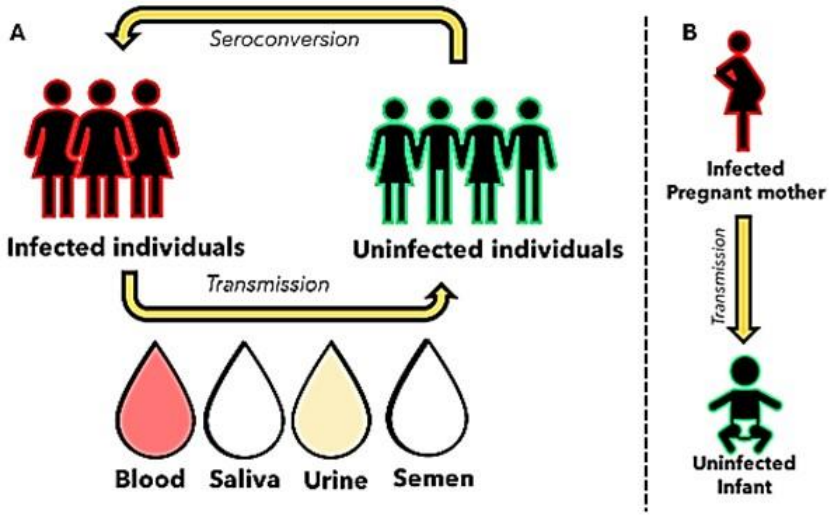

Figure 1.1. Mode of transmission of HCMV. HCMV is transmitted through the dissemination of bodily fluids such as blood, saliva, urine, and sex fluids. HCMV can undergo horizontal transmission of disease (Figure 1.1A), and lateral transmission of disease (Figure 1.1B). Abbreviations: HCMV – human cytomegalovirus [Adapted from (Okpoluaefe et al., (2024)].

In healthy individuals, HCMV infection is generally asymptomatic, granting a small fraction might exhibit subclinical symptoms such as fever, sore throat, fatigue, and swollen glands. However, recent evidence suggests that HCMV may cause breast cancer through the initiation of the giant cell cycle, which results in the generation of polyploid giant cells (Nauclér et al., 2019; Nehme et al., 2021). On the other hand, in the immunocompromised and immunologically immature group, HCMV is of major concern, as HCMV infection or reactivation can lead to end-organ disease (Gerna et al., 2012), severe respiratory infections, and pneumonia (Azoulay et al., 2020). In infants and fetuses, HCMV infection is associated with liver function abnormalities, pneumonia (Mu et al., 2023), HCMV-related sepsis-like syndrome (Namba et al., 2022), hearing impairment (Buxmann et al., 2017; Wang et al., 2022), and hampered neurological development (Pereira et al., 2017). HCMV replicates efficiently in fibroblasts (Wilkinson et al., 2015) but tropism is also present in smooth muscle cells, endothelial cells, macrophages, and epithelial cells (Sinzger et al., 1995) while residing latently in hematopoietic stem cells (Goodrum et al., 2002).

Global HCMV seroprevalence differs across populations and is influenced by socioeconomic factors. For example, the Centre for Disease Control and Prevention (CDC) reported that only about 50% of Americans are expected to be HCMV positive by the age of 40 (Centre for Disease Control and Prevention, 2020), while populations across Africa have been reported to have nearly 100% HCMV seropositivity (Bates & Brantsaeter, 2016). In Malaysia, the seropositivity in healthy Malaysian adults was last reported in 2012 at 92% (Camalxaman et al., 2012). Although the consensus of HCMV seropositivity is acute

causation of disease and poorer outcomes in immunocompromised patients (Bates & Brantsaeter, 2016), they pose potential advantages in healthy individuals, which are not yet extensively studied.

As a direct implication of HCMV infection, previous reports have shown an expansion of a unique set of natural killer (NK) cells, termed adaptive or 'memory-like' NK cells. More recently, tissue residency of adaptive NK cells, albeit limited, has also been described in the liver and lung of patients with primary or metastatic liver cancer, and suspected lung cancer respectively (Brownlie et al., 2021; Marquardt et al., 2015). In most cases, both peripheral, and tissue-resident adaptive NK cells, with varying phenotypes across medical conditions, described with the induction of NKG2-C type II integral membrane protein, NKG2C, CD16 or FcγRIII, and CD2 and suppression of NKG2A, FCεRγ and killer immunoglobulin-like receptors (KIRs) (Okpoluaefe et al., 2024).

The emergence of these subsets in HCMV infection is believed to aid in controlling HCMV in the body. NKG2A *(KLRC1),* and NKG2C (*KLRC2*) from the NKG2 family, are key receptors/genes that are responsible for NK cell activity, while the human leukocyte antigen-E or HLA-E, their cognate ligand, is responsible for antigen presentation. Upon interaction with HLA-E stabilized by HLA type I leader peptides, or UL40 peptide in HCMV infected cells, NKG2A leads to the inhibition of NK cell activity, while NKG2C leads to its activation. In the context of HCMV infection, the interaction between NKG2C and its cognate ligand results in the expansion of adaptive NK cells often characterized by increased expression of NKG2C. This unique NKG2C+ NK cell subset has potential

anti-tumor activity due to its persistence and heightened cytotoxicity, both being important focal points in the development of NK-based oncotherapies.

Despite HCMV being a worldwide phenomenon, its seroprevalence was last described in Malaysia in 2012 at 92%. The lack of recent reports raises questions regarding the current HCMV infection rate amongst Malaysians. Furthermore, given the wide variation in global seroprevalence, shaped by socioeconomic and demographic factors, it is essential to revisit the current status of HCMV infection amongst Malaysians. Unlike previous studies conducted in Kelantan (Ahmed et al., 2006), and Selangor and Kuala Lumpur (Camalxaman et al., 2012), this study focuses on the adult population in Penang, offering insights from a different geographic context. On the other hand, although NKG2A (*KLRC1*), NKG2C (*KLRC2*), and HLA-E are key genes and markers responsible for NK cell activity determination/modulation, the genotypic expression or single nucleotide polymorphisms (SNPs) within their respective regions have not been reported in Malaysia, and much of Southeast Asia. SNPs or variations in genotypic expression within these key gene regions may influence HCMV susceptibility and infection tolerance.

Based on the study by Camalxaman et al., (2012), and a meta-analysis by Zuhair et al., (2019), which estimated the general population's seroprevalence in Southeast Asia, we hypothesize that 80 to 90% of Malaysians are HCMV seropositive. We also hypothesize that the differential genotypic expression of NKG2C (*KLRC2*) and HLA-E will influence HCMV infectibility and infection tolerance, with a distinct difference in the distribution of these genotypes between HCMV seronegative and seropositive individuals. Furthermore,

it is theorized that novel SNPs in HLA-E, and NKG2A (*KLRC1*) will be discovered among healthy people in Northern population. Thus, this study aims to characterize the immunogenetic profiles of HCMV infection in the healthy population, with the objectives, i) To determine the current HCMV infection seroprevalence among a healthy adult cohort from northern Malaysia using anti-HCMV IgG Enzyme-linked immunosorbent assay (ELISA), and its association with age and ethnicity, respectively, ii) To identify NKG2C (*KLRC2*), and HLA-E genotypes from isolated peripheral mononuclear cells (PBMCs) from healthy Malaysian adult cohort from northern Malaysia with respect to HCMV infection via sequence specific primer-polymerase chain reaction (SSP-PCR), and Sanger sequencing, respectively, iii) To identify novel and previously reported single nucleotide polymorphisms (SNPs) in HLA-E, NKG2A (*KLRC1)*, and NKG2C (*KLRC2*) with respect to HCMV seropositivity by using Sanger sequencing, iv) To evaluate the associations of NKG2C (*KLRC2*) genotypes and SNPs identified in HLA-E exon 3 and NKG2A (*KLRC1*) exons 2 and 5 with gender, ethnicity, and NKG2C genotype, and v) To determine NKG2C (*KLRC2*) and HLA-E genotypes of select cell lines that have not been described before, as well as to identify SNPs within NKG2C (*KLRC2*) wildtype and deletion variants, NKG2A (*KLRC1*) exons 2 and 5, and HLA-E exon 3.

CHAPTER 2
LITERATURE REVIEW

2.1. Epidemiology of human cytomegalovirus

Human cytomegalovirus (HCMV), also known as herpes virus type-5 (HHV-5), is a widespread opportunistic pathogen belonging to the Herpesviridae family. It is ubiquitous worldwide, with seroprevalence rates varying significantly by region, population demographics, and socioeconomic factors (Bates & Brantsaeter, 2016; Camalxaman et al., 2012; Centre of Disease Control and Prevention, 2020; Vilibic-Cavlek et al., 2017) (**Table 2.1**). In developed countries, HCMV seroprevalence typically ranges from 40 to 80%, whereas in developing countries and regions in the Global South, seroprevalence can be as high as 100%. Higher seroprevalence rates are often observed in lower socioeconomic settings due to crowded living conditions and limited access to healthcare.

Table 2.1. Socioeconomically developed, developing, and underdeveloped populations and their reported HCMV seroprevalence. Abbreviations: POI – population of interest, USA – United States of America.

POI	Socioeconomy	Seroprevalence	Author(s)
Croatia	Developed	74.4%	(Vilibic-Cavlek et al., 2017)
USA	Developed	50% by age of 40	(Centre of Disease Control and Prevention, 2020)
Malaysia	Developing	92%	(Camalxaman et al., 2012)
Africa	Underdeveloped	~100%	(Bates & Brantsaeter, 2016)

Age, gender, and sexual behavior also significantly impact the spread of HCMV. For instance, women of reproductive age are at higher risk (Kling & Kabelitz, 2015; Siennicka et al., 2017; Stadler et al., 2010). Studies have shown that HCMV seroconversion is correlated with factors such as age (Francisse et al., 2009; Stadler et al., 2010), the serological status of sexual partners (Francisse et al., 2009), the presence of toddlers under the age of 3 (Francisse et al., 2009; Stadler et al., 2010), and being African American (Stadler et al., 2010). Understanding the epidemiological patterns is crucial for developing targeted prevention and intervention strategies, including screening pregnant women and improving hygiene practices in high-risk populations.

HCMV primarily spreads through the exchange of bodily fluids, such as saliva, urine, blood, and breast milk (**Figure 1.1, Chapter 1**). Primary infection generally occurs during childhood or early adulthood, often without noticeable symptoms. After infection, HCMV establishes lifelong latency and reactivates under certain conditions such as stress (Prösch et al., 2000) and in immune suppression (Frantzeskaki et al., 2015; Shimada et al., 2022). Primary HCMV infection occurs when an individual is first exposed to HCMV, whereas HCMV reactivation occurs when the latent virus re-enters its lytic cycle under specific triggers, and becomes active again. This condition typically happens when the immune system becomes weakened or compromised. However, primary infection can still occur in previously infected individuals, as people can be infected with more than one strain of the virus. Some commonly known HCMV strains include the AD169, Towne, Toledo, and Merlin strains, respectively. Different HCMV strains differ in their ability to cause infection and exhibit varying tropism toward various cell types (Cimato et al., 2024).

In individuals with rheumatoid diseases, advanced age, low albumin levels, high creatinine levels, cyclosporin use, and high doses of prednisolone are all risk factors for HCMV reactivation (Shimada et al., 2022). Both cyclosporin and prednisolone are commonly used in the management of rheumatoid diseases. Another study found that HCMV reactivation occurred in approximately 14% of critically ill immunocompromised patients, such as people living with HIV (Frantzeskaki et al., 2015). The inflammatory C-reaction protein levels and the amount of red blood cell units transfused were identified to be independent risk factors (Frantzeskaki et al., 2015). The study also noted that high levels of interleukin-10 (IL-10) were marginally related.

2.2. Mechanism and phases of HCMV infection

HCMV infection consists of 2 phases: the lytic phase and the latent phase. The lytic phase includes a distinct transient quiescence phase, which resembles a brief latent period. Quiescence is reported to last up to 14 days at a time. Primary HCMV infection is established in dendritic cells, endothelial, epithelial, fibroblasts, and smooth muscle cells (Sinclair & Reeves, 2014; Sinzger et al., 2008). CD13 antigen, an integral membrane protein equivalent to aminopeptidase N, is present in all cells permeable to HCMV infection (Söderberg et al., 1993). Although HCMV can infect a wide range of host cells, the site of infection plays a crucial role in determining the nature of the ensuing infection (Matthews et al., 2023). When macrophages, fibroblasts, and epithelial, and endothelial cells are infected, the virus typically leads to the lytic stage, characterized by robust viral transcription and genome replication (Matthews et al., 2023).

Upon infection of these primary sites, HCMV replicates rapidly with a doubling time of approximately 24 hours (Emery & Griffiths, 2000). In primary fibroblasts, the cell cycle state following primary infection plays a pivotal role in determining viral transcription (Zydek et al., 2010). The G0 or G1 cell state allows viral gene expression, whereas the S/G2 phase suppresses or delays gene expression until the next G1 state (Fortunato et al., 2002; Wiebusch et al., 2008) (**Figure 2.1A**). The exact reason for this level of stringency enforced by the cell cycle for viral gene expression is not known. Sinzger et al., (2008) suggested that the establishment of HCMV infection in epithelial and endothelial cells authorizes interhost transmission, while the establishment of the virus in CD34+ bone marrow cells facilitates the systemic spread within the host. They further iterate that the proliferation of the virus in ubiquitous cell types such as fibroblasts, and smooth muscle cells facilitates efficient propagation (**Figure 2.1B**).

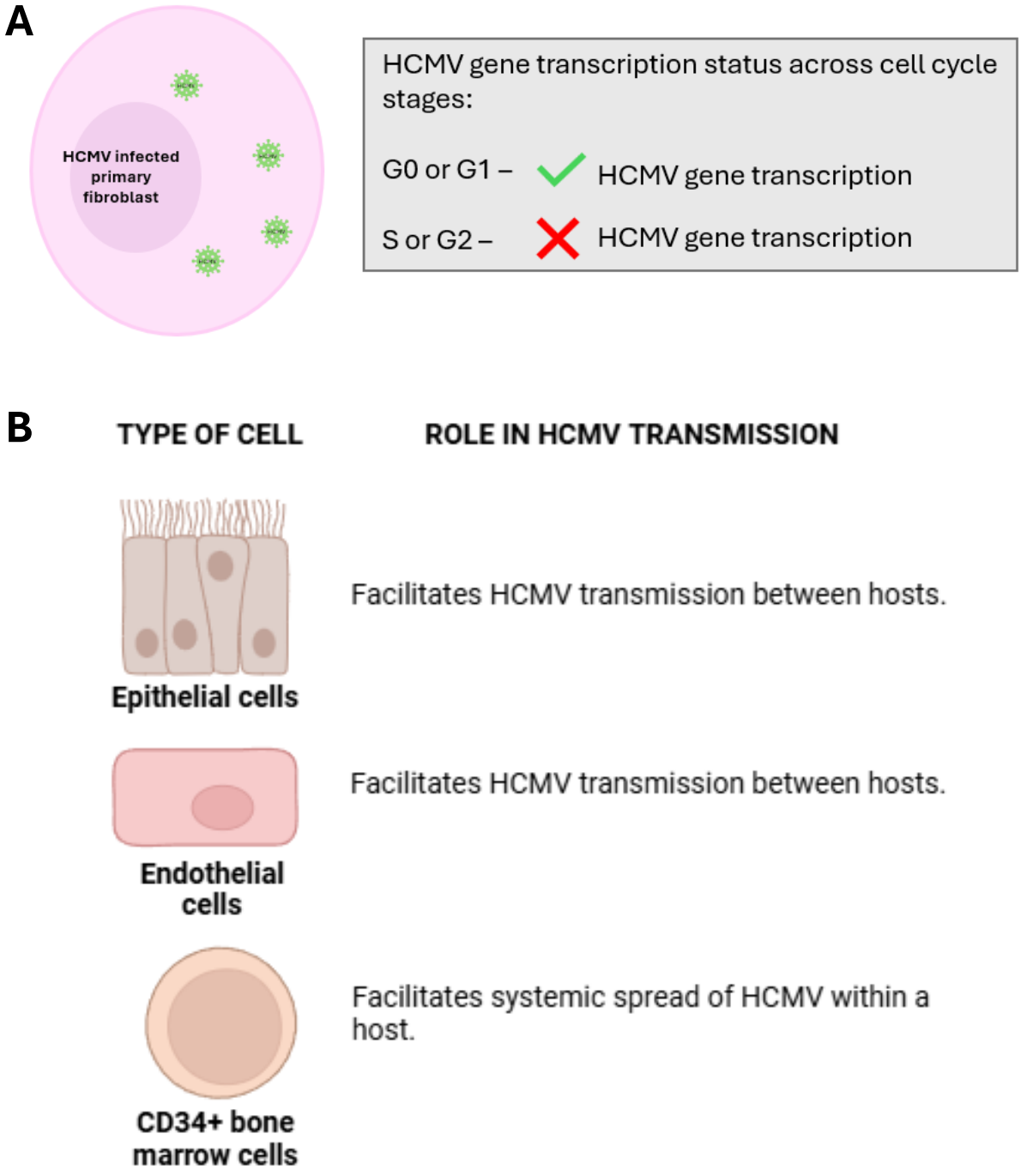

Figure 2.1. Factors influencing gene transcription, and transmission after primary HCMV infection. A) Relationship between cell cycle stage and HCMV gene transcription in infected primary fibroblasts. HCMV gene transcription is active (✓) during the G0 and G1 phases of the host cell cycle but is suppressed (X) during the S and G2 phases. B) Types of cells as site of HCMV infection, and their role in HCMV transmission/spread.

HCMV infection in the progenitor CD34+ hematopoietic stem cells and circulating CD14+ monocytes typically result in latency (Mendelson et al., 1996; Smith et al., 2021; Taylor-Wiedeman et al., 1991), where viral genome replication is usually absent and limited viral transcription occurs (Rossetto et al., 2013). Primary infection in monocytes is seen to cause the unique quiescence phase (Chan et al., 2012; Noriega et al., 2014). However, unlike the long-term latency phase, quiescence in monocytes is limited. The current understanding of HCMV reactivation involves the terminal differentiation of CD34+ progenitor hematopoietic cells and CD14+ monocytes into dendritic cells and macrophages (Sinclair & Reeves, 2014). ULb' and UL138 viral genes are essential for the establishment of latency (Montag et al., 2011). Dendritic cells are a common site where HCMV reactivation occurs (Sinclair & Reeves, 2014). Cells with reactivating HCMV can be identified by the co-expression of CD86 and CD83, cell surface markers for macrophages, and dendritic cells, respectively (Sinclair & Reeves, 2014). HCMV reactivation is triggered by cell differentiation (Matthews et al., 2023) and physical stresses like inflammation (Montag et al., 2011; Prösch et al., 2000) and hypoxia (Wise et al., 2021).

2.3. Clinical manifestations of HCMV infection/disease

In healthy individuals, HCMV infection is typically exhibited subclinically, with symptoms normally ranging from tiredness and fever (Okpoluaefe et al., 2024), with mononucleosis-like syndrome being the most common (Zedtwitz-Liebenstein et al., 2016). Interestingly, Zedtwitz-Liebenstein et al., (2016) further iterated that symptomatic display of HCMV infection in immunocompetent or healthy adults is due to reinfection with

another strain of the virus, the first undistinguishable by the latter in laboratories (Shimamura et al., 2006).

On another note, HCMV poses significant public health challenges, particularly in immunologically vulnerable or immunocompromised people, such as organ transplant recipients, individuals with HIV/AIDS, and pregnant women. Over 75% of organ transplant recipients acquire their first HCMV infection or reactivation of the virus (Fishman et al., 2007; Fishman & Rubin, 1998). In these groups, reactivation or primary infection can lead to severe complications including graft versus host disease leading to end-organ disease (Chan et al., 2001; Griffiths & Reeves, 2021; Wang & Zhao, 2020), organ rejection (Gerna et al., 2012), retinitis (Centers for Disease Control and Prevention, 2024), occurrence and progression of Rasmussen's encephalitis (Wang et al., 2021; Zhang et al., 2017), and congenital infections. One study even went to show that black infants were more likely to have contracted congenital HCMV compared to the non-Hispanic white cohort (Fowler et al., 2018), displaying that one’s ethnicity could be a possible predictor.

Congenital HCMV infection is a leading cause of birth defects and birth, and developmental abnormalities (Buxmann et al., 2017; Fowler & Boppana, 2006; Navti et al., 2021). According to the Centres for Disease Control and Prevention, (2024), approximately 1 in 200 children born infected, and of those, about 20% will have birth defects or exhibit long-term health conditions, such as hearing loss, developmental and

motor delay, vision loss, microcephaly, lack of coordination, and seizures. They further iterate that in severe cases, HCMV infection can cause loss of pregnancy or recurrent spontaneous abortions (Fotoohi et al., 2016).

2.4. Diagnostic methods for HCMV detection

As aforementioned, the virus cannot be completely removed from the body after HCMV infection. According to Razonable et al., (2020), in organ transplantation or hematopoietic stem cell transplantation, assays that directly detect HCMV are advised for HCMV surveillance, diagnosis, and monitoring. Conversely, tests that evaluate immune status, such as anti-HCMV Immunoglobulin M (IgM) or Immunoglobulin G (IgG), are used for assessing risk and stratifying risk factors. During the initial stages of viral infections, the body first produces IgM as an early-stage defense mechanism. This precedes the development of class-switched, high-affinity IgG antibodies (Li et al., 2003). IgG marks long-term immunity, and immune memory towards a pathogen.

The standard method for directly detecting and quantifying the HCMV genome involves using commercially available quantitative nucleic acid quantification tests (QNAQT) (Razonable et al., 2020). Whole blood and plasma are commonly used as specimens for QNAQT, although cerebrospinal and bronchoalveolar lavage fluids are also sometimes used (Beam et al., 2018; Boeckh et al., 2017).

2.5. HCMV management and treatment

At present, drugs for HCMV management and treatment are severely limited. Current antiviral treatments target the replicating phase of HCMV and are effective during reactivation, but do not provide prophylactic protection (Elder & Sinclair, 2019; Heald-Sargent et al., 2020). Combination drugs such as ganciclovir and valganciclovir, and foscarnet and cidofavir, have been used for the control of HCMV, but severe side effects or resistance often develop (Gerna et al., 2019). In 2017, a new drug that targets the HCMV DNA terminase complex, letermovir, was licensed for prophylactic treatment in HSCT recipients. Although its safety and efficacy are promising, long-term side effects and resistance should be investigated (Gerna et al., 2019; Goldner et al., 2011). Currently, there are no approved vaccines for the prevention of HCMV; however, numerous candidates have been identified for potential development. Interestingly, a mixed interventional and observational controlled study investigated the impact of hygiene education on reducing primary infection risk among pregnant women (Revello et al., 2015). The intervention group demonstrated a significant 6.4% decrease in seroconversion rates and fewer cases of congenital HCMV infection in newborns. For the time being, hygiene practices appear to be the safest and most reliable option for HCMV management.

2.6. HCMV-directed immune expansion

HCMV infection drives the expansion of CD8+IFNγ+ T cells (Klenerman & Oxenius, 2016; Sacre et al., 2008), and CD4+CD28- T cells (Pera et al., 2017), and a unique NKG2C+ NK cells. Sacre et al., (2008) reported that HCMV infection resulted in the

expansion of the HCMV pp65, and IE-1-specific CD8+ T cells, but the recovery of the latter subset was critical in the control of HCMV infection/reactivation post-allogeneic stem cell transplantation, where they noted that expansion of the IE-1-directed CD8+ T cells conferred protection towards uncontrolled HCMV replication. An expansion of a similar subset of these protective T cells was also reported in heart and lung transplant recipients (Bunde et al., 2005) and patients living with HIV with HCMV-caused retinitis (Sacre et al., 2005). Notably, HCMV disease appeared solely in heart and lung transplant patients exhibiting strong pp65-directed CD8+ T cell response (Bunde et al., 2005).

Pera et al., (2017) noted that HCMV infection is a driver of the expansion of the highly potent and cytotoxic CD28- CD4+ T cells, which only occurs limitedly in the absence of HCMV infection. Similarly, the expansion of the CD4+ CD28 devoid T cell subsets was reported in HCMV seropositive healthy people (Pera et al., 2018) and in Wegener's disease (Morgan et al., 2011). In these cohorts, the expansion of this T cell subset was highly undesirable, as it correlated with the severity of renal disease, and was associated with significantly increased events of mortality and risk of infection due to the significant immunological burden it carries; three fourth of CD4+ T cells were CD28 null, and the expansion led to low levels of naïve T cells in the circulatory system (Morgan et al., 2011; Pera et al., 2018). CD4+ CD28- T cells, along with the presence of HLA-DRB1*03:01, were also a dominant risk factor for cardiovascular pathology (Bunde et al., 2005). According to Pera, et al., (2018), this subset of T cells is directly responsible for damaging endothelium cells in the heart and instigating cardiovascular events (Pera et al., 2018).

HCMV infection also drives the expansion of a unique subset of NK cells that bridge innate and adaptive immunity, characterized by the expression of NKG2C+ (**Figure 2.2**). These innate cells share adaptive characteristics such as antigenic memory, persistence, high cytotoxicity, and the ability to be transplanted (Okpoluaefe et al., 2024). Conventional NK cells exhibit both the inhibiting NKG2A and activating NKG2C at varying proportions, with the net charges received deciding the effector functions on respective NK cells. The interaction between NKG2C on NK cells and HLA-E stabilized with the VL9 mimic, HCMV-derived UL40 nonamer on HCMV-infected cells, leads to the expansion of these adaptive NK cells (**Figures 2.2** and **2.3**), which are NKG2A dim/null, NKG2C bright, CD56 dim or null, CD16 bright, and have restricted killer immunoglobulin-like receptors (KIRs). As these adaptive NK cells mature, they gain CD57. The rearrangement of NK cell repertoire upon engagement with the HLA-E-UL40 axis occurs due to epigenetic factors as cleverly described by (Lee et al., 2015; Schlums et al., 2015). Even though HCMV infection evokes an expansion of T and NK cells as aforementioned, focusing on the latter is encouraged due to its lack of stringency or leniency and augmented anti-tumor activity (Okpoluaefe et al., 2024).

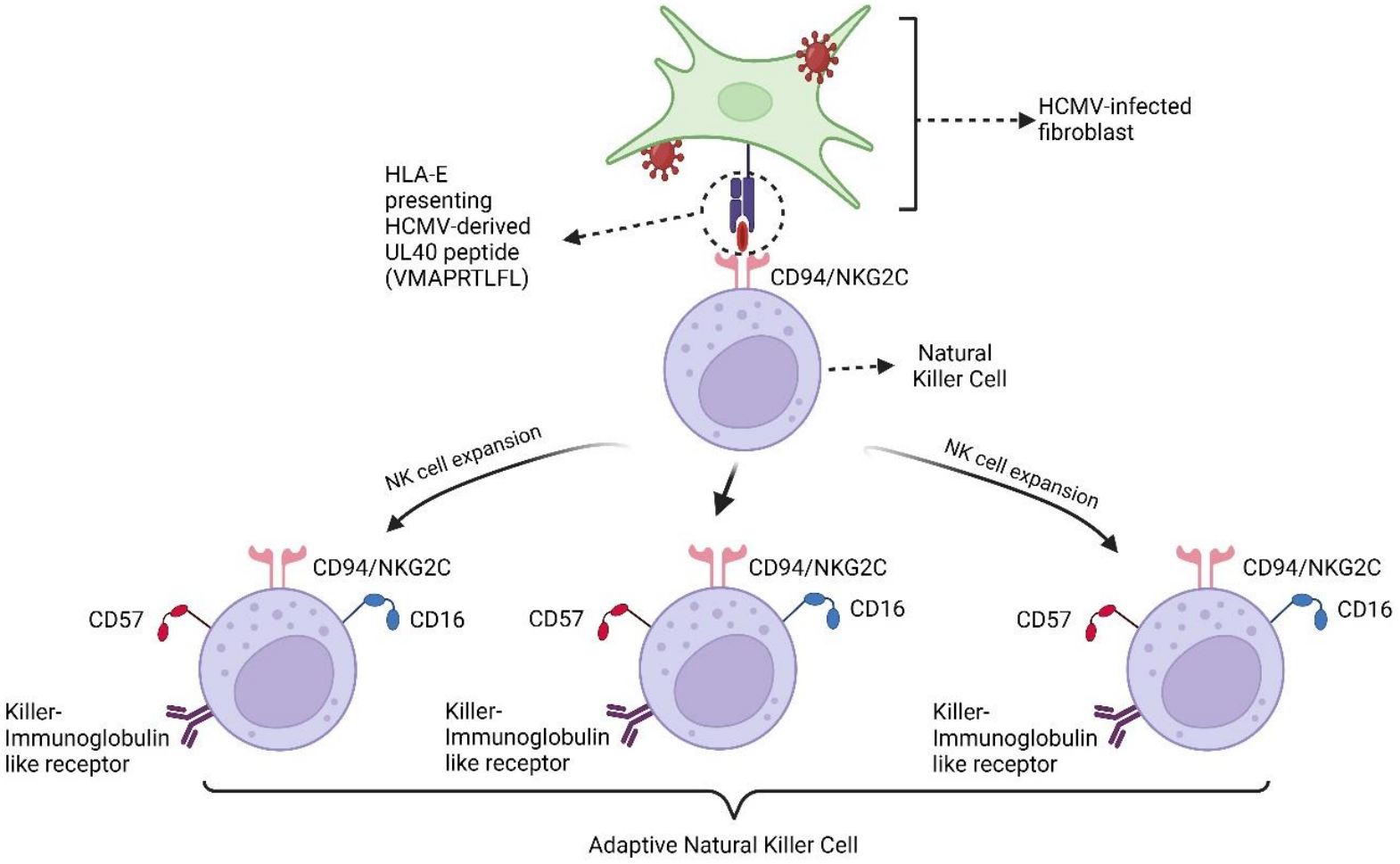

Figure 2.2. Activation and subsequent expansion of NK cells via CD94/NKG2C after interaction with HLA-E presenting HCMV viral peptide, UL40. Note: Image adapted from Okpoluaefe et al., (2024).

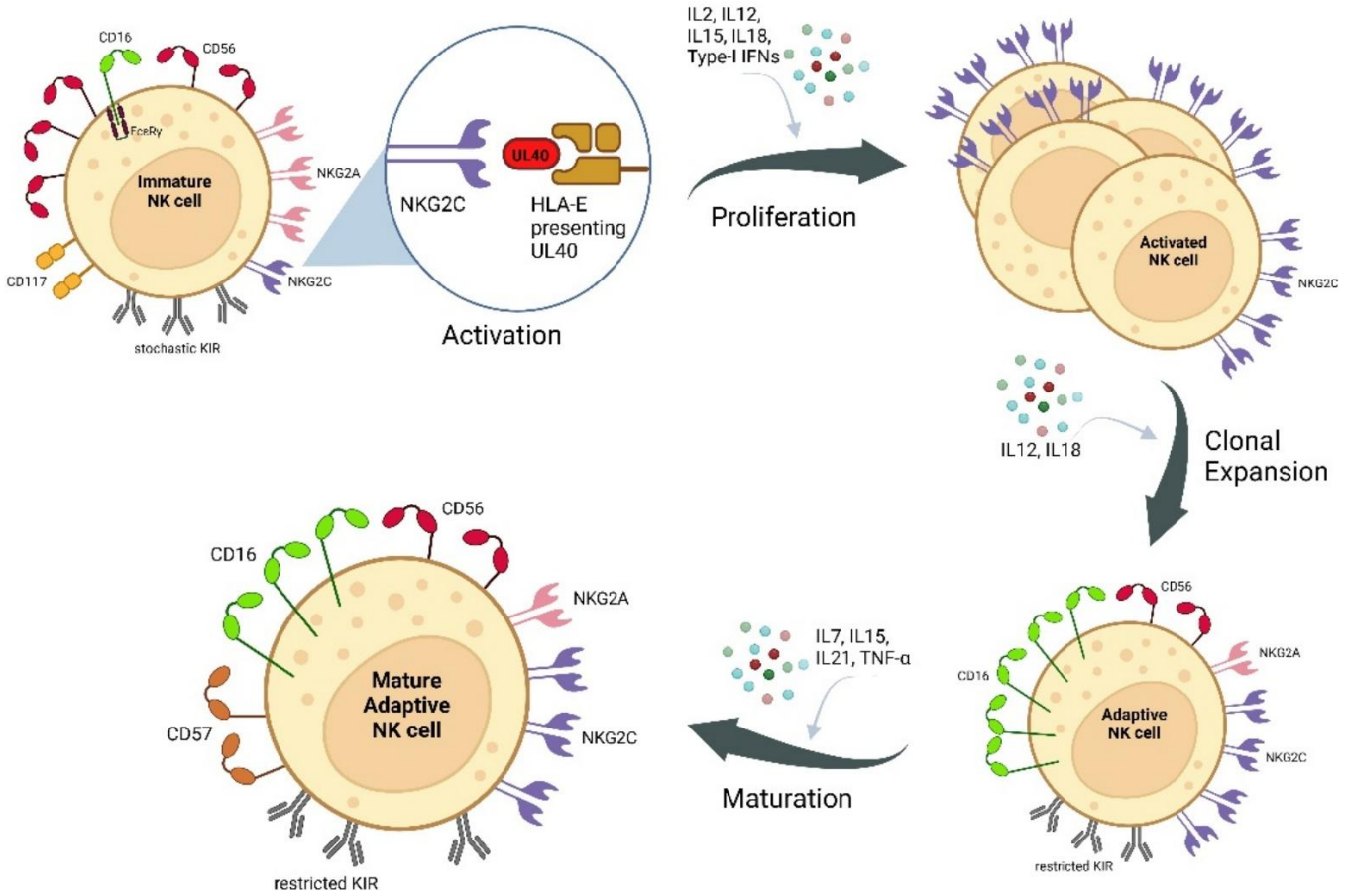

Figure 2.3. Activation, proliferation, clonal expansion, and maturation of NK cells after HCMV infection. The interaction between NKG2C (*KLRC1*) and HLA-E stabilized by the MHC class I VL9 mimic, UL40, combined with proinflammatory signals from Interleukin-2 (IL2), Interleukin-12 (IL-12), Interleukin-15(IL15), Interleukin-18 (IL-18), and type I interferons, activates immature NK cells and drives the proliferation of a substantial subset of NKG2C+ NK cells. CD16 bright/+ CD56 dim/- NKG2A dim/- NKG2C+ NK cells are preferentially amplified during clonal expansion. Subsequent maturation into adaptive NK cells is facilitated by cytokine signals from Interleukin-7 (IL-7), Interleukin-15 (IL-15), Interleukin-21 (IL-21) and tumor necrosis factor α (TNF-α). Note: Image adapted from Okpoluaefe et al., (2024).

2.7. NKG2C genotype and disease association

NKG2C is an activating receptor primarily found on NK cells (Gumá et al., 2006) and a few T cell subsets such as CD8+ T cells (Sottile et al., 2021), and γδ T cells (Fausther-Bovendo et al., 2008). This receptor is involved in the regulation of innate and adaptive immunity related to cytotoxicity and cytokine production (Park et al., 2008). NKG2C receptor levels have been linked to the NKG2C genotypes one possesses; homozygous wildtype (WT/WT), heterozygous wildtype-deletion (WT/DEL), and homozygous deletion (DEL/DEL). Individuals with WT/WT have the highest NKG2C levels in the lot, followed by WT/DEL and DEL/DEL (Ariyanto et al., 2018; Martínez-Rodríguez et al., 2016; Muntasell et al., 2016).

The occurrence of NKG2C deletion allele differs from population to population (**Table 2.2**), albeit the cause of occurrence, and deviance in percentages between populations remains unknown. The NKG2C DEL allele was unexpectedly discovered by Hikami et al., (2003) during a polymorphism screening of the Japanese population, where they identified homozygous NKG2C DEL genotype in about 4.3% of healthy individuals, with an estimated allele frequency of 20.7%. Miyashita et al., (2004), then later, mapped the 16 kilobase pairs long breakpoint to be within a 292bp region, roughly 1.5-1.8 kilobase pairs away from the 3' untranslated region of NKG2A, and moving towards the telomeric end (**Figure 2.4**). Interestingly, Liu et al., (2016) found that the loss of NKG2C in individuals with homozygous NKG2C deletion is compensated by CD2. They found that an adaptive CD2+ NK cell subset expanded in people with NKG2C homozygous deletion, where CD2 co-stimulation was found to be critical for adaptive NK cell response to

IgG-coated target cells. A brief literature search showed that NKG2C was not associated with trachoma disease presentation, determining the prognosis of nasopharyngeal carcinoma, and rheumatoid diseases (**Table 2.2**). Conversely, Park et al. (2008) reported that individuals with the NKG2C (*KLRC2*) rs1141715 serine/serine genotype and the NKG2D (*KLRK1*) rs225536 threonine/threonine genotype exhibited an increased risk towards rheumatoid arthritis, with. the coexistence of these two polymorphisms elevating the risk of the disease by 12-fold.

Table 2.2. Summary of African, Asian, and European populations with reported heterozygous NKG2C deletion (WT/DEL) and homozygous NKG2C deletion (DEL/DEL), and their association with different disease groups. Abbreviations: POI - population of interest; WT - wildtype; DEL - deletion type.

Author(s)	POI	WT/DEL (%)	DEL/DEL (%)	Disease group	Study size (n)	Finding
Goncalves et al., (2016)	East Africans (Tanzania)	20.9	4.7	Ocular *Chlamydia trachomatis* infection	1522	No correlation between NKG2C copy number and trachoma disease presentation.
	West Africans (Gambia, and Guinea-Bissau)	36.2	13.7			
Li et al., (2015)	Northern Chinese (Inner Mongolia Han and Inner Mongolia Mongol)	34.2*	4.9*	Healthy individuals with no family history of nasopharyngeal carcinoma (NPC), and no personal history of cancer.	1129	NKG2C deletion and HLA-E signalling pathways do not play an important role in determining the prognosis of NPC.
	Southern Chinese (Hunan Han and Guangdong Han)	37.0*	6.1*			
	Southeastern Chinese (Fujian Han)	39.3*	7.2*			
Miyashita et al., (2004)	Japanese	20.2	4.1	Healthy individuals	245	No relationship was found between NKG2C deletion and rheumatoid diseases. The results indicated that NKG2C deletion is common in the Japanese and Dutch and is nonessential for evolutionary fitness.
	Dutch	20.0	3.8		105	
	Japanese	34.8	5.2	Systemic lupus erythematosus	155	
	Dutch	38.2	5.6		89	

Table 2.2. Continued.

	Japanese	33.5	5.1	Rheumatoid arthritis	176
	Dutch	NIL	NIL		NIL

*Average percentage calculated using the Hardy-Weinberg equation, $p^2 + 2pq + q^2 = 1$, using the NKG2C WT, and DEL allele frequency of each population group provided in the publication

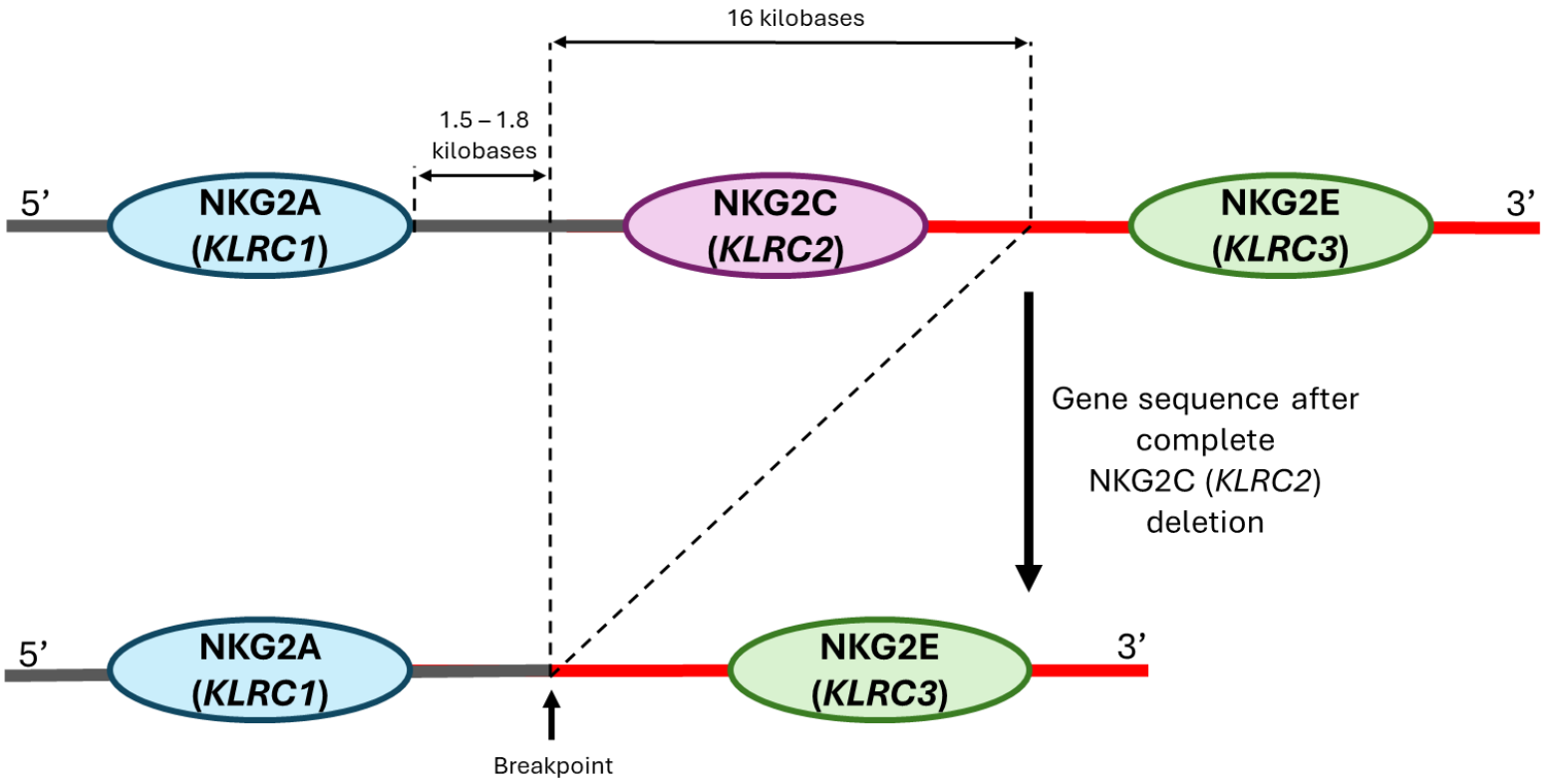

Figure 2.4. Graphical presentation of the complete deletion of *KLRC2* (NKG2C) on Chromosome 12, and the resulting gene sequence following the deletion. The breakpoint is mapped to the 1.5-1.8 kb telomeric region from the 3’ untranslated region of NKG2A (*KLRC1*), with a deletion spanning 16kb. Note: Image adapted from Miyashita et al., (2004).

2.8. HLA-E genotypes and associated diseases

The human leukocyte antigen-E (HLA-E) is a non-classical major histocompatibility complex (MHC) class I molecule responsible for the presentation of leader peptides from classical MHC class I molecules, such as HLA-A, HLA -B, and HLA-C, and the non-classical HLA-G (Lin et al., 2023). HLA-E is expressed on most nucleated cells, with its highest expression observed on immune cells (Boegel et al., 2018). In tissues, on the other hand, it has a restricted distribution pattern (Kedzierska & Turowski, 2001; Puppo et al., 2002). Soluble forms of HLA-E can also be found in sera (Kedzierska & Turowski, 2001), suggesting the importance of the role played by HLA-E in disease prevention, and progression.

Under normal circumstances, HLA-E is stabilized by nonamer signal peptides, derived from classical MHC class I molecules, bearing valine at position 1 and leucine at position 9 (VL9). During HCMV infection, the VL9 mimic, viral nonamer (9-peptide protein), UL40, is presented on HLA-E instead of class I HLA antigens, enabling escape from NK cell-mediated cytotoxicity via the CD94/NKG2A-HLA-E inhibition axis (Tomasec et al., 2000). The HLA-E-leader/viral peptide complex serves as a ligand for CD94/NKG2 receptors on NK cells and T cell subsets (Lin et al., 2023; Pietra et al., 2009). NKG2A ligates HLA-E on healthy cells for ‘self’ discrimination (Lin et al., 2023). Cells lacking MHC class I surface markers are recognized as foreign/non-self and become the target of NK cell-mediated killing (Ljunggren & Kärre, 1990). Similarly, on T-cells, the interaction of NKG2C and NKG2A leads to the enhancement and dampening of cytotoxic functions, respectively (Lin et al., 2023)).