Codon pair bias deoptimization of a major oncogene as an attenuation strategy for Marek's disease herpesvirus

Inaugural-Dissertation zur Erlangung des akademischen Grades

Doctor of Philosophy (Ph.D.) in Biomedical Sciences

an der

Freien Universität Berlin

vorgelegt von

Pratik Hemant Khedkar

aus Pune, Indien

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Jürgen Zentek

Erster Gutachter: Univ.-Prof. Dr. Nikolaus Osterrieder

Zweiter Gutachter: PD Dr. Sandra Blome

Dritter Gutachter: Prof. Dr. Benedikt B. Kaufer

Deskriptoren (nach CAB-Thesaurus):
Marek's disease, Marek's disease virus, genetics, open reading frame, codons,
virulence

Tag der Promotion: 27.08.2018

To the memories of my grandfather Bhagwan Khedkar (1935 – 2014)

and friend Ashish Kawade (1988 – 2014)

न जायते म्रियते वा कदाचित्

नायं भूत्वा भविता वा न भूय:|

अजो नित्य: शाश्वतोऽयं पुराणो

न हन्यते हन्यमाने शरीरे||

-भगवद्गीता २.२०

The Soul is neither born, nor does It die. Nor, having once existed, does It cease to exist.

For It is unborn, eternal, immortal and ageless. It does not die when the body dies.

-Bhagavad Gītā 2.20

1. Table of contents

2. List of abbreviations

amp^r	Ampicillin resistance gene
ATCC	American type culture collection
BAC	Bacterial artificial chromosome
BFP	Blue fluorescent protein
BR1	Basic region 1
BR2	Basic region 2
bZIP	Basic leucine zipper
CD	Cluster of differentiation
CDK2	Cyclin-dependent kinase 2
CEC	Chicken embryonic cell
CPB	Codon pair bias
CPBD	Codon pair bias deoptimization
CPS	Codon pair score
CREB	cAMP response element binding protein
DAP5	Death-associated protein 5
	(Eukaryotic translation initiation factor 4 gamma 2)
ddH$_2$O	Double distilled water
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dpi	Days post infection
ds	Double stranded
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGFP	Enhanced green fluorescent protein
FBS	Foetal bovine serum
GaHV-1	Gallid herpesvirus 1
GaHV-2	Gallid herpesvirus 2
GaHV-3	Gallid herpesvirus 3
gB	Glycoprotein B
gC	Glycoprotein C
gD	Glycoprotein D
gE	Glycoprotein E
GFP	Green fluorescent protein
gH	Glycoprotein H

gI	Glycoprotein I
h	Hour
HB-EGF	Heparin-binding epidermal-growth-factor-like growth factor
hFerH	Gene encoding the heavy subunit of human ferritin protein
hFerL	Gene encoding the light subunit of human ferritin protein
HIV-1	Human immunodeficiency virus 1
HRSV	Human respiratory syncytial virus
Hsp70	Heat-shock protein 70
HSV-1	Herpes simplex virus 1
HVT	Herpesvirus of turkeys
ICP4	Infected cell protein 4
IE	Immediate early
IFA	Immunofluorescence assay
IL	Interleukin
iNOS	Inducible nitric oxide synthase
IR_L	Internal repeat long
IR_S	Internal repeat short
kb	Kilobase pairs
kcal/mol	Kilocalories per mole
kV	Kilovolt
LB	Lysogeny broth
Leu	Leucine
MATSA	Marek's disease tumour-associated surface antigen
MCS	Multiple cloning site
MD	Marek's disease
MDV	Marek's disease virus
MeHV-1	Meleagrid herpesvirus 1
MEM	Minimum essential medium (Eagle's)
MERE	Meq responsive elements harbouring CRE/TRE cores
meq-D	CPB deoptimized version of vvMDV strain RB-1B *meq* gene
meq-O	CPB optimized version of vvMDV strain RB-1B *meq* gene
meq-R	CPB randomized version of vvMDV strain RB-1B *meq* gene
meq-W	Wild-type *meq* gene encoded by vvMDV strain RB-1B
MHC	Major histocompatibility complex
min	minute
MIP	Macrophage inflammatory protein

mM	Millimolar
mMDV	Mild Marek's disease virus
mRNA	Messenger ribonucleic acid
OD_{600}	Optical density at 600 nm wavelength
ORF	Open reading frame
p	Plasmid
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimine
pi	Post infection
pmol	Picomole
pp38	Phosphoprotein 38
PRRSV	Porcine reproductive and respiratory syndrome virus
qPCR	Quantitative real time PCR
RANTES	regulated on activation, normal T cell expressed and secreted
RB	Retinoblastoma protein
RdRp	RNA dependent RNA polymerase
rev	Revertant
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rotations per minute
s	Seconds
S. cerevisiae	*Saccharomyces cerevisiae*
SDS	Sodium dodecyl sulphate
SPF	Specific pathogen free
ss	Single stranded
TAE	Tris-acetate-EDTA buffer
TCR	T-cell receptor
TGF-β	Transforming growth factor β
TR_L	Terminal repeat long
tRNA	Transfer ribonucleic acid
TR_S	Terminal repeats short
U_L	Unique long
U_S	Unique short
v	Reconstituted virus

List of abbreviations

vIL-8	Viral interleukin 8
vMDV	Virulent Marek's disease virus
VSV	Vesicular stomatitis virus
vTR	Virus-encoded RNA subunit of telomerase
vvMDV	Very virulent Marek's disease virus
vv+MDV	Very virulent plus Marek's disease virus
v/v	Volume per volume
w.r.t.	with respect to
w/v	Weight per volume
ΔG	Free folding energy
µF	Microfarad
µl	Microliter
Ω	Ohm
§	Section

3. List of figures

4. List of tables

5. Introduction

The race between viral diseases and vaccination has been going on for centuries. Immunization was described for the first time in China in the eleventh century. The disease targeted was smallpox and the inoculation methods used included: (i) inhalation of powdered smallpox sores, and (ii) rubbing the fluid from smallpox pustules on the skin. The latter method was introduced to Turkey in 1692, from where it came to Great Britain in 1721. The method, however, was not full proof and was associated with a significant death rate. As a result, physicians were on a constant search for safer methods to fight smallpox until Edward Jenner (1748 – 1823), an English physician, observed that certain individuals who had contracted cowpox were not susceptible to smallpox. Thus began the story of modern immunization also known as vaccination, a term coined after cowpox (Latin: *Variolae vaccinae*) (1). It took almost two centuries, after Jenner's first experiments in 1796, for the complete elimination of smallpox which was undeniably a glorious landmark in the history of vaccination. The twentieth century also saw successful introduction of vaccines against other viral diseases such as polio, measles, mumps, rabies, rubella, yellow fever, influenza, varicella, rotavirus, and hepatitis A and B. Several other viral vaccines like adenovirus, cytomegalovirus, dengue virus are in the pipeline (2,3).

The great optimism brought by the success story of smallpox vaccination was further heightened by the elimination of poliomyelitis from many countries through either live-attenuated or inactivated vaccines. Vaccination also helped making measles an extremely rare disease (2,4). However, scientists soon realized that not all disease-causing viruses can be eliminated by vaccination, e.g. human immunodeficiency virus (HIV). Dengue, a flaviviral disease, is another peculiar example where vaccination is difficult because of antibody dependent enhancement of the disease, a phenomenon where individuals with pre-existing immunity have a more severe disease outcome than naïve individuals (5). Nevertheless, several other diseases can be controlled by regular vaccination. For example, elimination of influenza through vaccination is virtually impossible owing to the characteristic antigenic shifts and drifts of the virus which render any previously acquired immunity useless (2), yet regularly updated vaccines can effectively control the spread of the disease. Such variations among aetiological agents necessitate improvements in the available vaccine technology. Several types of viral vaccines have indeed emerged through this necessity including live-attenuated, inactivated (killed), subunit, nucleic acid, and chimeric live virus vaccines, each with their own advantages and disadvantages (6).

Recent developments in molecular biology have further improved vaccine technology with successful strategies like subunit vaccination for hepatitis B and virus like

particles for human papilloma virus (6). Coleman *et al.*, in 2008, described another such novel approach towards virus attenuation by genome wide alteration of 'codon pair bias', which is based on the genetic code and its degeneracy (7). The present study dealt with the application of this technology to a highly virulent and oncogenic Gallid herpesvirus 2 (GaHV-2), historically also known as Marek's disease virus (MDV).

5.1. Codon pair bias

The genetic code is degenerate or redundant whereby 20 amino acids and termination signals are encoded by 61 and 3 codons, respectively (**Table 1**). This degeneracy allows living systems to minimize the effects of mutations as well as base pairing errors during protein synthesis (8). Furthermore, it also allows for one protein to be coded in a vast number of ways. For example, a protein made up of 300 amino acid residues can be coded in 10^{151} different ways. However, biological systems prefer to use certain synonymous codons over their counterparts, consequently not using most of these vast number of ways but one (or a few). This preference for synonymous codons, termed 'codon bias', varies among species. The simplest example for codon bias is the encoding of the amino acid alanine in the human genome, where there is a four times higher preference for the codon GCC over its synonymous counterpart GCG (7).

Codon bias was shown to correlate with specific tRNA abundance in bacteria (*Escherichia coli*), yeast (*Saccharomyces cerevisiae*) as well as several multicellular eukaryotes including humans (9). Two other factors that might influence codon bias are: the levels at which particular genes are expressed within the same organism; and the optimization of the fidelity as well as the kinetic efficiency of translation (10). The association of codon bias with translation efficiency is correlated with the availability of

Table 1. Degeneracy of the genetic code.

First letter		Second letter										Third letter
		U		**C**		**A**		**G**				
U	UUU	Phenyl-	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U		
	UUC	alanine	UCC		UAC		UGC		C		
	UUA	Leucine	UCA		UAA	Stop	UGA	*Stop*	A		
	UUG		UCG		UAG		UGG	Tryptophan	G		
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U		
	CUC		CCC		CAC		CGC		C		
	CUA		CCA		CAA	Glutamine	CGA		A		
	CUG		CCG		CAG		CGG		G		
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U		
	AUC		ACC		AAC		AGC		C		
	AUA		ACA		AAA	Lysine	AGA	Arginine	A		
	AUG	Methionine	ACG		AAG		AGG		G		
G	GUU	Valine	GCU	Alanine	GAU	Aspartic acid	GGU	Glycine	U		
	GUC		GCC		GAC		GGC		C		
	GUA		GCA		GAA	Glutamic acid	GGA		A		
	GUG		GGG		GAG		GGG		G		

specific aminoacyl-tRNAs (10), that is, if mRNA depends on underrepresented tRNAs for translation, peptide elongation would slow down leading to stalling and premature dissociation of ribosomes followed by degradation of the resulting truncated peptide (11). Another factor that could influence the choice of codons in an open reading frame (ORF) is the inability of the translational machinery to discriminate between certain codons, e.g. ribosomes discriminate between the codons UUC (phenylalanine) and CUC (leucine) less efficiently than their synonymous codons UUU and CUU, respectively. In other words, the accuracy of translation varies with different synonymous codons (12,13). Codon bias has been widely utilized for optimizing gene expression in target species, e.g. the translation of the L1 and L2 capsid protein genes of human papilloma virus, which is restricted only to epithelial cells, was enhanced in other mammalian cells by codon optimization (14,15).

A similar but independent phenomenon was discovered in *E. coli* where the usage of adjacent codon pairs was shown to be biased, codon pair bias (CPB) (10), that is, certain codon pairs appear in coding sequences more often than expected based on the overall frequencies of their constituent codons (7,10,16,17). In other words, some codon pairs are preferred in ORFs and occur more frequently (overrepresented) than their synonymous counterparts (underrepresented). For instance, the amino acid pair alanine-alanine is expected to be encoded by the codon pairs GCT-GCC and GCC-GCT equally often based on the overall frequencies of the two codons, however in human ORFs, the former codon pair occurs almost six times as often as the latter (**Table 8**).

CPB has been shown to exist in all the species that have been studied so far (13,17) and varies drastically among different species (18,19). However, phylogenetically related species show a highly similar CPB indicating its co-evolution along with the respective species as they diverged (17). Although the existence of CPB has been known for many years (10), the exact mechanisms governing it remain largely elusive. Many hypotheses have been put forward over the years. The first and foremost reason for why CPB exists is, similarly to codon bias, its role in the accuracy and speed of translation (20,21). Besides, CPB seems to have a stronger influence on the translational accuracy and speed than codon bias. It is important to note, however, that the correlation between CPB and codon bias is negligible, thus CPB influences translation independently of codon bias. Specific codon pairs have been shown to lead to missense, non-sense and frameshift errors during translation (21). It was also shown that, under starvation, the *E. coli* translation machinery misreads the asparagine codons AAU and AAC as lysine. Moreover, the rate of lysine misincorporation is nine times higher with AAU than with AAC. The authors also noted that such misincorporation is 2-fold higher if the misread codon occurs in the order GAC-nnn-GGC, implicating CPB in the misincorporation (22). Certain underrepresented codon pairs followed by a particular

nucleotide in *S. cerevisiae*, namely, CUU-AGG-C, CUU-AGU-U and GGU-CAG-A, are responsible for ribosomal slippage leading to frameshifts (23). The properties of aminoacyl-tRNAs, especially the interaction between two aminoacyl-tRNAs that are juxtaposed on the A- and P-sites of a ribosome during translation, determine the peptide elongation rate. Due to this property, it has also been proposed that tRNA properties play an important role in shaping CPB (24). These data indicate that CPB is governed, at least partially, by the translation machinery of the host.

Another theory to explain the CPB phenomenon is based on the frequency of certain dinucleotides in the ORF, especially CpG and TpA (13,17). These dinucleotides are involved in DNA repair and replication and may be related to genetic mutational drift (13). The suppression of these dinucleotides in viral sequences is hypothesized to be because of their recognition by some largely elusive innate immune mechanisms (17,25,26). Despite these theories, the exact mechanism behind CPB remains to be identified.

5.2. CPB in virus attenuation

Even though the underlying mechanisms are poorly understood, the phenomenon of CPB has nonetheless been utilized to attenuate viruses. The attenuation strategy involves CPB deoptimization (CPBD) of the coding sequences. CPBD is achieved by reshuffling synonymous codons in an ORF without altering the amino acid sequence of the encoded protein. This reshuffling preserves the codon bias, i.e. the overall frequencies of individual codons, while simultaneously altering the CPB. The reshuffling is performed to increase the severity and frequency of underrepresented codon pairs, which leads to inefficient protein production and, as a result, attenuation. Viruses attenuated by CPBD are antigenically identical to the parental pathogenic virus. Moreover, they harbour hundreds of silent single nucleotide mutations which make them genetically extremely stable with virtually no chance of reversion (7).

Poliovirus, a positive sense single stranded (ss) RNA virus of the family *Picornaviridae*, was the first virus to be attenuated using CPBD (7). Other viruses that were attenuated using the strategy include Influenza A virus, HIV-1, human respiratory syncytial virus (HRSV), porcine reproductive and respiratory syndrome virus (PRRSV), vesicular stomatitis virus (VSV), etc. Viruses that have been attenuated using CPBD to date are enlisted in **Table 2**. Besides viruses, the pneumonia causing bacterium, *Streptococcus pneumoniae*, was rendered less virulent in mice by CPBD of its virulence factor pneumolysin (27). Thus, CPBD as an attenuation strategy has been employed for one bacterium and several ssRNA viruses but had never been tested on large double stranded (ds) DNA viruses such as poxviruses and herpesviruses. This brings us to the central objective of the present study which was to determine if CPBD is suitable for attenuation of

herpesviruses. To answer this, a unique small animal model was used, which involves a highly virulent and oncogenic herpesvirus, MDV, and its natural host, the chicken.

5.3. Marek's disease

Marek's disease (MD) was first reported in 1907 by Jószef Marek (1868 – 1952), a Hungarian veterinarian, after whom the disease was named in 1960. He reported a hallmark symptom of the disease, namely polyneuritis, a generalized inflammation of

Table 2. Viruses that have been attenuated using CPBD.

Virus (Family)	Genomic material	CPB deoptimized gene(s)	Attenuation	Ref.
Poliovirus (*Picornaviridae*)	(+) ssRNA	Capsid protein	Attenuated in mice	(7)
Influenza A (*Orthomyxoviridae*)	(-) ssRNA Segmented	PB1 (RdRp subunit), HA, NA	Attenuated in mice, No disease in ferrets	(18, 28, 29)
HIV-1 (*Retroviridae*)	ssRNA-reverse transcriptase	*gag*, *pol*	Attenuated in cell culture	(30)
HRSV (*Pneumoviridae*)	(-) ssRNA	Entire genome	Attenuated in mice and African green monkeys	(31)
PRRSV (*Arteriviridae*)	(+) ssRNA	GP5 (Envelope glycoprotein)	Attenuated in pigs	(32)
VSV (*Rhabdoviridae*)	(-) ssRNA	L gene (partial)	Modest attenuation in cell culture and mice	(33)
Dengue virus (*Flaviviridae*)	(+) ssRNA	Envelope (E) protein, Non-structural proteins 3 (serine protease) and 5 (RdRp)	Attenuated in new-born mice	(34)

the nerves. It was later identified that the symptom, which is observed in chickens with MD even today, is a result of infiltration of the nervous tissue by mononuclear cells. Visceral lymphoma was proposed to be a second syndrome of MD in the 1920s (35). Initially, MD was indistinguishable from another neoplastic chicken disease, retrovirus-induced lymphatic leucosis. It was only in the late 1960s that the herpesviral aetiology of MD was proven (36) and confirmed (37,38).

Owing to its oncogenic nature, MDV was thought to be closely related to Epstein-Barr virus, and hence classified as a member of the sub-family *Gammaherpesvirinae.* However, later studies revealed that MDV genome possesses repeat structures which are characteristic of the sub-family *Alphaherpesvirinae* that also includes important human pathogens such as herpes simplex virus type 1 (HSV-1), HSV-2 and varicella-zoster virus (39–41). Within the sub-family, MDV belongs to the genus *Mardivirus*, which includes other closely related viruses such as Meleagrid herpesvirus 1 (MeHV-1, also known as herpesvirus of turkeys (HVT)) and the non-oncogenic Gallid herpesvirus 3 (GaHV-3). GaHV-3 was initially believed to be another serotype of MD (hence named MDV-2), however, later studies on their genome sizes, base compositions as well as pathogeneses, made it apparent that the viruses are, in fact, two distinct species and that they have evolved in parallel for millions of years (42). This led to a change in the nomenclature of the viruses from MDV-1 and MDV-2 to GaHV-2 and GaHV-3, respectively (43). However, to avoid confusion, the name MDV for GaHV-2 is still prevalent in literature and for the same reason will be used in this thesis. The causative agent of infectious laryngotracheitis, GaHV-1, is significantly different than all the mardiviruses, both in terms of pathogenesis and genomic content, and hence is placed in a separate genus – *Iltovirus* (43).

5.3.1. Clinical presentation

MD is primarily a disease affecting chicken although other avian species including quail, turkey and pheasant also occasionally succumb to the disease. The clinical manifestation of MD depends not only on the pathotype of the virus and the genetic make-up of the chicken but also on the environment (44). MDV infected chicken show a wide range of symptoms including non-specific symptoms such as weight loss, diarrhoea, anorexia and pallor. The MD specific symptoms include paralysis caused by lymphoid infiltration of the peripheral nerves, lymphomas in internal organs, immunosuppression, lethargy, blindness and skin lesions. MD symptoms are currently classified as: (i) 'classical' MD or the so-called fowl paralysis, (ii) acute leucosis with lymphomas in internal organs; and (iii) transient paralysis and acute transient paralysis (45,46).

Classical MD (neurolymphomatosis or fowl paralysis) presents as asymmetrical paralysis of one or both legs because of lymphoid infiltration of the brachial and the lumbosacral nerves. Lack of physical coordination due to paralysis is an early sign characterized by one leg held forward and the other backward (**Figure 1A** and **B**) (47). Affected nerves appear grey to yellow and swollen due to inflammation and lymphoid infiltration. Muscular atrophy is also observed (46,48). Lymphomas can also be seen in several internal organs including spleen, kidneys, liver (**Figure 1C** and **D**), gonads, skeletal muscles, and proventriculus (45).

In the 1950s and 1960s, the lymphomatous presentation of MD became more prevalent and the severity of the disease has been increasing ever since. This form of
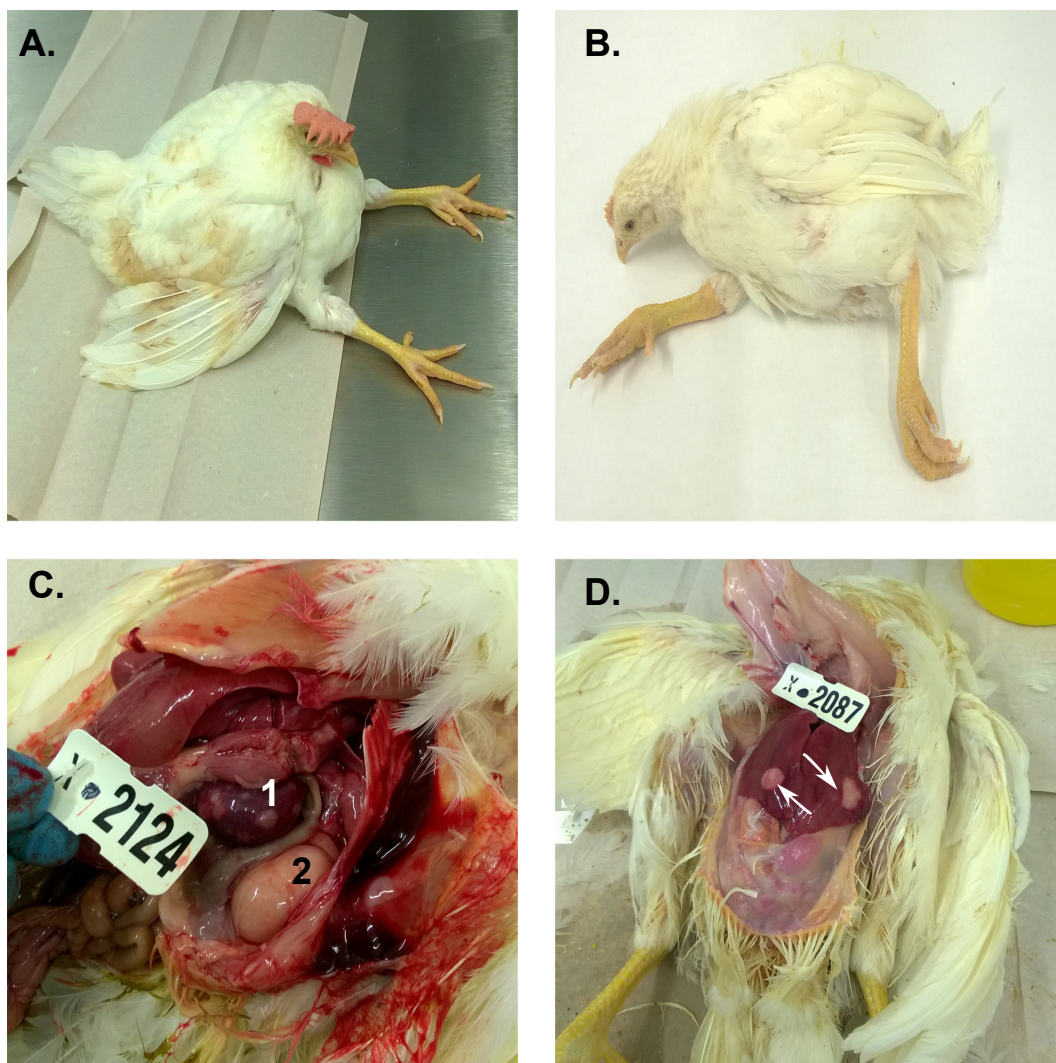
Figure 1. Clinical presentation of MD and macroscopic lesions.

Affected chicken showing **(A)** lethargy and inability to stand, **(B)** asymmetrical paralysis characterized by one leg held forward and the other backward. Typical necropsy findings include disseminated lymphomas on several internal organs including spleen **(C1)**, kidneys **(C2)** and liver **(D)**.

(N.B. The images were taken during the present study.)

the disease was termed 'acute leucosis' or 'acute MD' and is characterized by sudden death of the infected birds as early as 9 – 10 days post infection (dpi), sometimes even without the development of either neurological symptoms or tumours (35).

5.3.2. MDV: Structure and genome organization
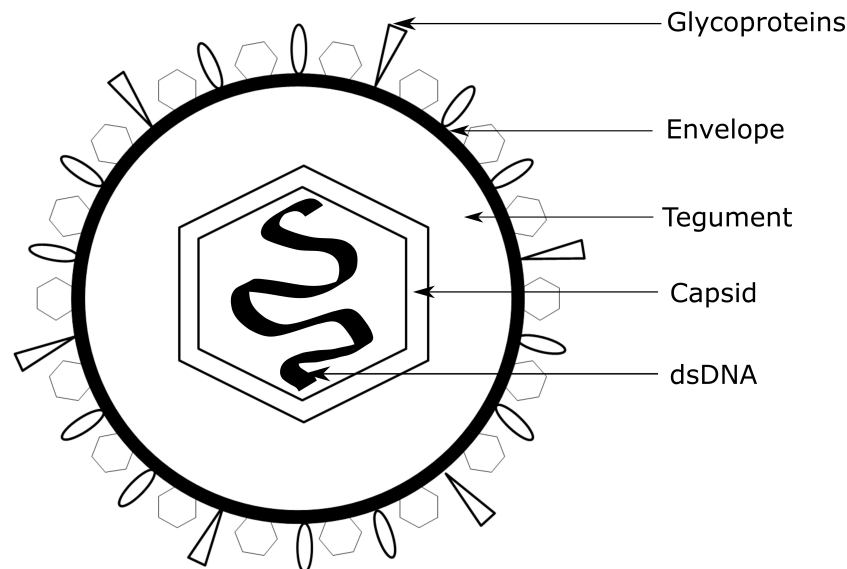
Figure 2. Schematic representation of MDV virion structure

The virion structure of MDV (**Figure 2**) is characteristic of the order *Herpesvirales*. It consists of a linear, dsDNA genome contained in an icosahedral (T = 16) nucleocapsid. The capsid is surrounded by a proteinaceous matrix of variable size, the tegument. Surrounding the tegument is a host cell derived lipid envelope containing membrane associated viral glycoproteins that appear as spikes in electron micrographs (36,43).

The linear, dsDNA genome of MDV is approximately 180 kilobase pairs (kb) in length and shows a typical type E organization depicted in **Figure 4B**. It resembles the genome of HSV-1, the prototype virus of the sub-family *Alphaherpesvirinae*, and consists of two unique regions, unique long (U_L) and unique short (U_S). Each of the unique regions is flanked by inverted repeats called terminal and internal repeats long (TR_L and IR_L) and terminal and internal repeats short (TR_S and IR_S), respectively (35,49).

5.3.3. MDV replication cycle

In vitro, MDV replicates at a slow rate giving rise to visible plaques several days after infection. It is also highly cell associated. MDV replication cycle *in vivo*, on the other hand, is very complex and remains poorly understood to date. The currently accepted 'Cornell model' of MDV replication (**Figure 3**) involves transmission of the virus among different cell types. The virus infects chicken via the respiratory route upon inhalation of cell-free, infectious MDV particles from environments contaminated by infected birds.

The primary source of this cell-free virus is the feather follicle epithelium, the only site where MDV infection is fully productive and leads to the release of cell-free virus particles which are shed into the environment with dead cell debris (dander) and moulted feathers. MDV can remain infectious in such contaminated environments for several months (50,51). The Cornell model describes MD pathogenesis in four stages: early cytolytic phase (2 – 7 dpi), latent phase (7 – 10 dpi onwards), late cytolytic and immunosuppressive phase (18 dpi onwards), and proliferative phase (28 dpi onwards) (51).

5.3.3.1. Early cytolytic phase

Following inhalation, MDV is predicted to infect phagocytes, namely, macrophages or dendritic cells, in the respiratory tract either directly or following a round of replication in the epithelial cells (35). Within 24 hours (h) after infection, the infected phagocytes carry the virus to the primary lymphoid organs, thymus and bursa of Fabricius, through the
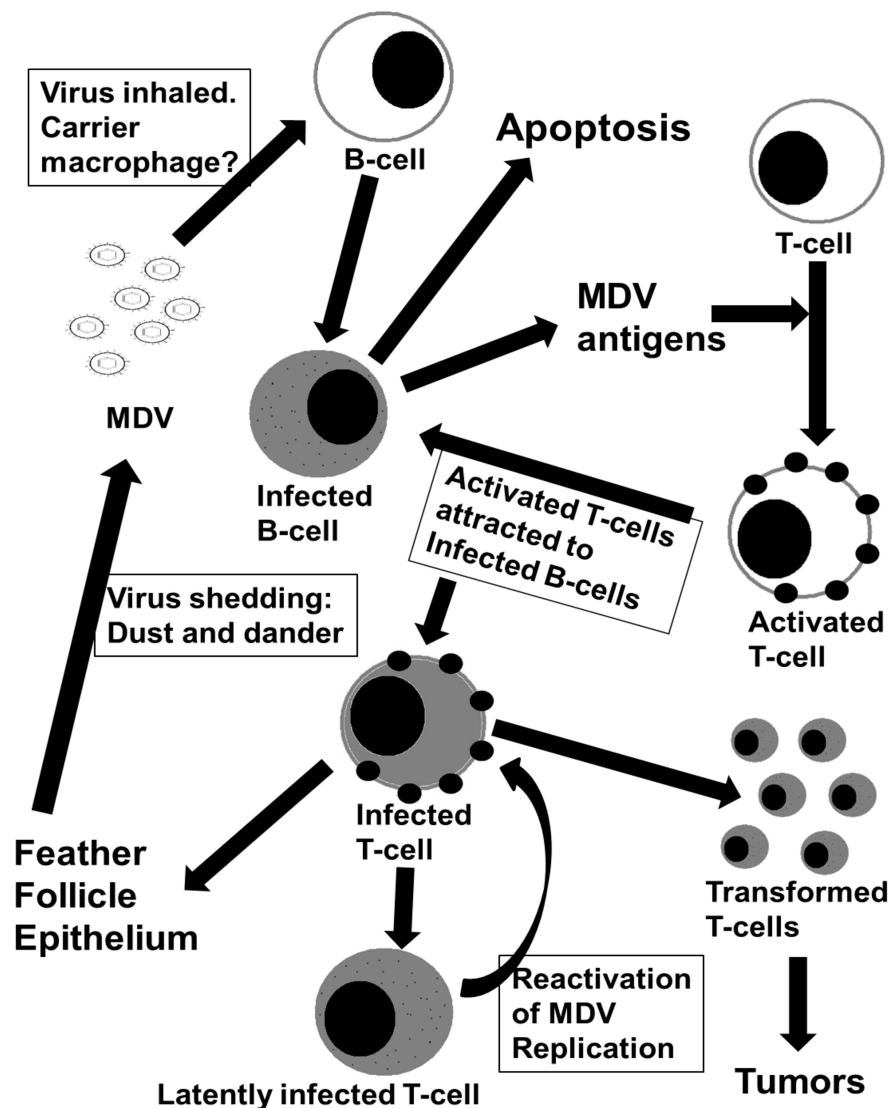
Figure 3. Schematic representation of MDV replication cycle

bloodstream (35). The virus is detectable in the secondary lymphoid tissues (spleen, gut-associated lymphoid tissue, Harderian gland, caecal tonsil and conjunctiva-associated lymphoid tissue) from 2 to 7 dpi, peaking at 4 dpi (51–53). It is in the lymphoid tissues that the virus infects lymphocytes, B-cells and activated T-cells, which are primary targets for the initial phase of cytolytic replication. MDV rarely also infects CD8$^+$ T-cells and CD4$^-$ CD8$^-$ T-cells (35).

In lymphoid tissues, majority (~ 90 %) of the lytically infected cells are B-cells (51). These cells not only act as virus amplification factories but also activate resting T-cells. This activation is necessary for the T-cells to become susceptible to MDV infection since resting T-cells are fairly refractory to infection (54). As mentioned earlier, most of these infected T-cells are CD4$^+$ T-cells, which serve the following functions in MDV pathogenesis: (i) act as a target for transformation, (ii) act as a reservoir for latent MDV genomes, (iii) mediate spread of the virus within the host, and (iv) facilitate viral transport to the skin (35). The last function is especially crucial since MDV replication is semi-productive, which means that these cells are unable to produce certain structural components of the virus thereby being unable to produce fully mature, cell-free virus (51). CD4$^+$ T-cells appear to be the 'Trojan horse' which allows the transfer of the virus to the feather follicle epithelium where infectious cell-free virus is produced and shed into the surroundings, which in turn is responsible for the spread of the virus to chickens in the vicinity leading to infection and disease in susceptible birds (35).

5.3.3.2. Latency

Latency is defined as the presence and maintenance of viral genomes in host cells without the production of infectious progeny virus (35). Along with the advantage of passive replication during cell division, latency allows the virus to escape immunity due to minimized viral protein production (55). MDV latency is primarily established in CD4$^+$ T-cells beginning around 7 dpi, however, latent infection has also been reported in B-cells, CD4$^-$ CD8$^-$ T-cells and CD8$^+$ T-cells (35,56). MDV enters latency by integrating its genome into the host telomeres using two arrays of telomeric repeats present at the ends of its linear genome (57–59). It is worth noting that, besides MDV, the ability to integrate their genome into the host genome is observed in only two other members of the family *Herpesviridae*: human herpesvirus 6 and Epstein-Barr virus (60). Recently, possible chromosomal integration has also been reported in human herpesvirus 7 (61). Several genes, estimated between 10 and 30, are transcribed during the latent phase of infection. Latency associated transcripts that have been studied so far arise from the following three genomic regions: transcripts in antisense

orientation to *ICP4* (infected cell polypeptide 4) (62–64), the so-called 1.8 kb family of transcripts (42,65), and transcripts originating from the *meq* region (35,42).

Our knowledge of the factors that govern the fate of an MDV infection: whether it will be latent or lytic, remains limited (35). The outcome of MDV infection depends on intricate interactions between the virus and the host. Resistant chicken can develop a latent MDV infection without clinical signs while in susceptible and/or immunocompromised birds the early cytolytic phase of MDV infection is prolonged (66). It was also shown that conditioned medium from concanavalin A-activated chicken spleenocytes could maintain latency in MDV infected spleenocyte cultures, thus proving the role of host cytokines in the maintenance of latency (67). Latently infected T-cells not only act as reservoirs for the virus but also disseminate the virus in the body of the infected host through lymph and bloodstream. In susceptible birds, these cells can undergo neoplastic transformation leading to gross lymphomas in almost all of the internal organs (51).

5.3.3.3. Late cytolytic and immunosuppressive phase

The early cytolytic phase, in which majority of the lytically infected cells are B-cells, lasts from 7 – 14 dpi. Severe immunolysis during this phase leads to the suppression of both humoral and cell-mediated branches of immunity. This immunosuppression is seen around 18 dpi and coincides with a second round of cytolytic infection in MD-susceptible birds. This late cytolytic phase affects tissues like bursa, thymus, some epithelial tissues including feather follicle epithelium, kidney, adrenal gland and proventriculus. The affected organs show inflammation, infiltration of mononuclear cells and heterophils, and, necrosis of lymphocytes and epithelial cells. Thymus and bursa of Fabricius undergo atrophy. It is suggested that immunosuppression causes reactivation of the virus carried to these organs by the latently infected T-cells, which in turn leads to late cytolytic infection (51).

5.3.3.4. Transformation

Neoplastic transformation and lymphoma formation in almost all visceral organs is the hallmark of MDV infection (68,69). Lymphomas appear in infected birds 3 – 4 weeks post infection (51). The primary cells which undergo transformation upon MDV infection are CD4$^+$ CD8$^-$ T-cells (70). The fact that the same T-cell subpopulation is also the site for latent MDV infection suggests that establishment of latency is an essential criterion for neoplastic transformation and formation of lymphomas (35). Although early cytolytic infection of B-cells is not essential for MDV replication, the formation of lymphomas certainly seems to depend on its magnitude (51). A higher number of cytolytically infected cells in the initial phase ensures the development of sufficient latently infected cells,
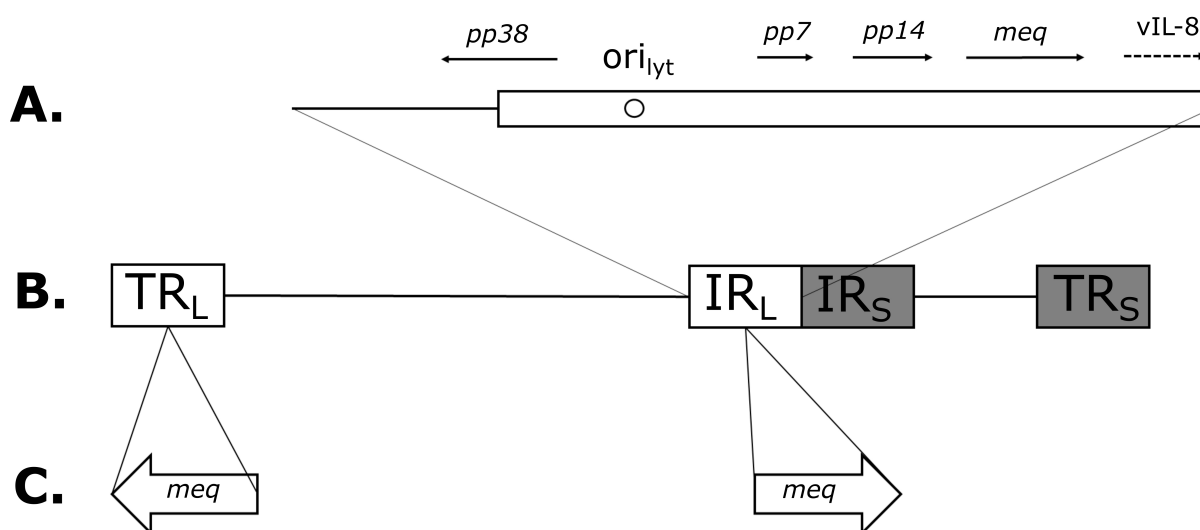
Figure 4. Organization of MDV genome.

(A) ORFs in the R_L region unique to MDV. (B) Genome organization of MDV. (C) Location of the two *meq* ORFs on the MDV genome.

which in turn induce lymphomas (35). However, only a few of these latently infected cells are thought to undergo transformation leading to oligoclonal tumours. In some birds, the tumours can even be monoclonal (57,71). MDV transformed cells are characterized by the presence of several cell-surface molecules such as major histocompatibility complex (MHC) classes I and II (MHC-I^{hi}, MHC-II^{hi}), T-cell receptors αβ (TCRαβ$^+$), Hodgkin's disease antigen CD30^{hi}, CD28$^{lo/-}$, interleukin-2 receptor α (IL-2α$^+$), MD tumour-associated surface antigen (MATSA) and AV37 antigen (51,70,72,73).

Several MDV ORFs are known to be expressed in transformed lymphocytes and have been shown to have transforming properties. These include the BamHI-H family of transcripts, phosphoprotein 38 (pp38), ICP4 and its related transcripts, virus-encoded CXC chemokine (vIL-8), virus-encoded RNA subunit of telomerase (vTR), and *meq* (See § 5.3.3.5.). Barring *ICP4*, all of these genes are unique to MDV and are encoded in the R_L region (repeats flanking the U_L region) (**Figure 4A**) (74).

The BamHI-H region produces at least three transcripts of varying sizes (1.8 kb, 3.0 kb and 3.8 kb). Several other ORFs have been predicted in the region, however, only two of these are reported to yield protein products. The transcripts are involved in prolonging proliferation and serum dependence of chicken embryonic fibroblasts which suggests their role in development and maintenance of lymphomas (75). Although pp38 plays a crucial role in the early cytolytic phase, studies using antisense oligonucleotides against pp38 have demonstrated its implication in lymphoblastoid cell proliferation as well (76). Viral chemokine vIL-8 has been shown to promote lymphoma

formation by recruiting target cells and expanding the infected cell population (77). vIL-8 may also maintain the T-cells in an activated state prolonging their survival and trigger several kinase cascades. These kinases, in turn, activate the Jun/Fos pathway, and thus vIL-8 may act together with Meq to induce lymphomas (74). vTR negative viruses show severely impaired tumour formation, suggesting that vTR plays an important role in tumorigenesis, the exact mechanism behind which remains to be elucidated (58). ICP4 and its related transcripts are detectable at high levels in latently infected cells (63). Moreover, inhibiting ICP4 transcripts using antisense oligonucleotides curtails the proliferation of lymphoblastoid cells (76). Both results suggest that ICP4 plays a strong role in transformation.

5.3.3.5. *meq:* The major oncogene of MDV

Among all the genes associated with MDV induced transformation, *meq* is the most consistently expressed gene in both latently infected as well as transformed cells (74,78,79). The gene is named after its location on the Q fragment of the EcoRI digested MDV genome (**M**DV ***Eco*Q**: *meq*) (79). Two copies of the *meq* gene, *MDV005* and *MDV074*, are encoded in the MDV genome in the IR$_L$ and TR$_L$ regions, respectively (**Figure 4C**). The gene encodes a 339 amino acid long nuclear phosphoprotein, Meq, characterized by a basic-leucine zipper (bZIP) domain at the N-terminus and a proline-rich transactivation domain at the C-terminus (**Figure 5**) (79). It is the only herpesviral bZIP protein within the Jun/Fos family of transcription factors (74,76). It is expressed during both the lytic and latent phases of infection and performs several functions such as transactivation, DNA binding, chromatin remodelling and regulation of transcription (35,80).

The bZIP domain allows Meq to dimerize with other proteins that also contain a bZIP domain. Due to this property, Meq can not only form homodimers but also high affinity heterodimers with several proteins including c-Jun, JunB, JunD, ATF2 and Fos. This dimerization facilitates the basic region adjacent to the ZIP domain to interact with DNA (74). The most stable and perhaps the most favourable dimer formed is Meq-c-Jun. The stability of this dimer allows c-Jun to function like v-Jun (a retroviral oncoprotein) and activate the following genes: cathepsin-like protein JTAP-1, JAC and heparin-binding epidermal-growth-factor-like growth factor (HB-EGF), which c-Jun alone does not usually activate. Each of these genes is a proto-oncogene and can transform chicken cells independently. Another important function of Meq-c-Jun dimers is transactivation by interacting with AP-1 sites or the so-called MERE sites (Meq responsive elements harbouring CRE/TRE cores). This results in the upregulation of the proto-oncogenes mentioned above as well as *meq* itself. Similarly, Meq transactivates

IL-2 and CD30 through MERE sites. This interference with the CD30 pathway is a feature that MDV shares with Epstein-Barr virus, another oncogenic herpesvirus (72). It has also been shown that antibodies against CD30, which is a cell-surface co-stimulatory molecule, could be used in anticancer therapy (81).

Besides activating the Jun pathway, Meq also activates c-Ski, the cellular homologue of the retroviral oncogene v-Ski (82). Moreover, Meq also upregulates antiapoptotic genes like Bcl-2 and downregulates proapoptotic genes like Fas and DAP5 (74,83,84). Another important role of Meq in maintenance of lymphomas is immunoregulation, which it plays by upregulating IL-8, RANTES, MIP and TGF-β (74). In addition to activation of proto-oncogenes, negative regulation of apoptosis, and immunoregulation, Meq also interacts with factors responsible for cell cycle control, namely, RB, p53, and cyclin-dependent kinase 2 (CDK2) (80). Meq has also been shown to interact with the molecular chaperone protein Hsp70 which, along with its several other functions, is also involved in neoplastic transformation (85).

Perhaps more convincing evidence for the transforming properties of Meq comes from overexpression studies using *meq* expressing plasmids, in which Meq is indeed able to transform Rat-2 and chicken fibroblast DF-1 cells (78,84). Moreover, the inability of *meq*-negative MDV mutants to induce tumours in infected birds underlines its
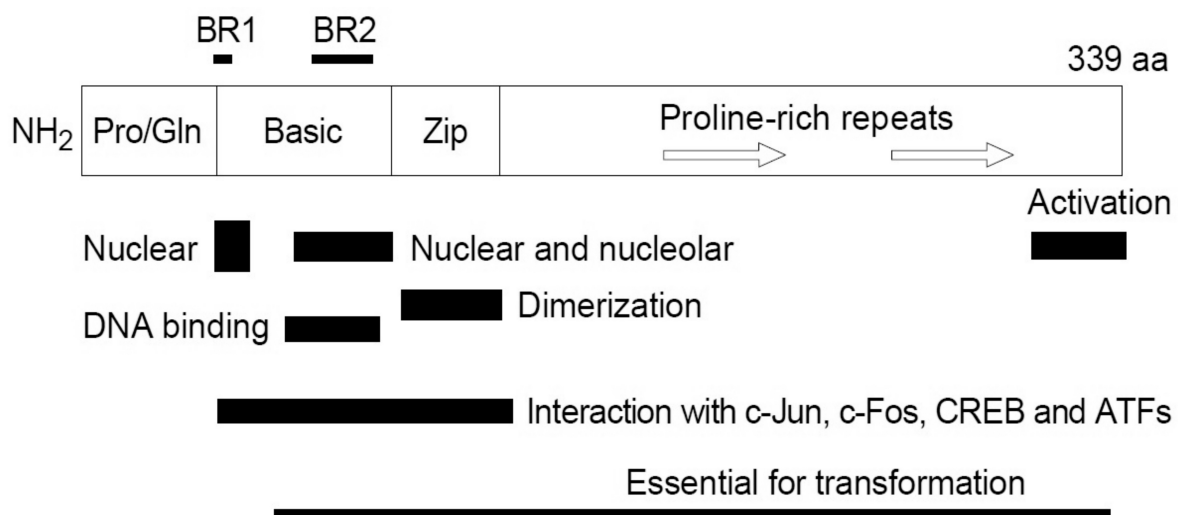
Figure 5. Schematic representation of MDV Meq protein.

The domains required for nuclear and nucleolar localization, transactivation, dimerization and interaction with c-Fos/c-Jun and DNA elements. Figure adapted from N.L.J. Ross, 1999 (73). Abbreviations: BR1: basic region 1; BR2: basic region 2; CREB: cAMP response element binding protein.

oncogenic potential. It is worth noting that these *meq*-negative mutant viruses were able to replicate *in vitro*, however had lost their oncogenicity *in vivo* (86).

5.4. Vaccines against MD

Although classical MD, fowl paralysis or neurolymphomatosis, was described in 1907, it did not have a large impact on the poultry industry until the 1960s, when MD almost devastated the poultry industry (44). Intensification of the poultry industry around this time had led to an increase in the virulence of the virus. By then, the lymphomatous form of the disease had become the dominant form. It occurred in younger birds causing very high mortality, and hence came to be referred to as the acute form (87). Identification and isolation of the causative agent in 1967 (36,37) were the key steps in development of anti-MD vaccines. The first live attenuated vaccine was produced within two years of identification of the virus by repeated passage of a virulent MDV strain, HPRS-16 (88,89). This made MD the first naturally occurring cancer in any species to be prevented by effective use of vaccines (87). The vaccine was licensed in the UK for commercial use. However, it was rapidly replaced by a vaccine formulated using HVT, a non-pathogenic but related virus (87,90). HVT soon gained popularity over MDV as a vaccine because it can be produced as a cell-free suspension enabling its easier storage in a lyophilized state. MDV, on the other hand, is highly cell-associated and requires storage in liquid nitrogen, which makes both its storage as well as transportation difficult and expensive (87).

The decade after the introduction of MD vaccines saw a great decrease in morbidity and mortality. However, first vaccine breaks were reported in the following decade. Moreover, viruses isolated from these vaccine breaks had increased virulence (very virulent [vv] MDV) (87,91). The problem intensified in the regions with a high density of poultry farms, especially in the USA. To tackle the problem of vvMDV, newer vaccine strains of MDV were introduced, which were usually administered in combination with HVT (e.g. HVT and the non-oncogenic SB-1 strain of MDV) (87,92). Nevertheless, the solution proved to be only temporary after viruses with even more virulence (very virulent + [vv+] MDV) emerged and caused vaccine breaks in the decade that followed (87).

The evasion of vaccine strategies by the virus by rapidly evolving into more pathogenic strains led Witter (93–95) to suggest that the increase in the potency of the vaccine regimen drives virus evolution (**Figure 6**). The imperfect or leaky vaccination hypothesis has been suggested as the evolutionary mechanism behind the increase in MDV virulence (96,97). An imperfect or leaky vaccine is defined as a vaccine that does not prevent infection but merely reduces disease symptoms (97). Even if experimental evidence for the leaky vaccination hypothesis is still unavailable, it has

been shown that MDV vaccination prolongs the infectious phase of the hyperpathogenic strains (98).

The emergence of vv+ MDV strains through vaccine breaks, nevertheless, demanded for a better and more potent vaccine. This demand was met by a new, very effective MD vaccine, CVI988/Rispens. Considered the gold standard vaccine (35,87,99), it was developed in the Netherlands based on a low pathogenic strain of MDV and remains the most effective vaccine against MD to date (100,101). Several alternative vaccine strains have been developed and are licensed for use in different parts of the world, e.g. MDV BH16 strain, HVT NSW 1/70 (Australia), MDV-2 (GaHV-3) Z4 strain (China) and a vaccine strain based on vvMDV Md11 (USA). However, none of them has been able to surpass the efficacy of the benchmark strain: CVI988/Rispens, or the bivalent vaccine (SB-1 and FC126 HVT) (87).

Perhaps a major breakthrough in both the study of the virus as well as vaccine development came after bacterial artificial chromosome (BAC) technology was introduced to MDV research in the early 2000s (102). The technology has allowed for the preparation of an experimental DNA vaccine against MD (103). Recombinant DNA technology and the use of large DNA viruses, including avipox and herpesviruses such as MDV and
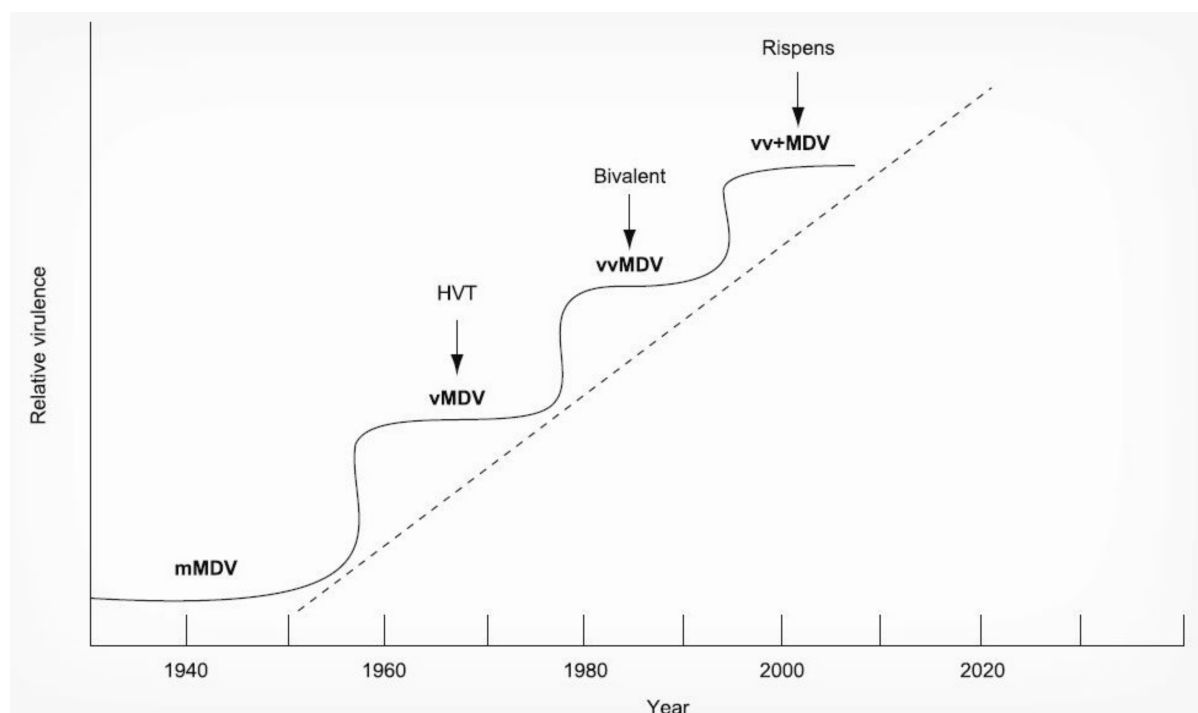
Figure 6. Evolution of MDV towards increased virulence since the 1940s.

Arrows indicate the time points at which vaccine were introduced: HVT (Herpesvirus of Turkeys), Bivalent (HVT and GaHV-3 strain SB-1) and Rispens (CVI988 strain). MDV pathotypes are represented with the prefixes: m: moderate, v: virulent, vv: very virulent and vv+: very virulent +. From Witter, 1998 (95).

HVT, as vectors has also been tested to produce experimental vaccines. Several different approaches have been tried with varying degrees of success. MDV genes, especially glycoprotein genes like gB, gC, gD, gE, gI and gH as well as U_L47 and U_L48, have been expressed in fowl pox virus and used as vaccines (87,104–107). The feasibility of expressing chicken genes, e.g. IL-2 and CD30, in recombinant MDV and HVT vaccines is also being studied in order to enhance specific immune responses (87). Another strategy studied is the augmentation of vaccines with interferon-gamma (γ) in order to prolong the secondary antibody response (108). None of these experimental vaccines, however, could outdo the efficacy of the gold standard and the hunt for a better vaccine continues.

5.5. Specific aims of the thesis

Live attenuated vaccines have played a pivotal role in MD control. Nevertheless, if MDV continues its trajectory of overcoming vaccine protection and increasing virulence, newer vaccines with improved potential would be required on a short notice (100). Consequently, one of the paramount goals of MD research today is the development of a vaccine superior to the gold standard, CVI988/Rispens. The fact that none of the new vaccines licensed after CVI988/Rispens can supersede its efficacy clearly demonstrates that the goal is difficult to achieve. Exploring newer avenues towards vaccine development might provide us with suitable vaccine candidates and steps are indeed being taken in this direction.

The goal of the present study was to examine if CPBD can be used as a viable strategy to attenuate MDV. Witter (109), in 1991, stated the properties of an ideal MD vaccine as follows: The vaccine should be derived from a vv+ MDV strain to assure maximum antigenic similarity to the highly pathogenic strains that chickens are likely to encounter in the field. It should not spread horizontally. It should be genetically stable and unable to recombine with the wild-type strains. It should also be able to replicate easily *in vitro* in order to ease large scale production. Strikingly, CPBD allows us to fulfil all these attributes. Moreover, by using CPBD, circulating highly pathogenic strains can be attenuated in a very short time to obtain vaccine strains with a perfect antigenic match, which might become very important given the history of vaccine breaks and the fear of such breaks occurring more frequently in the future (100).

The application of CPBD to attenuate viruses has so far only been studied in RNA viruses with much smaller genome size compared to MDV (**Table 2**). Therefore, obtaining comprehensive knowledge about the effects of CPBD in large DNA viruses, particularly MDV in the present study, is essential. However, CPBD of the entire viral genome is (still) impossible owing to its large size, and hence individual genes need to be

targeted. MDV genes can be divided into two classes based on their role in replication: the genes that are essential for efficient viral replication and those that are non-essential. The non-essential genes are usually involved in host control (immune evasion or pathogenesis) and herpesviruses, including MDV, can dedicate more than half of their genes to host control because of their large genomes (110). Since these genes are dispensable for viral growth in cell culture, their deoptimization should not interfere with viral growth *in vitro*, but at the same time due to their role in host control, such deoptimization would lead to reduced virulence and can possibly enhance immune response. To test this hypothesis **meq**, the major oncogene of MDV (see § 5.3.3.5.), was selected as the candidate gene for CPBD.

Besides being non-essential for viral growth *in vitro*, *meq* is one of the few genes that are actively expressed during latency and encodes a crucial immunogen (108,111,112). Remarkably, the vaccine strain CVI988/Rispens expresses two functional copies of *meq* (113). The properties of the Meq proteins produced by CVI988/Rispens are different than those produced by virulent MDV strains, however, they are able to transform human and rodent cells *in vitro* (114). Furthermore, given the fact that the CVI988 Meq proteins are unable to induce tumours in vaccinated chickens, it can be safely assumed that MDV strains possessing the *meq* gene can be rendered non-oncogenic by attenuation while retaining the characteristics of the parental virus, which is necessary for the induction of protective immune responses. To determine whether CPBD of the *meq* gene is a suitable strategy for MDV attenuation, three goals were set forth in the present study:

1. Generation and *in vitro* characterization of recombinant vvMDV strain RB-1B with CPB deoptimized *meq* gene,
2. To study the effects of CPBD on *meq* expression *in vitro,*
3. To study the effects of CPB deoptimized *meq* gene on MD progression and disease outcome *in vivo.*

6. Materials and Methods

6.1. Materials

All materials indicated below were used as per the instructions given by the manufacturer.

6.1.1. Chemicals, consumables and equipment

6.1.1.1. Chemicals

Name	Cat. No.	Manufacturer
Acetone ($(CH_3)_2CO$)	A160, 2500	Applichem, Darmstadt
Agar (bacteriological)	2266.2	Carl-Roth, Karlsruhe
Agarose - Standard Roti® grade	3810.4	Carl-Roth, Karlsruhe
Arabinose L (+)	A11921	Alfa Aesar, Karlsruhe
BSA (albumin bovine fraction V)	A6588.0100	Applichem, Darmstadt
CH_3COOH (acetic acid)	A3686, 2500	Applichem, Darmstadt
Chloroform	411 K3944831	Merck, Darmstadt
Dimethyl sulfoxide (DMSO)	1.02952.2500	Merck, Darmstadt
dNTP Mix (10 mM total)	BIO-39053	Bioline, Luckenwalde
Ethylenediaminetetraacetic acid (EDTA)	A2937, 1000	Applichem, Darmstadt
Ethidium bromide 1 %	2218.2	Carl-Roth, Karlsruhe
Ethanol, absolute for analysis C	A1613	Applichem, Darmstadt
Formamide deionized, Molecular biology grade	A2156	Applichem, Darmstadt
Glucose (α-D (+) glucose monohydrate)	303 K1468642	Merck, Darmstadt
Glycerol	A2926, 2500	Applichem, Darmstadt
Hydrochloric acid (HCl), 37 %	4625.2	Carl-Roth, Karlsruhe
Isopropyl alcohol (2-propanol)	A0892	Applichem, Darmstadt
β-mercaptoethanol (2-mercaptoethanol)	28625	Serva, Heidelberg
Magnesium chloride hexahydrate ($MgCl_2 \cdot 6H_2O$)	5833.025	Merck, Darmstadt
Sodium chloride (NaCl)	A3597, 5000	Applichem, Darmstadt
Sodium hydroxide (NaOH)	1.06462	Merck, Darmstadt
Peptone/Tryptone	A2210, 0250	Applichem, Darmstadt
Phenol/Chloroform	A0889, 0500	Applichem, Darmstadt
Roti™-Phenol	0038.3	Roth, Karlsruhe
Sodium dodecyl sulphate (SDS)	75746	Sigma-Aldrich, St Louis

Name	Cat. No.	Manufacturer
Sodium phosphate, monobasic, monohydrate ($NaH_2PO_4 \cdot H_2O$)	S9638	Sigma-Aldrich, St Louis
di-Sodium hydrogenphosphate dodecahydrate ($Na_2HPO_4 \cdot 12H_2O$)	A3906	Applichem, Darmstadt
Tris	A1086,5000	Applichem, Darmstadt
Triton X-100 detergent	8603	Merck, Darmstadt
Tween-20	9127.2	Roth, Karlsruhe
Water, molecular biology grade	A7398	Applichem, Darmstadt
Yeast extract granulated	212750	Becton-Dickinson, Heidelberg

6.1.1.2. Consumables

Name	Manufacturer
Cell culture dishes: 6-well, 24-well	Startsedt, Nümbrecht
Cell culture flasks: 25 ml, 75 ml	Startsedt, Nümbrecht
Conical test tubes: 17 × 120 (15 ml)	Startsedt, Nümbrecht
Conical test tubes 30 × 115 (50 ml), with and without skirted base	Startsedt, Nümbrecht
Cryotubes 1.8 ml	Nunc, Kamstrupvej
BD Falcon Cell Strainers [Cat. No. 352340]	BD Falcon, San Jose
Eppendorf tubes 1.5 and 2 ml	Sarstedt, Nümbrecht
Expendable cuvettes	Biodeal, Markkleeberg
Latex gloves	Unigloves, Troisdorf
Kimtech Science, Precision Wipes [Cat. No 05511]	Kimberly-Clark, Roswell
Nitrile gloves	Hansa-Medical 24, Hamburg
Parafilm® M	Bems, Neenah
Pipettes 5, 10, 25 ml	Sarstedt, Nümbrecht
Pipette tips P1000, 200, 100 and 10	VWR International, West Chester
Cell culture dishes: 60 mm, 100 mm, 150 mm	Sarstedt, Nümbrecht
Petri dishes for bacterial culture	Sarstedt, Nümbrecht
Sterile syringe filters PVDF 0.45 µm	VWR International, W. Chester
Transfection polypropylene tubes	TPP, Trasadingen
Whatmann blotting paper (WM Whatmann 3MM)	GE Healthcare, Freiburg

6.1.1.3. Equipment

Name	**Manufacturer**
General equipment	
ABI Prism 7500 Fast Real-time PCR system	Invitrogen Life Technologies, Grand Island
Bacterial incubator 07-26860	Binder, Turtlingen
Bacterial incubator shaker Innova 44	New Brunswick Scientific, New Jersey
Bunsen burner Type 1020	Usbeck, Radevormwald
Cell incubators Excella ECO-1	New Brunswick Scientific, New jersey
Centrifuge 5424, Rotor FA-45-24-11	Eppendorf, Hamburg
Centrifuge 5804R, Rotors A-4-44 and F45-30-11	Eppendorf, Hamburg
CytoFLEX Flow Cytometer	Beckman Coulter Life Sciences, Krefeld
Electroporator Genepulser Xcell	Bio-Rad, Munich
Electrophoresis power supply Power Source 250 V	VWR International, W. Chester
Freezer -20°C	Liebherr, Bulle
Freezer -80°C	GFL, Burgwedel
Galaxy mini centrifuge	VWR International, W. Chester
Gel electrophoresis chamber Mini Elektroforese System	VWR International, W. Chester
Gel electrophoresis chamber SUB-Cell GT	Bio-Rad, Munich
Ice machine AF100	Scotsman, Vernon Hills
INTEGRA Pipetboy	IBS Integrated Biosciences, Fernwald
Magnetic stirrer RH basic KT/C	IKA, Staufen
Nanodrop 1000	Peqlab, Erlangen
Neubauer counting chamber	Assistant, Sondheim/Rhön
Nitrogen tank ARPEGE70	Air liquide, Düsseldorf
Orbital shaker 0S-10	PeqLab, Erlangen
Pipetman P1000, P100, P10	VWR International, W. Chester
Perfect Blue™ Horizontal Maxi-Gel System	PeqLab, Erlangen
pH-meter RHBKT/C WTW pH level 1	Inolab, Weilheim
Sterile laminar flow chambers	Bleymehl, Inden
Thermocycler Flexcycler	Analytik Jena, Jena
Thermocycler GeneAmp PCR System 2400	PerkinElmer, Waltham

Name	**Manufacturer**
Thermocycler T-Gradient	Biometra, Göttingen
UV Transiluminator Bio-Vision-3026	PeqLab, Erlangen
Transiluminator printer P93D	Mitsubishi, Rüsselsheim
Transiluminator VL-4C, 1x4W-254 nm	Vilber-Lourmat, Eberhardzell
Vortex Genie 2™	Bender&Hobein AG, Zurich
Water baths TW2 and TW12	Julabo, Seelbach
Water bath shaker C76	New Brunswick Scientific, New Jersey

Microscopes	
Axiovert S 100 fluorescence microscope	Carl Zeiss MicroImaging GmbH, Jena
Axio-Observer.Z1 fluorescence microscope	Carl Zeiss MicroImaging GmbH, Jena
Microscope AE20	Motic, Wetzlar

6.1.1.4. Software

Name	**Version**	**Author/Company**
Axiovision software for Zeiss microscopes	4.8	Carl Zeiss MicroImaging GmbH, Jena
Chemi-Capt	-	Vilber-Lourmat, Eberhardzell
CytoFLEX CytExpert Software	1.2.11.0	Beckman Coulter Life Sciences, Krefeld
Graphpad Prism	7.02	Graphpad Software Inc., USA (115)
Image J	1.41	NIH, Bethesda
ND-1000	V.3.0.7	PeqLab, Erlangen
pDRAW32	V1.1.130	AcaClone Software (116)
Vector NTI	9	Invitrogen Life Technologies, Grand Island
Vision-Capt	-	Vilber-Lourmat, Eberhardzel

6.1.2. Enzymes and markers

Name	**Cat. No.**	**Manufacturer**
Enzymes		
BamHI	R0136	New England Biolabs, Ipswich
BamHI HF	R3136	New England Biolabs, Ipswich
BglII	R0144S	New England Biolabs, Ipswich
DpnI	ER1701	New England Biolabs, Ipswich

Name	**Cat. No.**	**Manufacturer**
EcoRI	R0101	New England Biolabs, Ipswich
EcoRI HF	R3101	New England Biolabs, Ipswich
EcoRV	R0195	New England Biolabs, Ipswich
HindIII	R0104	New England Biolabs, Ipswich
LongAmp® *Taq* DNA polymerase	M0323S	New England Biolabs, Ipswich
MluI	R0198S	New England Biolabs, Ipswich
NcoI	R0193S	New England Biolabs, Ipswich
NheI	R0131S	New England Biolabs, Ipswich
Phusion® High-Fidelity DNA Polymerase	M0530S	New England Biolabs, Ipswich
Proteinase K	7528.2	Carl-Roth, Karlsruhe
RNase A	A2760	Applichem, Darmstadt
SacI	R0156S	New England Biolabs, Ipswich
SnaBI	R0130S	New England Biolabs, Ipswich
Taq DNA Polymerase	01-1020	PeqLab, Erlangen
PerfeCTa® qPCR FastMix®, UNG, low ROX™	950778-01240	Quanta Biosciences inc, USA
XhoI	R0146S	New England Biolabs, Ipswich

Markers

Generuler ™ 1kb Plus DNA Ladder	SM0311	Fermentas, Mannheim
PageRuler™ Plus Prestained Protein Ladder, 10 to 250 kDa	26619	Thermo Fisher Scientific, USA

6.1.3. Plasmids

Cloning procedure for all the plasmids listed as 'present study' is described in § 6.2.2.4.

6.1.3.1. Cloning and expression vectors

Name	**Features**	**Company/Reference**
pEPkan-S	pEP vector containing kanamycin resistance gene *aphAI* and an I-SceI restriction site	(117)
pUC19	*E. coli* cloning vector; *amp^r*, ColE1 ori	New England Biolabs, Ipswich
pVitro2-Hygro-MCS	Multigenic eukaryotic expression plasmid with 2 multiple cloning cites (MCS) and hygromycin resistance gene	Invivogen, San Diego [Cat. No. pvitro-mcs]

6.1.3.2. Plasmid vectors containing *meq* gene variants

Name	**Features**	**Company/Reference**
pUC19-*meq*-W	pUC19 with *meq* gene from vvMDV strain RB-1B (*meq*-W)	Present study
pUC57-*meq*-D	pUC57 cloning plasmid containing synthesized CPB deoptimized *meq* gene (*meq*-D) of vvMDV strain RB-1B	Present study, Genes were recoded in house (§ 6.2.1.3.), synthesized commercially and delivered as pUC57 constructs by Bio Basic Inc., Canada
pUC57-*meq*-O	pUC57 cloning plasmid containing synthesized CPB optimized *meq* gene (*meq*-O) of vvMDV strain RB-1B	
pUC57-*meq*-R	pUC57 cloning plasmid containing synthesized CPB randomized *meq* gene (*meq*-R) of vvMDV strain RB-1B	

6.1.3.3. Transfer plasmids for *en passant* mutagenesis

Name	**Features**	**Reference**
pUC19-*meq*-W-*aphAI*	pUC19-*meq*-W with the kanamycin resistance gene and an I-SceI restriction site derived from pEPkan-S	Present study
pUC57-*meq*-D-*aphAI*	pUC57-*meq*-D with the kanamycin resistance gene and an I-SceI restriction site derived from pEPkan-S	Present study
pUC57-*meq*-O-*aphAI*	pUC57-*meq*-O with the kanamycin resistance gene and an I-SceI restriction site derived from pEPkan-S	Present study
pUC57-*meq*-D-*aphAI*	pUC57-*meq*-D with the kanamycin resistance gene and an I-SceI restriction site derived from pEPkan-S	Present study

6.1.3.4. Expression plasmids for immunoblotting

Name	**Features**	**Reference**
pVitro2-*EGFP*	pVitro2-Hygro-MCS with an enhanced green fluorescent protein gene cloned into one of the MCSs	(118)
pVitro2-*EGFP*-*meq*-W	pVitro2-*EGFP* with *meq*-W in the second MCS	Present study

Name	Features	Reference
pVitro2-*EGFP*-*meq*-D	pVitro2-*EGFP* with *meq*-D in the second MCS	Present study
pVitro2-*EGFP*-*meq*-O	pVitro2-*EGFP* with *meq*-O in the second MCS	Present study
pVitro2-*EGFP*-*meq*-R	pVitro2-*EGFP* with *meq*-R gene of vvMDV strain RB-1B in the second MCS	Present study

6.1.3.5. Expression plasmids for flow cytometry

Name	Features	Reference
pVitro2-*mtagBFP*	pVitro2-Hygro-MCS containing a blue fluorescent protein gene – *mtagBFP* (Evrogen, Russia)	Present study
pVitro2-*mtagBFP*-*meq*-W-*EGFP*	pVitro2-*tagBFP* with *meq*-W fused with enhanced green fluorescent protein gene (*EGFP*) in the second MCS	Present study
pVitro2-*mtagBFP*-*meq*-D-*EGFP*	pVitro2-*tagBFP* with *meq*-D fused with *EGFP* in the second MCS	Present study
pVitro2-*mtagBFP*-*meq*-O-*EGFP*	pVitro2-*tagBFP* with *meq*-O fused with *EGFP* in the second MCS	Present study
pVitro2-*mtagBFP*-*meq*-R-*EGFP*	pVitro2-*tagBFP* with *meq*-R fused with *EGFP* in the second MCS	Present study

6.1.4. Antibodies

Name	Cat. No.	Company/Reference
Chicken anti-MDV U_S2, polyclonal	-	(119)
Goat anti-mouse IgG (Fc specific)-peroxidase	A0168	Sigma-Aldrich, USA
Goat anti-Rabbit IgG, HRP linked	7074	Cell Signaling Technology, USA
Mouse anti-GFP IgG$_{2a}$ (B-2), monoclonal	sc-9996	Santa Cruz Biotechnology, Inc., USA
Rabbit anti-MDV Meq[1], polyclonal	-	(120)

---

[1] Rabbit anti-MDV Meq polyclonal antibody was kindly provided by Dr. Yoshihiro Izumiya, Department of Dermatology, University of California, Davis.

6.1.5. Animals, bacteria, cells, viruses

6.1.5.1. Animals

Name	Features	Reference
VALO® SPF chickens	Hatched in house from VALO® SPF layer chicken eggs	Lohmann Tierzucht, Cuxhaven

6.1.5.2. Bacteria

Name	Features	Company/Reference
DH10B	F⁻ endA1 recA1 galE15 galK16 nupG rpsL ΔlacX74 Φ80lacZΔM15 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) λ	Invitrogen, USA
GS1783	DH10B λcI857 Δ(cro-bioA)<>araC-P$_{BAD,}$ I-SceI	(121)

6.1.5.3. Cells

Name	Features	Company/Reference
CEC	Chicken embryonic fibroblasts/cells, primary cells, VALO specific pathogen free (SPF) strain	Primary cells
HEK 293T	Human epithelial kidney cell line transformed by human adenovirus 5, SV-40 T-antigen	ATCC CRL-11268

6.1.5.4. Viruses

Name	Features	Reference
vΔIR$_L$	BAC of vvMDV strain RB-1B, with the internal repeat long (IR$_L$) deleted	(77)
vΔIR$_L$-*meq*-D	BAC of vΔIR$_L$ with CPB deoptimized version of the *meq* gene replacing the parental version	Present study.
vΔIR$_L$-*meq*-D-rev	Revertant BAC of vΔIR$_L$-*meq*-D with parental version of the *meq* gene restored	(For procedures, see § 6.2.2.5.
vΔIR$_L$-*meq*-O	BAC of vΔIR$_L$ with CPB optimized version of the *meq* gene replacing the parental version	Generation of recombinant MDV
vΔIR$_L$-*meq*-O-rev	Revertant BAC of vΔIR$_L$-*meq*-O with parental version of the *meq* gene restored	and § 6.2.3.3.
vΔIR$_L$-*meq*-R	BAC of vΔIR$_L$ with CPB randomized version of the *meq* gene replacing the parental version	Reconstitution of viruses from BAC
vΔIR$_L$-*meq*-R-rev	Revertant BAC of vΔIR$_L$-*meq*-R with parental version of the *meq* gene restored	DNA)

6.1.6. Molecular biology kits

Name	Cat. No.	Manufacturer
E-Z® 96 DNA isolation kit	D1192-01	Omega Biotek, USA
Hi Yield Gel/PCR DNA Fragments Extraction Kit	30 HYDF100-1	SLG, Gauting
GF-1 AmbiClean PCR/Gel Purification Kit	GF-GC-200	Vivantis, USA
Qiagen Plasmid Midi kit	12145	Qiagen, Hilden
PeqGOLD Plasmid Mini Kit	12-6942-02	VWR Peqlab, Erlangen

6.1.7. Buffers and media

6.1.7.1. Buffers

Phosphate buffered saline (PBS)

2 mM KH_2PO_4

10 mM Na_2HPO_4

137 mM NaCl

2.7 mM KLC, pH 7.3

Tris-acetate-EDTA buffer (TAE)

40 mM Tris

1 mM $Na_2EDTA \cdot 2H_2O$

20 mM Acetic acid 99 %, pH 8.0

0.8 % Agarose Gel

0.8 % w/v Agarose

TAE buffer

0.5 µg/ml Ethidium bromide

10x SDS-PAGE running buffer

250 mM Tris, pH 8.5

1.9 M Glycine

1 % SDS

SDS-PAGE sample preparation

RIPA I

50 mM Tris-HCl, pH 7.5

150 mM NaCl

1 % Triton X-100

1 % Na-desoxycholate

0.1 % SDS

SDS-PAGE sample preparation

RIPA II

RIPAI

1mM EDTA

Buffer P1 (Resuspension buffer)

50 mM Tris HCL, pH 8.0

10 mM EDTA

100 µg/ml RNAse

Buffer P2 (Lysis buffer)

200 mM NaOH

1 % w/v SDS

Buffer P3 (Neutralization buffer)

3 M K-Acetate pH 5.5

Buffer TE

10 mM Tris HCl pH 7.4

1 mM Na_2EDTA

6.1.7.2. Media for bacterial cultures

LB medium (1 l)	**SOB medium (1 l)**	**SOC medium**
10 g BactoTM Tryptone	20 g BactoTM Tryptone	SOB medium
5 g BactoTM Yeast Extract	5 g BactoTM Yeast Extract	20 mM Glucose
10 g NaCl	0.584 g NaCl	
15 g BactoTM Agar	0.186 g KCl	
	pH 7.0	

6.1.7.3. Media and supplements for cell culture

Name	**Cat. No.**	**Manufacturer**
Dulbecco's MEM (DMEM)	F 0435	Biochrom AG, Berlin
Foetal bovine serum (FBS)	S 0415	Biochrom AG, Berlin
L-alanyl-L-Glutamine	K 0302	Biochrom AG, Berlin
Minimum essential Medium Eagle (MEM)	F 0315	Biochrom AG, Berlin
Sodium Pyruvate	L 0473	Biochrom AG, Berlin
Trypsin	L 2103-20G	Biochrom AG, Berlin

CEC Medium	**HEK 293T Medium**	**Trypsin**
MEM	DMEM	1.5 M NaCl
10 % FBS	10 % FBS	0.054 M KCl
1x Penicillin/Streptomycin	2 mM Na-Pyruvate	0.055 M $C_6H_{12}O_6$
	1 % L-Glutamine	0.042 M $NaHCO_3$
	1x Penicillin/Streptomycin	106 U Penicillin (P)
		1457.4 Streptomycin (S)
		0.0084 M Versene (EDTA)
		Trypsin 1:250

6.1.7.4. Antibiotics

Name [Cat. No.]	**Working concentration**	**Company**
Ampicillin [K0292]	100 µg/m diluted in ddH2O	Roth, Karlsruhe
Hygromycin B		
Kanamycin sulphate [T832.3]	50 µg/ml diluted in ddH2O	Roth, Karlsruhe
Chloramphenicol [3886.3]	30 µg/ml diluted in 96 % EtOH	Roth, Karlsruhe
Penicillin (P) [A1837]	100 U/ml diluted in MEM	Applichem, Darmstadt
Streptomycin (S) [A1852]	100 U/ml diluted in MEM	Applichem, Darmstadt

6.1.8. Primers

Table 3. Primers used for molecular cloning.

[a]Denotes the annealing orientation of the primer to the template. [b]Recognition site for the restriction enzyme in parenthesis has been underlined.

Construct	Orientation[a]	Primer sequence (5' → 3')[b]
pUC19-*meq*-W	Forward	ATATA<u>GAGCTC</u>CACCGTTGAAGTGATGTAATAT (SacI)
	Reverse	TATAT<u>GAGCTC</u>ATAGAAGCCATATCCCGAATAATACTC (SacI)
pUC19-*meq*-W-*aphAI*	Forward	AATT<u>AGATCT</u>AAGGACTGAGTGCACGTCCCTGCGTGTACA TAGGGATAACAGGGTAATCG (BglII)
pUC19-*meq*-W-*aphAI*	Reverse	AATT<u>AGATCT</u>CGAATTTCCTTACGTAGGTGTTCATTGGCC GCCAGTGTTACAACCAATTA (BglII)
pUC57-*meq*-D-*aphAI*	Forward	AATT<u>GGATCC</u>CGAACAGGATTCCCTATATTCCGGCCAGAT TAGGGATAACAGGGTAATCG (BamHI)
	Reverse	AATT<u>GGATCC</u>TCGGTAAGCCTAGCGTAAAGTCTCTCGGG TGCCAGTGTTACAACCAATTA (BamHI)
pUC57-*meq*-O-*aphAI*	Forward	AATT<u>CCATGG</u>CTGTTCCTTTGACTGTAACTCTTGGTCTTCT AGGGATAACAGGGTAATCG (NcoI)
	Reverse	TTAA<u>CCATGG</u>ACAGACAGGTTCATGGCAGGCCAGCTGAA GCCAGTGTTACAACCAATTA (NcoI)
pUC57-*meq*-R-*aphAI*	Forward	AATT<u>AGATCT</u>ATATCCAGTTTCCGTCGGACACCCAGTCTA TAGGGATAACAGGGTAATCG (BglII)
	Reverse	AATT<u>AGATCT</u>GGCCAGAGTAAAGAGAGTCCTGTTCAGGAT GCCAGTGTTACAACCAATTA (BglII)
pVitro2-*EGFP*-*meq*-W/D/O/R	Forward	ATATA<u>ACGCGT</u>CGGTACAGGTGTAAAGAGATG (MluI)
	Reverse	TATAT<u>ACGCGT</u>GTTGATCTTCCCGAAACTATG (MluI)
pVitro2-*tagBFP-meq*-W-*EGFP*	Forward	TATAT<u>ACGCGT</u>ATGTCTCAGGAGCCAGAGCC (MluI)

Construct	Orientation[a]	Primer sequence (5' → 3')[b]	
pVitro2-*tagBFP-meq*-D-*EGFP*	Forward	TATAT<u>ACGCGT</u>ATGTCTCAGGAACCCGAGCC	(MluI)
pVitro2-*tagBFP-meq*-O-*EGFP*	Forward	TATAT<u>ACGCGT</u>ATGTCGCAGGAGCCAGAACC	(MluI)
pVitro2-*tagBFP-meq*-R-*EGFP*	Forward	TATAT<u>ACGCGT</u>ATGTCCCAGGAGCCGGAG	(MluI)
pVitro2-*tagBFP-meq*-W/D/O/R-*EGFP*	Reverse	TATAT<u>ACGCGT</u>TTACTTGTACAGCTCGTCCATGC	(MluI)

Table 4. Primers used for *en passant* mutagenesis.

[a]Denotes the annealing orientation of the primer to the template. [b]Underlined nucleotides denote the annealing site of the primer to the template.

Construct	Orientation[a]	Primer sequence (5' → 3')[b]
pΔIR$_L$-Δ*meq*-*amp^r*	Forward	CAGGGAGAAGGCGGGCACGGTACAGGTGTAAAGAG<u>ATGTTCAAATATGTATCCGCTCATG</u>
pΔIR$_L$-Δ*meq*-*amp^r*	Reverse	TAGACGATGTGCTGCTGAGAGTCACAATGCGGATC<u>A TCAGAACGAAAACTCACGTTAAGG</u>
pΔIR$_L$-*meq*-D/O/R	Forward	<u>CCTAGGCAGGCGTCTCTTGC</u>
pΔIR$_L$-*meq*-D/O/R	Reverse	<u>ATCGATAAATAATGCCTTTAACCCTTTC</u>

Table 5. Primers used for sequence confirmation of constructs.

aDenotes the annealing orientation of the primer to the template.

Construct	Orientation[a]	Primer sequence (5' → 3')
pUC19-*meq*-W	Forward	CCCAGTCACGACGTTGTAAAACG
	Reverse	GAGCGGATAACAATTTCACACAGG
pUC19-*meq*-W-*aphAI* and pUC57-*meq*-D/O/R-*aphAI*	Forward	GTTGATGCGCTGGCAGTGTTCCTG
	Reverse	CAGGAACACTGCCAGCGCATCAAC
pVitro2-Hygro-MCS constructs	Forward	GCTAATTCAAAGCAACCGGTG
	Reverse	GAAACCTGCTCCTAGGGTCGAC
pΔIR$_L$-Δ*meq*-amp^r	Forward	TATGCAGTGCTGCCATAAC
	Reverse	ATTCTCTTACTGTCATGCCATC
pΔIR$_L$-*meq*-D/O/R	Forward	CCTAGGCAGGCGTCTCTTGC
	Reverse	ATCGATAAATAATGCCTTTAACCCTTTC

Table 6. Primers and probes used for qPCR.

Target gene	Primer/Probe	Sequence (5' → 3')
iNOS	Forward	GAGTGGTTTAAGGAGTTGGATCTGA
	Reverse	TTCCAGACCTCCCACCTCAA
	Probe	6-FAM-CTCTGCCTGCTGTTGCCAACATGC-TAMRA
ICP4	Forward	CGTGTTTTCCGGCATGTG
	Reverse	TCCCATACCAATCCTCATCCA
	Probe	6-FAM-CCCCCACCAGGTGCAGGCA-TAMRA

6.2. Methods

6.2.1. *In silico* methods

6.2.1.1. Calculation of CPS and CPB score

Codon pair score (CPS) is defined as the natural logarithm of the ratio of the observed to the expected number of occurrences of a particular codon pair over all coding regions in a species (7). The observed number of occurrences of a codon pair can be obtained by simply counting its actual occurrences in each set of genes. On the other hand, determining the number of its expected occurrences involves further calculations. It is calculated independently of codon bias as well as the amino acid frequency. In other words, the number of times an amino acid is encoded by a specific codon determines the expected frequency (**Figure 7A**) (7,10). A negative CPS signifies that a given codon pair is underrepresented in the ORFeome while a positive CPS signifies that it is overrepresented. The CPB score of an ORF can be then calculated as the arithmetic mean of the CPSs of all the codon pairs in that ORF (**Figure 7B**) (7).

Since MDV infects chicken, we assumed that it has adapted to the chicken CPB through co-evolution. Therefore, an algorithm was developed to calculate the CPS for each of the 3,721 codon pair combinations (61 × 61 codons, excluding stop codons) in context of the chicken ORFeome (*Gallus gallus*, breed Red jungle fowl, line UCD001) using

$$(\textbf{A}) \quad \text{CPS} = \ln\left(\frac{F(AB)_O}{\dfrac{F(A) \times F(B)}{F(X) \times F(Y)} \times F(XY)}\right)$$

$$(\textbf{B}) \quad \text{CPB score} = \sum_{i=1}^{k} \frac{\text{CPS}_i}{k-1}$$

Figure 7. Equations used to calculate codon pair score (CPS) and codon pair bias (CPB) score of chicken protein coding sequences.

(A) The equation used to calculate CPS, where A and B denote the codons, X and Y denote the amino acids and F denotes the frequency (number of occurrences). $F(AB)_O$ is the observed frequency of the codon pair AB and F(XY) is the frequency of the amino acid pair XY. The number of expected occurrences of a codon pair were calculated independent of the codon bias and amino acid frequency. This CPS score denotes if the given codon pair is overrepresented (+) or underrepresented (-) in the chicken ORFeome. **(B)** The equation used to calculate the CPB score for an entire ORF. CPB score is calculated as the arithmetic mean of the individual CPS of all the codon pairs present in the ORF.

44

15,762 predicted protein-coding ORFs (17). The calculations were done exactly as described (7,17). The CPSs obtained were used to determine the CPB score of the said 15,762 ORFs to analyse the distribution of the CPB of chicken genes. CPB scores were also calculated for 131 genes of the vvMDV strain RB-1B using the chicken CPSs.

6.2.1.2. Algorithm for gene recoding

A computer program to recode a protein coding sequence such that the resulting sequence has a desired CPB without altering the original amino acid sequence, was developed using the calculated CPSs. Recoding of genes to attain maximum CPB deoptimization/optimization is computationally extremely difficult and tedious, given the huge number of ways in which a protein can be encoded (10^{151} ways for a peptide containing 300 amino acid residues). A near-optimal solution to this problem, however, can be achieved by heuristic and metaheuristic approaches. Our recoding program utilizes a fast metaheuristic algorithm, called simulated annealing (122), to find a near-optimal approximation of the actual CPB extreme. The program recodes the sequences to maximize, minimize or randomize the CPB of a given sequence while preserving the codon bias, thus the codons are exactly the same as the parental sequence, however, synonymous codons are reshuffled, and hence occur in an altered order.

A foreseeable consequence of this reshuffling is inadvertent changes in RNA secondary structures, for instance, formation of stem-loop or hairpin-like structures. To ensure that the changes in protein expression are a result of the recoding and not a consequence of the undesired changes in RNA secondary structure, the free energy (ΔG) of folded RNA of the recoded gene needs to be maintained close to that of the parental gene. Our algorithm, similar to the algorithm by Coleman *et al.* (7), controls the ΔG ofthe RNA within a narrow range. The program was tested extensively to confirm its functionality and robustness. The algorithms described above were designed and written by Dr. Dušan Kunec, *Institut für Virologie, Freie Universität Berlin* (17).

6.2.1.3. Recoding of the major oncogene *meq* of MDV

The major oncogene of MDV, *meq,* was recoded to obtain its CPB deoptimized (*meq*-D) and optimized (*meq*-O) versions with near maximal levels of CPB deoptimization and optimization, respectively. As a control, a CPB 'randomized' version (*meq*-R was designed to have the codons so reshuffled as to maintain a comparable CPB score to that of the parental *meq* gene (*meq*-W). Sequences are listed in § 12.1.

As several splice variants of *meq* have been reported, it was essential that recoding does not hamper alternative splicing. Several potential splice donor and acceptor sites have been proposed within the *meq* ORF, but only one of them, the splice donor site D2 (cacctacGTaagga), has been shown to be functional (123,124).

This functional splice donor site, which has been identified for the Meq-sp variant, was preserved during recoding. The alternative splice variants, whose acceptor/donor sites lie outside of the *meq* ORF, were unaffected during the recoding.

To avoid inadvertent changes in RNA secondary structures that can be caused by recoding, the program maintains the ΔG of the recoded sequences close to that of the parental sequence. In order to confirm this, the sequences were scanned through mFold program (125) exactly as described (7). In short, an array of short sequences was generated from the coding sequences, 100 nucleotides in length with an 80-nucleotide long overlap with the preceding as well as the succeeding fragments. ΔG is calculated for each of these fragments by the program. If any of these fragments had ΔG lower than -30 kcal/mol, we recoded the fragment to raise its ΔG. Multiple sequence alignment was performed to assess the extent of the silent mutations that were introduced during recoding using Clustal Omega (v1.2.4.), an online tool available from EMBL-EBI (126,127). The recoded sequences were produced synthetically and cloned into pUC57 plasmids by Bio Basic Canada, Inc. To facilitate *en passant* mutagenesis, the sequences sent for commercial synthesis also included 80-nucleotide long arms homologous to the non-coding regions up- and downstream of the *meq* ORF.

6.2.2. Molecular biology methods

6.2.2.1. Plasmid and BAC DNA preparation

Isolation of BAC/plasmid DNA from bacterial cultures was performed using alkaline lysis as described previously (128). Bacterial cultures (3 – 5 ml) were grown overnight in lysogeny broth (LB) supplemented with appropriate antibiotics at either 32 °C (BACs, see § 6.2.2.5.) or 37 °C (plasmids) in a shaker incubator at 220 rpm. The cells were sedimented by centrifugation at maximum speed on a table top centrifuge and the pellets were resuspended in 250 µl of buffer P1 (See § 6.1.7.1. for buffer compositions). To this, 250 µl of buffer P2 (lysis buffer) were added and the tubes were incubated at room temperature for 5 minutes (min) for alkaline lysis. The mixture was neutralized by adding 250 µl of buffer P3 and incubating on ice for 5 min. The mixture was centrifuged at 20,000 × g for 5 min to remove cellular debris. Equal volume of Tris-buffered phenol-chloroform was added to the supernatants and the samples were thoroughly mixed followed by centrifugation at 20,000 × g for 10 min. The aqueous (upper) phase was transferred to a new tube and 0.7 volumes of 2-propanol were added to it. The samples were mixed by inverting the tubes and were centrifuged at 20,000 × g for 30 min for DNA precipitation. The resulting DNA pellet was washed twice with 70 % ethanol (ice-cold) and was air-dried at room temperature. The pellet was then dissolved in nuclease-free water and stored at -20 °C until further use. Medium scale

preparation (Midiprep) of plasmid/BAC DNA was carried out using Qiagen plasmid mini kit as per the instructions provided by the manufacturer. Small scale plasmid preparations were performed using peqGOLD plasmid miniprep kit, VWR Peqlab. The quality of the isolated DNA was assessed using a nanodrop spectrophotometer.

6.2.2.2. Polymerase chain reaction (PCR)

All PCRs were carried out as per the protocols provided by the respective manufacturer of the enzyme used. For cloning, Red mediated recombination and sequencing, the PCRs were performed using Phusion® High-Fidelity DNA Polymerase (New England Biolabs). Colony PCR and other such diagnostic PCR assays were performed using *Taq* polymerase (Peqlab). **Table 7** describes the cycling conditions used for PCR.

6.2.2.3. Preparation and transformation of competent bacteria

6.2.2.3.1. Chemically competent bacterial cells

Chemically competent bacterial cells were prepared using the calcium chloride ($CaCl_2$) method as described (128). Briefly, a single bacterial colony (*E. coli* strain DH10B) was grown in 5 ml LB overnight at 37 °C in a shaking incubator at 220 rpm. Then, 1 ml of this starter culture was inoculated into 100 ml of LB in a 500-ml conical flask and incubated in a shaking incubator at 220 rpm at 37 °C until the cells were in the logarithmic phase of growth, that is, until the optical density of the culture measured at a wavelength of 600 nm (OD_{600}) was in the range of 0.45 – 0.60. The cells were then rapidly chilled by incubation on ice for 10 min with agitation. After cooling,

Table 7. PCR cycling conditions.

[a]Specific annealing temperatures were determined for each primer set w.r.t. the polymerase used for the reaction by using the online Tm calculator provided by NEB (129).

Step	Phusion® High-Fidelity DNA Polymerase		*Taq* DNA Polymerase		Cycles
	Temperature	Duration	Temperature	Duration	
Initial denaturation	98 °C	30 s	95 °C	30 s	1
Denaturation	98 °C	15 s	95 °C	20 s	25–35
Annealing	45 – 72 °C [a]	30 s	45 – 72 °C [a]	45 s	
Elongation	72 °C	30 s/kb	72 °C	60 s/kb	
Final elongation	72 °C	300 s	72 °C	300 s	1

the cell suspension was aliquoted into pre-chilled 50-ml polypropylene tubes and centrifuged at 2,700 × g for 10 min. All the centrifugation steps were performed at 4 °C. Each of the cell pellets was resuspended in 30 ml ice-cold $CaCl_2$-$MgCl_2$ solution (20 mM $CaCl_2$, 80 mM $MgCl_2$) and was centrifuged at 2,700 × g for 10 min. After discarding the supernatant, each cell pellet was resuspended in 2 ml of ice-cold 0.1 M $CaCl_2$ supplemented with sterile glycerol (15 % v/v). These competent cells were aliquoted into pre-chilled 1.5-ml microcentrifuge tubes (100 µl/tube) and were either transformed immediately or stored at -80 °C until further use.

For transformation, ligation mixture (§ 6.2.2.4). was added to the competent cells and the cells were incubated on ice for 30 min. The cells were then subjected to a heat shock at 42 °C for 45 s followed by rapid chilling on ice for 5 min. The cells were suspended in SOC medium (1 ml) and incubated in a shaker-incubator at 37 °C for 1 h. The cells were then pelleted by centrifugation and plated on LB agar with appropriate antibiotic for selection and incubated at 37 °C overnight. On the following day, the bacterial colonies were replica plated and inoculated in LB for screening. Plasmid DNA was extracted from these overnight cultures by alkaline lysis as described (§ 6.2.2.1.) and the correct clones were identified by diagnostic restriction digestion followed by confirmation by sequencing. Clones containing the sequence confirmed plasmid constructs were stored at -80 °C in 20 % glycerol supplemented LB.

6.2.2.3.2. Electrocompetent bacterial cells

Electrocompetent *E. coli* cells were prepared as previously described (130). Single colony of the *E. coli* strain GS1783 containing the BAC to be modified were grown overnight in 5 ml of LB containing suitable antibiotics at 32 °C in a shaking incubator at 220 rpm. On the following day, 100 µl of this overnight culture were inoculated into 5 ml of fresh LB containing suitable antibiotics and incubated at 32 °C. Once the OD_{600} of these cultures reached 0.5 – 0.6, the tubes were transferred to a water bath at 42 °C shaking at 220 rpm for 15 min. This step is required to induce the Red recombination system (See § 6.2.2.5.). The cultures were then transferred to an ice-water bath and incubated on a shaker at 220 rpm for 20 min. The cells were centrifuged at 1,000 × g for 10 min at 4 °C. The pellets were washed three times with sterile ice-cold 10 % glycerol in ddH_2O. The electrocompetent cells thus prepared were suspended in 50 µl of ice-cold 10 % glycerol.

For electroporation, 100 – 200 ng of desalted, purified PCR product was added to the competent cells. The cells were then subjected to electroporation at 1.25 kV, 25 µF and 200 Ω. The electroporated cells were then suspended in 1 ml of pre-warmed

SOC medium and were incubated at 32 °C, 220 rpm on a shaking incubator for 2 h, followed by plating on selective LB agar.

6.2.2.4. Molecular Cloning

Primers used for PCR amplification of inserts for both transfer plasmids and eukaryotic expression plasmid vectors are listed in **Table 3**. The PCR products were separated based on their size by agarose gel electrophoresis. The products with correct sizes were excised from the gels and purified using Hi Yield Gel/PCR DNA fragments extraction kit as instructed by the manufacturer. The concentration and quality of the DNA was assessed using Nanodrop spectrophotometer (Peqlab). Plasmid vectors and the PCR amplified inserts were digested using appropriate restriction enzyme(s) for 1 h using the recommended buffer(s) at indicated temperature(s) as per the manufacturer's guidelines. The digested plasmid vectors were dephosphorylated using calf intestine phosphatase (CIP) for 1 h at 37 °C. The digested inserts and vectors were then gel purified and ligated using T4 DNA ligase (New England Biolabs) according to manufacturer's recommendation. The ligations were transformed into chemically competent cells and clones were screened as described in § 6.2.2.3.1.

To generate pUC19-*meq*-W-area, the wild-type *meq* (*meq*-W) gene was amplified using vRB-1B BAC (131) and was inserted into the SacI site of pUC19 (New England Biolabs). This plasmid was used to produce the transfer plasmid, pUC19-*meq*-W-area-*aphAI*, for *en passant* mutagenesis (described in § 6.2.2.5.) by inserting the *aphAI* kanamycin resistance gene and an I-SceI restriction site, hereafter together referred to as 'kanamycin cassette', into the BglII site present within the *meq* gene. The kanamycin cassette was amplified using pEP-Kan-S plasmid (117).

The commercially synthesized recoded versions of the *meq* gene, *meq*-D, *meq*-O and *meq*-R, were delivered as pUC57 clones by Bio Basic Inc., Canada (§ 6.1.3.2.). The sequences sent for commercial synthesis also included homologous arms up- and downstream of the *meq* gene to facilitate recombination during *en passant* mutagenesis. Transfer plasmids for these genes, pUC57-*meq*-D-*aphAI* (BamHI)*,* pUC57-*meq*-O-*aphAI* (NcoI) and pUC57-*meq*-R-*aphAI* (BglII)*,* were produced similarly by inserting the kanamycin cassette into the indicated enzyme sites of the respective plasmid.

The mutant *meq* genes as well as *meq*-W were cloned under the control of one of the two constitutive promoters of the eukaryotic dual expression vector pVitro2-hygro-MCS (Invivogen) using the MluI restriction site. A copy of the enhanced green fluorescent protein (*EGFP*) gene was already inserted into the other multiple cloning site (MCS) of the vector by Dr. Timo Schippers (118). These plasmids, namely, pVitro2-*EGFP-meq*-W, pVitro2-*EGFP-meq*-D*,* pVitro2-*EGFP-meq*-O and pVitro2-

EGFP-meq-R (**Figure 13A**), were used to assess the gene expression *in vitro* by immunoblotting (§ 6.2.4.1.).

For the assessment of gene expression by flow cytometry (§ 6.2.4.2.), each of the mutant versions and the wild-type *meq* gene were fused to an *EGFP* gene and inserted into the MluI site in one of the MCSs of pVitro2-hygro-MCS. A blue fluorescent protein gene, *mtagBFP* (Evrogen), was inserted in the second MCS (**Figure 14C**).

6.2.2.5. Generation of recombinant MDV

BAC technology enables rapid manipulation of cloned DNA sequences with the aid of well-established techniques under well-controlled conditions (132). We used '*en passant*' mutagenesis, a technique developed in our laboratory by Tischer, *et al.* (117). The technique is based on Red mediated recombination of dsDNA. The Red mediated recombination system is derived from bacteriophage lambda (λ) and consists of three proteins – Exo, Bet and Gam*.* The *E. coli* strain GS1783, a derivative of the DH10B strain, was developed to express the Red system under a temperature sensitive promoter which induces expression at 42 °C (121). Upon induction, Gam protein protects free dsDNA from degradation by the *E. coli* Rec B/C/D system. The 5' – 3' exonuclease, Exo, creates free single stranded 3'-overhangs in the introduced DNA, which are in turn protected by Bet. The Bet protein also facilitates strand invasion during the replication of BAC DNA which is an important step in homologous recombination. GS1783 cells also express a *S. cerevisiae* derived homing endonuclease, I-SceI, under the control of an arabinose inducible promoter. I-SceI recognises and cleaves an 18 base pair (bp) long sequence. Since this recognition sequence is rather long, it occurs rarely in genomes, and hence allows specific cleavage of the mutated region. This specific cleavage mediates the removal of the kanamycin cassette and the final recombination event. The major advantage of this two-step method is that the recombination happens in a scarless manner.

As the recoded *meq* variants have a high similarity with the wild-type sequence (See **Table 10**), there was a possibility of generating chimaeras during *en passant* mutagenesis as it is based on homologous recombination. Therefore, the mutagenesis was carried out in two steps as follows: (i) the wild-type *meq* gene was replaced with an ampicillin resistance marker (*amp^r*) by homologous recombination; (ii) the *amp^r* was replaced with the mutant *meq* variants by *en passant* mutagenesis.

Briefly, the *amp^r* gene was amplified by PCR from a pUC19 plasmid (§ 6.2.2.2.). The forward and reverse primers used were designed to have overhangs homologous to the non-coding regions upstream and downstream of the *meq* ORF, respectively (**Table 4**). The PCR product was gel purified, desalted and used to

transform electrocompetent *E. coli* GS1783 cells (§ 6.2.2.3.2.) containing the BAC, p∆IR$_L$ (77), a BAC of the vvMDV strain RB-1B in which the IR$_L$ has been deleted, and hence contains only one copy of *meq*. Colonies were screened for the presence of the modified BAC p∆IR$_L$-∆*meq-amp*^{r} by selecting for chloramphenicol resistance present in the BAC and the introduced ampicillin resistance. The clones were analysed by restriction fragment length polymorphism (RFLP) to check the integrity of the BAC. Successful deletion of the *meq* gene was confirmed by PCR and sequencing.

In the second step, the target DNA, consisting of the recoded variant of the *meq* gene (*meq*-D, *meq*-O or *meq*-R) with 80-nucleotide long homologous arms for recombination on either side, and a kanamycin cassette, was amplified using the transfer plasmids (§ 6.1.3.3.) as template and primers listed in **Table 4**. The PCR products were gel purified, desalted and used to transform electrocompetent *E. coli* GS1783 cells containing the BAC p∆IR$_L$-∆*meq-amp*^{r} (§ 6.2.2.3.2.). Kanamycin resistant clones were analysed by RFLP with at least two different restriction enzymes. After the second recombination step, in which the kanamycin cassette is removed, kanamycin sensitive clones were analysed by RFLP. The BACs with the expected restriction pattern were further confirmed for the insertion of the target *meq* variant by PCR and sequencing. The BACs, thus obtained, were p∆IR$_L$-*meq*-D, p∆IR$_L$-*meq*-O, and p∆IR$_L$-*meq*-R, with CPB deoptimized, optimized and randomized versions of the *meq* gene.

Revertants for each of these BACs, p∆IR$_L$-*meq*-D-rev, p∆IR$_L$-*meq*-O-rev, and p∆IR$_L$-*meq*-R-rev, were generated similarly by first replacing the mutant *meq* gene with an *amp*^{r} marker by homologous recombination followed by *en passant* mutagenesis to remove the *amp*^{r} and restore the wild-type *meq* gene.

6.2.3. Virological methods

6.2.3.1. Cell culture and maintenance

Primary chicken embryonic cells (CEC) were prepared from 11-day-old VALO specific pathogen free (SPF) eggs (VALO®, Lohmann Tierzucht, Cuxhaven) as described (133). Briefly, eggs were surface sterilized with 70 % ethanol and their shells were carefully cracked. The embryos were transferred to a petri dish containing sterile phosphate buffered saline (PBS). The internal organs as wells as the extremities of the embryo were removed. The remaining torso was teased with the help of forceps to obtain small pieces of the tissue, which were then washed with PBS on a magnetic stirrer for 10 min. The washed tissue was digested with 100 ml of 0.05 % trypsin solution. The cell suspension thus obtained was filtered through a sterile gauze into Eagle's

minimum essential medium (MEM, Biochrom AG, Berlin) containing 10 % foetal bovine serum (FBS) and 1 % penicillin/streptomycin. The tissue was digested two more times. The cell suspension was centrifuged at 1,200 rpm for 10 min. The cell pellets were pooled and washed again with MEM + 10 % FBS before seeding at a desired density onto sterile cell culture dishes. The cells were maintained in MEM + 1 – 10 % FBS at 37 °C in a 5 % CO_2 atmosphere. To passage confluent CEC, the medium was discarded, cells were washed with PBS followed by detachment with 0.05 % trypsin at 37 °C. The detached cells were resuspended in MEM + 10 % FBS and split in desired ratios onto new culture dishes. The FBS in the medium also served to inactivate any residual trypsin.

Human embryonic kidney 293T (293T ATCC® CRL-3216) cells were grown in Dulbecco's minimal essential medium (DMEM, Biochrom AG, Berlin) supplemented with 10 % FBS and 1 % penicillin/streptomycin.

6.2.3.2. Transfection of CEC and 293T cells

Cells grown in 6-well plates (1×10^6 cells/well) were transfected using polyethelenimine (PEI) as described (134). Briefly, the cells were washed once with sterile PBS and 2 ml of antibiotic-free MEM + 10 % FBS were added per well. DNA (1 μg) was added to 100 μl of Opti-MEM (Gibco, Thermo Fischer Scientific) and mixed briefly. A 0.1 % solution of PEI (pH 7.0, 10 μl) was added to this DNA solution and incubated at room temperature for 20 min. Following incubation, the transfection mixture was added dropwise to the cells. The cells were incubated at 37 °C in a 5 % CO_2 environment for 4 h after which the medium was changed to MEM + 10 % FBS + 1 % penicillin/streptomycin.

6.2.3.3. Reconstitution of viruses from BAC DNA

CEC were transfected as described in § 6.2.3.2. above with 1 μg of high quality BAC DNA. The FBS concentration was reduced to 0.5 % once the cells reached confluency. The cells were observed regularly for viral plaque formation. Once visible viral plaques had developed around 5 to 6 dpi, the viruses were grown further by co-culture of infected and uninfected cells at appropriate ratios on 10 cm dishes. The dishes were then incubated until development of visible plaques. Since MDV is a cell associated virus, virus stocks were prepared by cryopreservation of infected cells as described in § 6.2.3.4. below.

6.2.3.4. Cryopreservation and revival of cryopreserved cells

Infected or uninfected cells were trypsinized, washed with PBS and suspended in MEM supplemented with 10 % FBS, 1 % penicillin/streptomycin and 10 % dimethyl

sulfoxide (DMSO). The cells were aliquoted in cryovials and slowly frozen overnight in cryocontainers filled with 2-propanol at -80 °C before being transferred to liquid nitrogen for long term storage.

For revival, the cryopreserved cells were rapidly thawed in a water bath at 37 °C and immediately transferred to pre-warmed MEM + 10 % FBS and seeded onto cell culture dishes (uninfected cells) or co-seeded (infected cells) with fresh CEC.

6.2.3.5. Immunofluorescence assay

Standard immunofluorescence assay (IFA) using chicken anti-MDV-U$_S$2 polyclonal antibody was performed as described (119). Briefly, infected cells were fixed with 4 % formaldehyde followed by blocking for 20 min at room temperature with 3 % bovine serum albumen (BSA) in PBS. Subsequently, the cells were incubated with chicken polyclonal anti-MDV-US2 antibody (119) diluted 1:2,000 in PBS + 1 % BSA for 45 min at room temperature. The monolayers were then washed three times with PBS before incubation with the secondary antibody, goat anti-chicken IgG (heavy and light chains) Alexa 488, diluted 1:1,000. The plates were then observed under the Axiovert S 100 fluorescence microscope (Carl Zeiss GmbH).

6.2.3.6. Virus titration

One vial each of the frozen virus stocks to be titrated was rapidly thawed and diluted in MEM + 5 % FBS in a 10-fold series. Fresh CEC (1×10^6 cells/well) were mixed with 100 µl of the diluted virus and seeded onto 6-well plates in duplicates. Once the cells were confluent, the serum concentration was reduced to 0.5 %. After 6 days, the cells were fixed with a 4 % formaldehyde solution and IFA was performed for the visualization of plaques as described in § 6.2.3.5. Plaques were counted under the Axiovert S 100 fluorescence microscope and virus titres for the stocks were determined as plaque forming units per ml (PFU/ml).

6.2.3.7. Plaque size assay

To determine the cell to cell spread of the viruses, plaque size assays were performed as described (135). Fresh CEC (1×10^6) were mixed with virus suspensions containing 100 PFU and seeded on to 6-well plates in MEM + 5 % FBS. The serum concentration was reduced to 0.5 % once the cells were confluent. After 6 days, IFA was performed as described in § 6.2.3.5. Pictures of a total of 153 individual plaques were taken for each virus at a 100-fold magnification in three independent experiments with the help of the Axiovert S 100 fluorescence microscope (Carl Zeiss GmbH). The corresponding plaque areas were measured using ImageJ software version 1.48v (136) and were mathematically transformed

into plaque diameter lengths. Diameter values were plotted relative to those of vΔIR$_L$ using GraphPad Prism 7.02 (115).

6.2.3.8. Multi-step growth kinetics

Multi-step growth kinetics were performed by co-seeding fresh CEC (1 × 10^6) with 100 PFU of the respective virus as described (135). One well of infected cells was trypsinized every day until 6 dpi and resuspended in 1 ml of culture medium. This suspension of infected cells i.e. the resulting virus was titrated as described in § 6.2.3.6.

6.2.4. *In vitro* gene expression assay

6.2.4.1. Immunoblot analysis

293T cells were transfected with 1 µg of pVitro2-*EGFP-meq*-W, pVitro2-*EGFP-meq*-D, pVitro2-*EGFP-meq*-O or pVitro2-*EGFP-meq*-R as described in § 6.2.3.2. Transfected cells were harvested 24 h post transfection and lysed in RIPA buffer supplemented with Complete® mini protease inhibitor and phosphatase inhibitor cocktail. The lysates were subjected to SDS-polyacrylamide gel electrophoresis (SDS-PAGE), and the separated proteins were transferred to nitrocellulose membranes by semi-dry blotting. The membranes were blocked with 5 % non-fat dry milk in PBS for 1 h on a shaker at room temperature followed by overnight incubation at 4 °C with rabbit polyclonal anti-Meq antibody (120) or mouse anti-GFP monoclonal IgG$_{2a}$ (Santa Cruz Biotechnology, Inc), diluted 1:1,000 and 1:500, respectively, in blocking buffer. The membranes were washed with PBS containing 0.1 % Tween 20 (PBS-T) followed by incubation for 1 h with either goat anti-rabbit IgG-HRP (Cell Signaling Technology) or goat anti-mouse IgG-peroxidase antibody (Sigma-Aldrich), respectively. Both the secondary antibodies were diluted 1:10,000 in blocking buffer. After washing with PBS-T, the membranes were incubated with enhanced chemiluminescence western blotting detection reagent (ECL-Prime, GE Healthcare) and signal was recorded using a Chemi-Start 5100 detection system.

6.2.4.2. Flow cytometry

For quantitative analysis of gene expression by flow cytometry, 293T cells were transfected as described (§ 6.2.3.2.) with 1 µg of pVitro2-*tagBFP-meq*-W-*EGFP*, pVitro2-*tagBFP-meq*-D-*EGFP*, pVitro2-*tagBFP-meq*-O-*EGFP* or pVitro2-*tagBFP-meq*-R-*EGFP*. The cells were washed with PBS 24 h post transfection, trypsinized and suspended in FACS buffer. The cell suspensions were analysed using a CytoFLEX flow cytometer (Beckmann Coulter Life Sciences) equipped with 405 nm and 488 nm

lasers and the following bandpass filters: 525/40 nm (GFP) and 450/45 nm (BFP). The data were analysed using CytoFLEX CytExpert Software version 1.2.11.0.

6.2.5. *In vivo* methods

6.2.5.1. Infection experiment

To study the effect of CPBD on MD progression and disease outcome *in vivo*, an infection experiment was performed on SPF layer chickens. The animal experiment was approved by the *Landesamt für Gesundheit und Soziales*, Berlin, Germany (Approval No. G 0218/12) and was performed at the animal facility of the *Institut für Virologie*, *Freie Universität Berlin*. SPF layer chicken eggs (VALO®, Lohmann Tierzucht, Cuxhaven) were incubated and hatched in house. One-day old chickens were neck-tagged and randomly distributed in different groups. They were housed in isolated rooms at 25 °C with a 12 h light program in stainless steel cages on wooden bedding enriched with straw. Up to 10 days post hatch, an additional heating lamp was used per cage. Food and water were provided *ad libitum*.

One-day old chickens were infected subcutaneously in a blinded fashion with 5,000 PFU of the mutant viruses: vΔIR$_L$-*meq*-D, vΔIR$_L$-*meq*-O or vΔIR$_L$-*meq*-R; or their respective revertants: vΔIR$_L$-*meq*-D-rev, vΔIR$_L$-*meq*-O-rev, vΔIR$_L$-*meq*-R-rev; or the parental virus vΔIR$_L$. Each group consisted of 25 birds for the mutant and revertant viruses and 10 for the parental virus. Eleven uninfected chickens (contact birds) were housed with each infected group to study the transmission of the virus via the natural route. Following infection, the inocula were back titrated on CEC to determine the minimum infectious dose per chicken. Animals that died in the first week post infection were excluded from experimental analysis as early mortality. Animals were monitored daily for clinical MD symptoms, which served as a criterion for termination. Animals were examined *post mortem* for tumorous lesions either upon development of MD symptoms or at the end of the experiment 91 dpi.

6.2.5.2. Quantification of MDV genome copies in chicken whole blood

To evaluate the ability of the mutant viruses to replicate *in vivo*, the number of MDV genome copies in whole blood samples was quantified using quantitative real-time PCR (qPCR). Blood samples (40 µl) were collected from wing veins of the infected as well as contact chickens in a 100 mM EDTA solution (20 µl) in deep well plates. Eight chickens per group were sampled at different time points; on days 7, 14, 21, and 28 post infection for infected chickens and days 21, 28, 35, and 42 for contacts. DNA was isolated using E-Z96 96-well blood DNA isolation kit (Omega Biotek, USA) according to manufacturer's instructions.

For determination of MDV genome copy number in the blood samples, the *ICP4* gene was used as a target for amplification by qPCR using a 7500 Fast Real-Time PCR System (Applied Biosystem, USA) as previously described (119). DNA loading for each sample was normalized by qPCR amplification of a chicken house-keeping gene, inducible nitric oxide synthase (*iNOS)*. The primers and probes used for qPCR are listed in **Table 6**.

For the generation of standard curves plasmids containing either *ICP4* or *iNOS* genes were used. Serial 10-fold dilutions of the respective plasmid DNA (range: $10 - 10^7$ copies/µl) were used for generating the standard curves. The standard curve was determined by potting the cycle threshold (C_T) value at each dilution against the total copy number in the respective dilution. A C_T value of > 40 indicated no amplification of the specific target DNA, thus any sample with a C_T value > 40 had 0 copies of the target DNA.

All qPCR assays were performed using PerfeCTa® qPCR FastMix®, UNG, low ROX™, 5 µl template DNA, 25 pmol of each gene-specific primer, and 10 pmol of the gene specific probe in a 20 µl reaction. Thermal cycling conditions were as follows: 95 °C for 20 s, followed by 40 cycles at 95 °C for 3 s and 60 °C for 30 s. The C_T values for the respective genes were used to determine the copy numbers of *ICP4* and *iNOS* in each individual sample with the help of the generated standard curves. MDV genome copies per cell were calculated by dividing the number of *ICP4* copies by the number of *iNOS* copies. This number was further multiplied by 10^6 and the final values were expressed as MDV genome copies per 1×10^6 cells.

6.2.6. Statistics

Statistical analysis was performed using GraphPad Prism 7.02 (115). Data were first tested for normal distribution. The data for plaque size assay were analysed for statistical significance by one-way ANOVA, with Bonferroni correction for multiple comparisons. Growth curves and qPCR data were analysed using the Kruskal–Wallis non-parametric test for significance. MD and tumour incidence data were analysed by Fisher's exact test and the p-values were adjusted using the Bonferroni correction for multiple comparisons.

7. Results

7.1. Calculation of CPS and CPB score

CPB has been found in all the species studied so far and varies among different species (7). We assumed that, since MDV infects chicken, it must have adapted to the chicken CPB through the process of co-evolution to achieve optimal translational efficiency. Data on CPB of several species are available and mammalian CPB has been shown to be, by and large, the same (18). However, chicken CPB was still unknown.

CPB score, the observed frequency of codon pairs in ORFs, is calculated for an ORF as the arithmetic mean of the CPSs of all the constituent codon pairs in that ORF (**Figure 7B**). Therefore, we calculated the CPS for all the possible 3,721 codon pair combinations (61 × 61, excluding stop codons) for the chicken ORFeome, based on 15,762 predicted ORFs (*Gallus gallus*, breed Red jungle fowl, line UCD001). CPS is defined as the natural logarithm of the ratio of the observed to the expected number

Table 8. Comparison of calculated CPS of 16 possible codon pair combinations that encode the amino acid pair alanine-alanine in human and chicken ORFeomes.

Obs and Exp denote the observed and expected number of occurrences, respectively. CPS is the calculated codon pair score. Positive CPS signifies that the codon pair is overrepresented while a negative CPS signifies that the codon pair is underrepresented. Individual codon pairs are colour coded for easier visualization of identical codon pairs. Human CPS data was published by Coleman *et al*., in 2008 (7).

Homo sapiens					*Gallus gallus*				
Codon pair	**Obs**	**Exp**	**Obs/Exp**	**CPS**	**Codon pair**	**Obs**	**Exp**	**Obs/Exp**	**CPS**
GCG-GCG	4195	874	4.80	1.57	GCG-GCG	3906	760	5.14	1.64
GCG-GCC	5603	3201	1.75	0.56	GCG-GCC	3394	2043	1.66	0.51
GCT-GCT	7477	5075	1.47	0.39	GCT-GCT	8457	5988	1.41	0.35
GCA-GCA	5569	3895	1.43	0.36	GCA-GCA	7427	5505	1.35	0.30
GCA-GCT	6292	4446	1.42	0.35	GCT-GCC	7029	5734	1.23	0.20
GCT-GCA	5807	4446	1.31	0.27	GCA-GCT	6597	5741	1.15	0.14
GCT-GCC	9505	7713	1.23	0.21	GCT-GCA	6574	5741	1.15	0.14
GCA-GCC	8180	6757	1.21	0.19	GCA-GCG	2229	2046	1.09	0.09
GCA-GCG	1923	1845	1.04	0.04	GCA-GCC	5492	5498	1.00	0.00
GCT-GCG	2103	2106	1.00	0.00	GCT-GCG	2018	2134	0.95	-0.06
GCG-GCT	2014	2106	0.96	-0.04	GCC-GCG	1923	2043	0.94	-0.06
GCC-GCG	2920	3201	0.91	-0.09	GCC-GCC	4505	5491	0.82	-0.20
GCG-GCA	1593	1845	0.86	-0.15	GCG-GCT	1314	2134	0.62	-0.48
GCC-GCC	7118	11723	0.61	-0.50	GCG-GCA	1128	2046	0.55	-0.60
GCC-GCT	1888	7713	0.24	-1.41	GCC-GCT	1243	5734	0.22	-1.53
GCC-GCA	1518	6757	0.22	-1.49	GCC-GCA	900	5498	0.16	-1.81

of occurrences of a codon pair (**Figure 7A**). The calculated chicken CPSs of all possible codon pairs are published (17). A negative CPS signifies that a given codon pair is underrepresented in the ORFeome while a positive CPS signifies that it is overrepresented. Upon comparison of the calculated chicken CPS with the previously published human CPS (7), it was found that the CPSs for both the species were different, yet show a certain degree of similarity. The extent of this similarity is demonstrated in **Table 8**, which compares the calculated CPSs of 16 possible codon pair combinations that encode the amino acid pair alanine-alanine in human and chicken ORFs. The most overrepresented (green cells) and the most underrepresented (red cells) codon pairs are identical in both species.

Using the calculated CPS values, we determined the CPB score for each of the 15,762-predicted chicken ORFs that were used to calculate the CPS. A negative CPB score indicates that the ORF contains largely underrepresented codon pairs while a positive CPB score indicates that the ORF contains mainly overrepresented codon pairs. We also calculated the CPB for 131 ORFs of the vvMDV strain RB-1B using the chicken CPS. Individual CPB scores analysed for each of the chicken as well as MDV genes were plotted against the length (number of codon pairs) of the respective genes (**Figure 8**). The overall distribution of chicken CPB scores proved to be similar to that of the human CPB scores (7,18). Majority of the chicken ORFs had a positive CPB score with
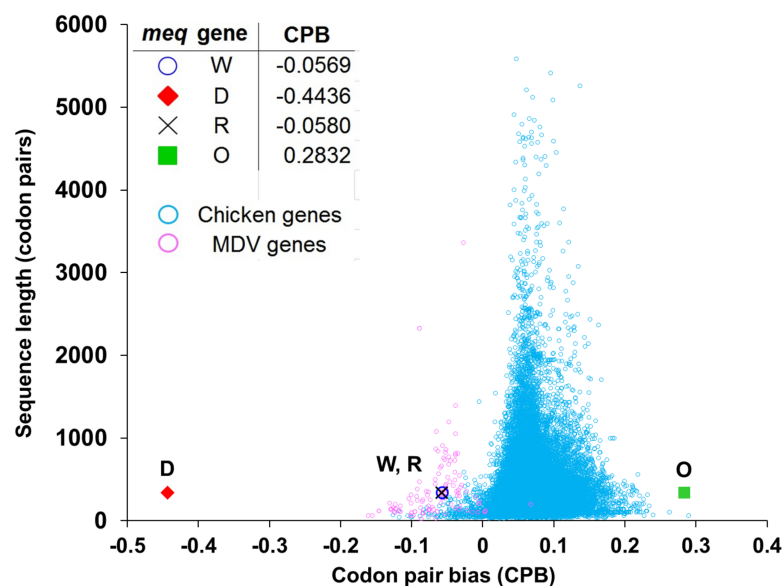
Figure 8. Distribution of calculated CPB scores of 15,762 predicted chicken ORFs and 131 vvMDV strain RB-1B ORFs.

The CPB scores of each the chicken genes (blue circles) and MDV genes (purple circles) were plotted against the length of the respective genes (in codon pairs). W represents the CPB score of the *meq* gene of vvMDV strain RB-1B, while R, D and O represent the CPB scores of the recoded versions of the *meq* gene – *meq*-R, *meq*-D and *meq*-O, respectively.

an average of 0.0755. Most of the MDV genes, however, had a slightly negative CPB score with an average of -0.0646.

7.2. Recoding of MDV *meq* gene

Based on the calculated chicken CPS, an algorithm was developed to recode a given protein coding sequence to obtain a new one with a desired CPB score, while maintaining the original amino acid sequence. The major oncogene *meq* of the vvMDV strain RB-1B was recoded to deoptimize (*meq*-D), optimize (*meq*-O) and preserve (*meq*-R) the CPB of the parental (*meq*-W) sequence. The CPB scores for the recoded genes were -0.442 for *meq*-D, 0.283 for *meq*-O while *meq*-R and *meq*-W had highly similar CPB scores of -0.058 and -0.057, respectively (**Figure 8**). The recoded sequences can be found in the supplementary information (§ 12.1.) **Table 9** illustrates the outcome of recoding, namely the introduction of point mutations, with the help of the first 48 nucleotides of the sequences.

To ensure that recoding does not lead to stable changes in the RNA secondary structure, the recoding algorithm keeps the free folding energy (∆G) of the sequences tightly controlled and close to that of the parental sequence. To confirm this, the recoded sequences were scanned through mFold program (§ 6.2.1.3.) (125). The sequence fragments (100 nucleotides in length, with an 80 nucleotide overlap on either side) that had a ∆G of less than -30 kcal/mol were recoded again to increase their ∆G. The final sequences had a similar ∆G distribution to the parental sequence (**Figure 9**).

7.2.1. Characterization of the recoded *meq* sequences

The three recoded versions of the *meq* gene were designed to contain exactly the same codons as the parental version, but the order of the codons was reshuffled to obtain the desired CPB. To assess the effect of this reshuffling on the sequences a multiple sequence alignment was performed using Clustal Omega version 1.2.4. (126,127) (For sequence alignment, see § 12.2.) and the proportion of the nucleotides and codons which are

**Table 9. First 48 nucleotides of the recoded variants of the MDV *meq* gene.**

The protein encoded by the recoded sequences *meq*-D, *meq*-O and *meq*-R is identical to that encoded by parental sequence, *meq*-W. [a]Nucleotide changes, w.r.t. the parental sequence, are highlighted.

Name	Sequence[a]
meq-W	ATG-TCT-CAG-GAG-CCA-GAG-CCG-GGC-GCT-ATG-CCC-TAC-AGT-CCC-GCT-GAC…
meq-D	ATG-TCT-CAG-GA**A**-CC**C**-GAG-CC**A**-GGC-GCT-ATG-CC**T**-TA**T**-**TCG**-CC**A**-GC**C**-GAC…
meq-O	ATG-TC**G**-CAG-GAG-CCA-GA**A**-CC**A**-GG**G**-GC**C**-ATG-CCC-TAC-**TCT**-CC**A**-GC**A**-GA**T**…
meq-R	ATG-TC**C**-CAG-GAG-CC**G**-GAG-CC**T**-GGC-GCT-ATG-CC**A**-TA**T**-**TCA**-CC**A**-GC**C**-GA**T**…
Protein	M S Q E P E P G A M P Y S P A D…

identical at the same positions in the parental and the recoded sequences was calculated (**Table 10**). The three recoded sequences had an average difference of 57 % compared to the parental sequence with respect to codons that occupy the corresponding position. Since synonymous codons tend to share the first two nucleotides, the sequence similarity is higher at the nucleotide level than at codon level.

7.3. Generation of recombinant MDV BACs

To assess the effects of CPBD of the *meq* gene *in vitro*, the recoded sequences were commercially synthesized (Bio Basic Inc.) and were obtained as plasmid constructs in the vector pUC57. The three recoded versions of the *meq* gene, *meq*-D, *meq*-O and *meq*-R, were then used to replace the parental gene, *meq*-W, in the virus. This was done using the BAC system of the vvMDV strain RB-1B with the help of *en passant* mutagenesis (§ 6.2.2.5.) (121). MDV contains two copies *meq*, one each in
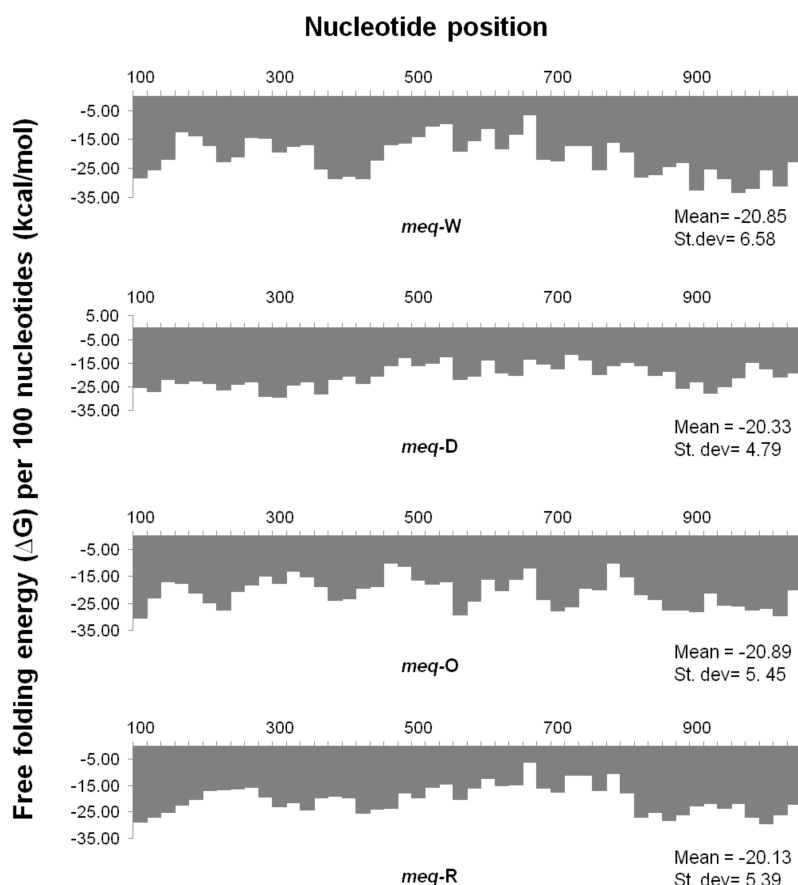
Figure 9. Free folding energy (ΔG) of the RNA encoded by the wild type (*meq*-W), CPB deoptimized (*meq*-D), optimized (*meq*-O) and randomized (*meq*-R) versions of the MDV *meq* gene.

ΔG was calculated for 100 nucleotide long sequence fragments with an 80 nucleotide overlap with each other. None of the fragments had a ΔG of less than -30 kcal/mol. All recoded versions of the *meq* gene have their mean ΔG similar to that of the parental sequence.

the IR_L and TR_L regions. To ensure complete replacement of *meq*-W with one of the mutants, a recombinant BAC of vvMDV RB-1B in which the IR_L has been deleted, ΔIR_L (77), was used. It has been shown that the reconstituted virus, $v\Delta IR_L$, is able to restore the deleted IR_L in cell culture as early as two passages after reconstitution (77), hence, ensuring a complete viral genome with two identical copies of either the parental *meq*-W gene (ΔIR_L BAC) or one of the recoded versions thereof (ΔIR_L-*meq*-D, ΔIR_L-*meq*-O or ΔIR_L-*meq*-R BACs). Another factor that was considered while generating the recombinant BACs was the average nucleotide similarity of 79.08 % (**Table 10**) between the recoded and the parental *meq* genes. This extent of similarity might have caused the generation of chimaeras with partial parental *meq*-W genes since *en passant* mutagenesis is based on homologous recombination. To avoid this, *meq*-W in the ΔIR_L BAC was initially replaced with an *amp^r*, which was in turn replaced with one of the recoded versions of the *meq* gene (§ 6.2.2.5.).

To confirm that the phenotypic changes in the mutant viruses are caused by the recoding and not because of unexpected mutations elsewhere in the sequence, revertant BACs were generated from each of the BACs containing the recoded *meq* genes. Revertant BACs, namely ΔIR_L-*meq*-D-rev, ΔIR_L-*meq*-O-rev and ΔIR_L-*meq*-R-rev, were generated by initially replacing the *meq*-D, *meq*-O and *meq*-R genes, respectively, with an *amp^r*, which was then replaced by the parental *meq*-W gene. The recombinant (mutant as well as revertant) BAC clones were tested for correct insertion of the recoded genes and complete deletion of the parental *meq* gene, or vice versa, by Sanger sequencing. The integrity of the BACs was confirmed by RFLP (**Figure 10**).

Table 10. Comparison of coding sequences of parental (*meq*-W), CPB deoptimized (*meq*-D), optimized (*meq*-O) and randomized (*meq*-R) versions of MDV *meq* gene.

All sequences consist of the same codons, but the order of the codons in each of the sequences is different. Reshuffling of synonymous codons lead to the introduction of hundreds of silent mutations. Exactly the same [a]nucleotides and [b]codons occupy the same position in the ORFs.

| ORF	ORF	Nucleotide identity (%)[a]	Codon identity (%)[b]	No. of silent mutations
meq-W	*meq*-D	78.24	41.47	221
meq-W	meq-O	79.90	44.71	204
meq-W	*meq*-R	79.12	41.47	212
meq-D	meq-O	73.14	37.06	273
meq-D	*meq*-R	77.06	27.35	233
meq-O	*meq*-R	79.12	42.65	212
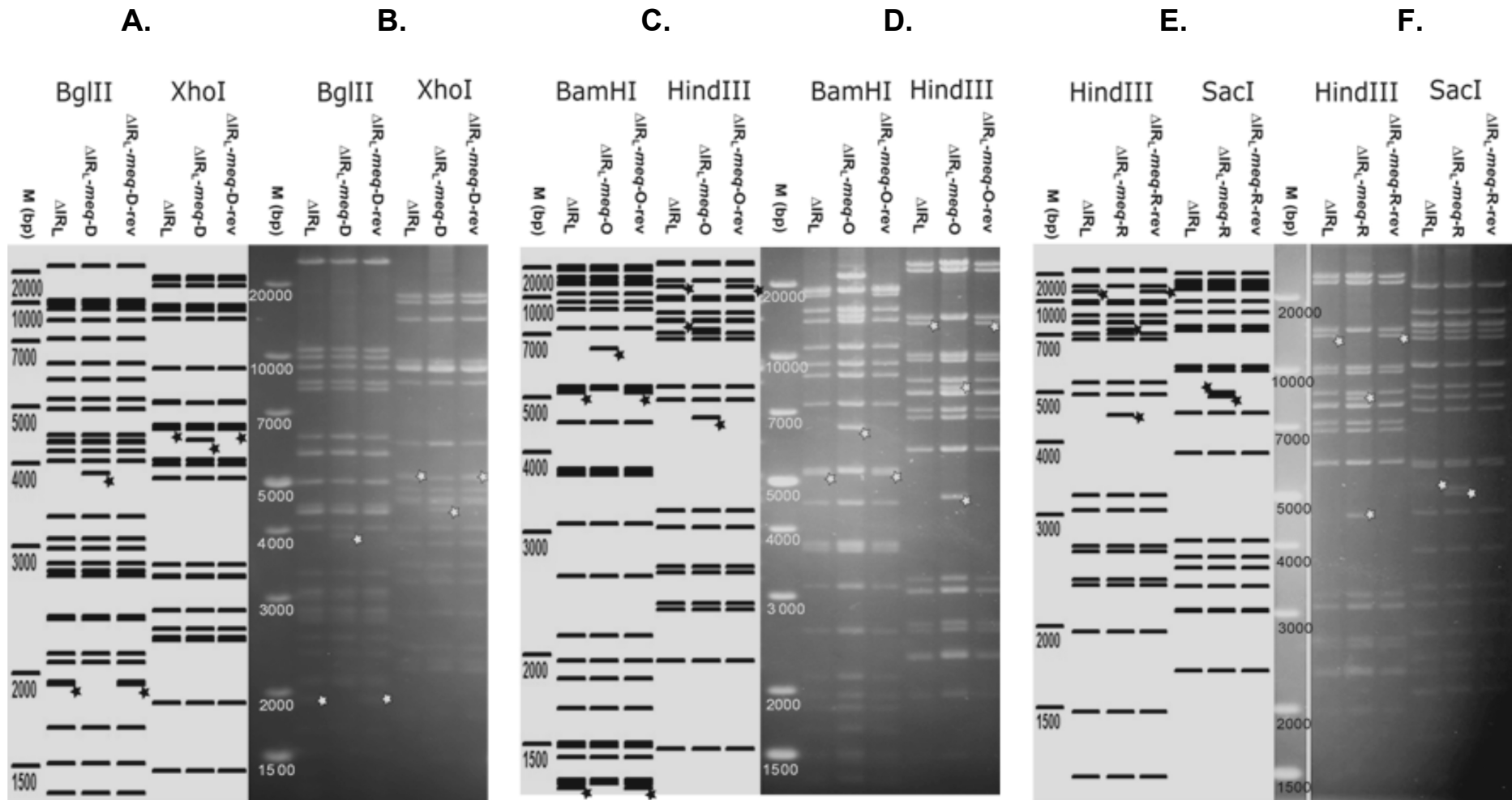
Figure 10. RFLP analysis of recombinant MDV BACs.

RFLP patterns of the recombinant BACs, **(B)** ΔIR$_L$-*meq*-D and ΔIR$_L$-*meq*-D-rev; **(D)** ΔIR$_L$-*meq*-O and ΔIR$_L$-*meq*-O-rev; and **(F)** ΔIR$_L$-*meq*-R and vΔIR$_L$-*meq*-R-rev; with the respectively labelled restriction enzymes were compared to the RFLP patterns of the parental BAC vΔIR$_L$. **(A), (C)** and **(E)** represent the pDRAW32 (116) predictions of the respective RFLP patterns.

7.4. *In vitro* characterization of reconstituted viruses

The recombinant BAC DNA was transfected in CEC to reconstitute the viruses containing either the parental *meq*-W gene, v∆IR$_L$, or one of its recoded versions, v∆IR$_L$-*meq*-D, v∆IR$_L$-*meq*-O and v∆IR$_L$-*meq*-R (§ 6.2.3.3.). To assess the ability of the viruses to grow *in vitro* and their cell-to-cell spread, plaque sizes of all the mutant viruses were compared to those of the parental virus. Indeed, the plaque sizes of the mutant viruses had no significant difference compared to those of the parental virus (**Figure 11A**, One-way ANOVA, n = 153, p = 0.0612). The revertant viruses v∆IR$_L$-*meq*-D-rev, v∆IR$_L$-*meq*-O-rev and v∆IR$_L$-*meq*-R-rev, were similarly reconstituted and their respective plaque sizes were compared to those of the parental virus v∆IR$_L$. No significant difference between the respective plaque sizes of the revertant viruses and those of the parental virus (**Figure 11B,** One-way ANOVA, n = 153, p = 0.3648) confirmed the absence of unexpected mutations elsewhere in the viral genome.
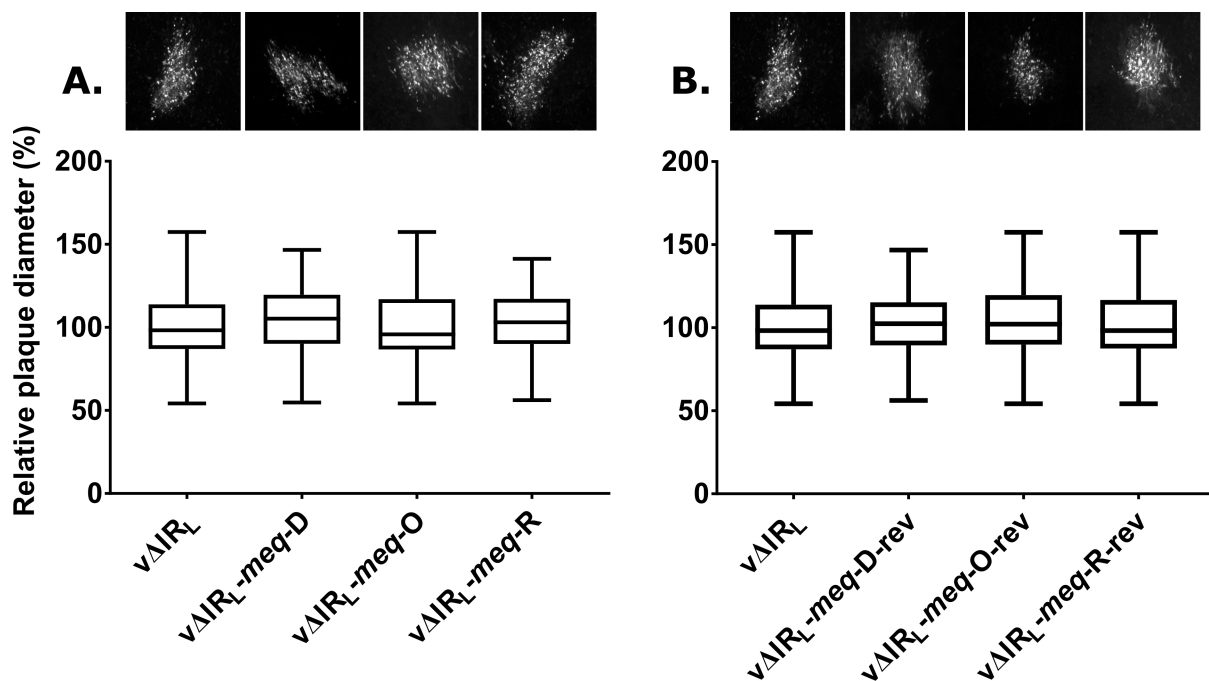
Figure 11. Plaque size assay.

Relative diameters of the plaques induced by the parental virus, v∆IR$_L$, were compared to those induced by **(A)** mutant viruses, v∆IR$_L$-*meq*-D, v∆IR$_L$-*meq*-O and v∆IR$_L$-*meq*-R (n = 153, One-way ANOVA, p = 0.0612) and, **(B)** revertant viruses, v∆IR$_L$-*meq*-D-rev, v∆IR$_L$-*meq*-O-rev and v∆IR$_L$-*meq*-R-rev (n = 153, One-way ANOVA, p = 0.3648). Relative plaque diameters were calculated w.r.t. the median plaque diameter of v∆IR$_L$ and are represented as box plots. Images above show a representative plaque induced by the respective virus.
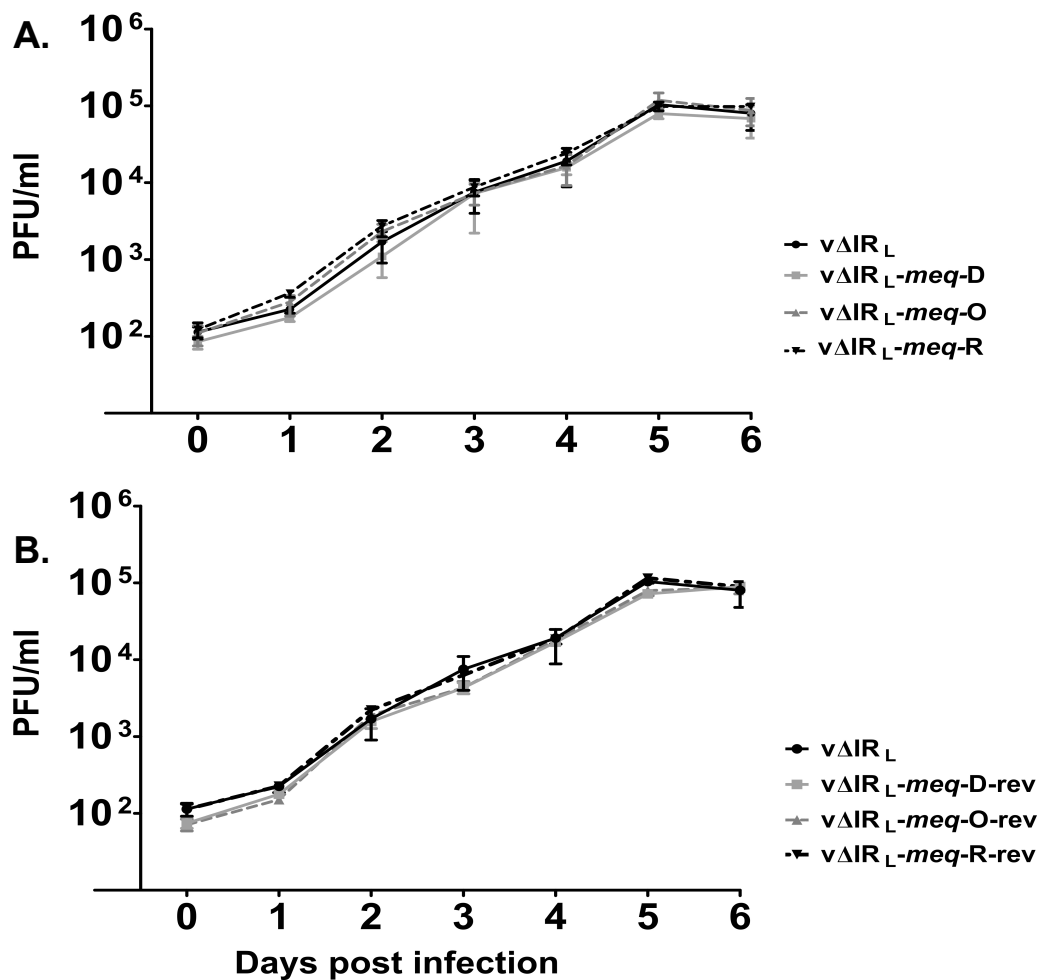
**Figure 12. Multi-step growth kinetics of recoded viruses.**

(A) Comparison of growth curves of the mutant viruses, v∆IR$_L$-*meq*-D, v∆IR$_L$-*meq*-O and v∆IR$_L$-*meq*-R with that of the parental virus, v∆IR$_L$ (n = 4, Kruskal-Wallis, p = 0.9123), **(B)** Comparison of growth curves of the revertant viruses, v∆IR$_L$-*meq*-D-rev, v∆IR$_L$-*meq*-O-rev and v∆IR$_L$-*meq*-R-rev with that of the parental virus, v∆IR$_L$ (n = 4, Kruskal-Wallis, p = 0.9277).

To confirm the results of the plaque assays, multi-step growth kinetics were performed simultaneously on all the mutant viruses as well as the parental virus. Indeed, the mutant viruses did not show any significant difference in the growth kinetics as compared to the parental virus (**Figure 12A,** Kruskal-Wallis Test, n = 4, p = 0.9123). Similarly, the growth curves of the revertant viruses were compared to that of the parental virus. No significant differences between the respective growth curves of the revertant viruses and the parental virus (**Figure 12B,** Kruskal-Wallis Test, n = 4, p = 0.9277) further confirmed the absence of undesired mutations elsewhere in the viral genome.

7.5. *In vitro* gene expression assays

7.5.1. Immunoblot analysis

A eukaryotic dual expression vector, pVitro2-hygro-MCS (Invivogen) (§ 6.1.3.1.), was used to assess the ability of the recoded *meq* genes to produce proteins. The vector consists of two composite promoters, hFerH and hFerL, which control the expression of the heavy and light subunits of the ubiquitously expressed ferritin gene in humans. The *meq* gene (parental or recoded) was cloned under the control of the hFerH promoter.
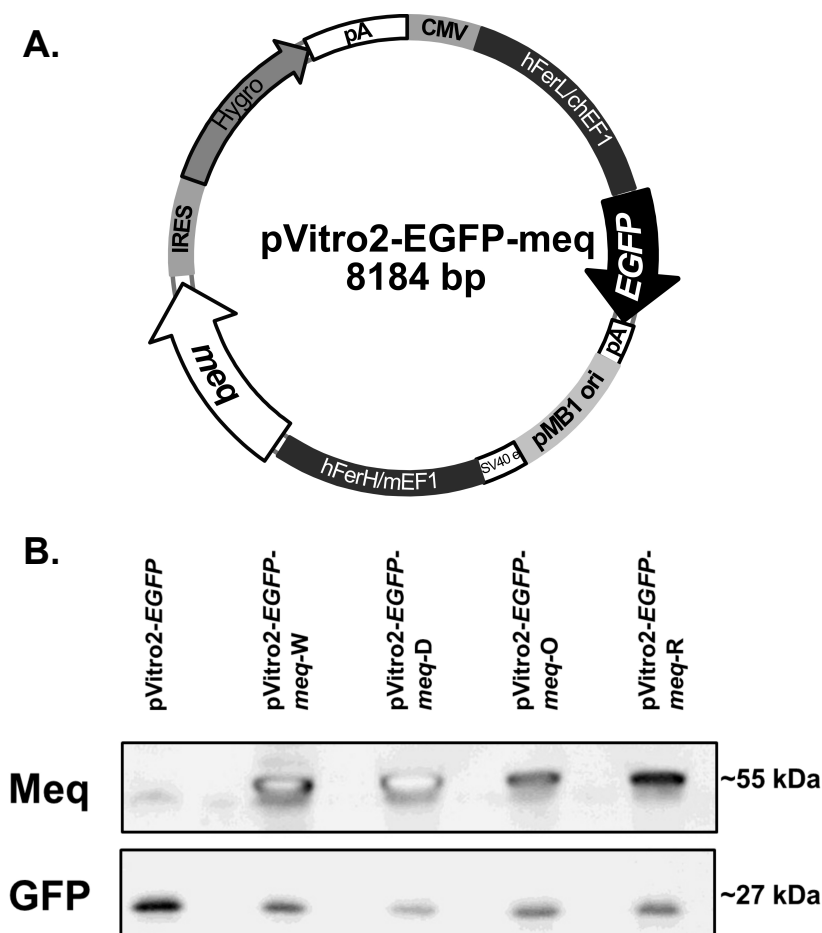
Figure 13. Immunoblot analysis.

(A) Schematic representation of the eukaryotic dual expression vector used to assess the ability of the recoded versions of *meq* to produce protein. CMV: major immediate early enhancer of human cytomegalovirus; *EGFP* and *meq* (recoded or parental) were respectively cloned under hFerL and hFerH: composite promoters of the light (L) and the heavy (H) subunits of the ubiquitously expressed human ferritin gene; to eliminate the iron regulation of the two promoters the 5'-UTRs of the promoters are respectively replaced with chimpanzee (chEF-1) and mouse (mEF-1) elongation factor 1 genes. pA: polyadenylation signal; pMB1 ori: a minimal *E. coli* origin of replication; *Hygro*: confers resistance to hygromycin B in both *E. coli* and mammalian cells; IRES: internal ribosome entry site of Foot and Mouth Disease virus. **(B)** Immunoblot analysis. The membranes stained with a polyclonal rabbit anti-Meq antibody (120) (upper panel) and with mouse anti-GFP monoclonal IgG$_{2a}$ (lower panel, control).

A green fluorescent protein gene (*EGFP*), which served as transfection control, was cloned under the control of the hFerL promoter (**Figure 13A**). All the four plasmids were transfected into 293T cells (§ 6.2.3.2.). An 'empty' vector with only *EGFP* was used for mock transfection. The cells were lysed 24 h post transfection and the lysates were used for immunoblotting as described in § 6.2.4.1. All of the recoded version of the *meq* gene could produce the Meq protein at amounts detectable by immunoblotting (**Figure 13B**).

7.5.2. Flow cytometry

After confirming that the recoded genes are, in fact, able to produce protein, the next step was to determine if CPBD influences the levels of protein produced and, if it does, to what extent. For this, a flow cytometry based assay for protein production was designed. 293T cells were transfected with the eukaryotic dual expression vector, pVitro2-hygro-MCS, in which one of the recoded or the parental *meq* genes were cloned under the control of the hFerH[2] promoter. An *EGFP* gene was fused to the *meq* gene. A blue fluorescent protein gene, *mTagBFP*, was cloned under the influence of hFerL promoter (**Figure 14C**). The cells were harvested 24 h post transfection and were analysed by flow cytometry for both BFP as well as GFP signals as described in § 6.2.4.2. (**Figure 14A**).

As *EGFP* was fused to the C-terminal end of *meq* (either recoded or parental, **Figure 14C**), it was simultaneously expressed with *meq* providing a measurable marker for Meq protein production. Therefore, the expression levels of all the four *meq* variants were measured as the mean fluorescence intensity (MFI) of the EGFP fused to them by flow cytometry. The MFI values for EGFP were then normalized for transfection efficiency with the help of the MFI values of BFP for the respective samples. The values, thus obtained, were the absolute expression levels of the respective *meq* variant. The expression levels of the recoded genes, *meq*-D, *meq*-O and *meq*-R, were then expressed as percent relative expression w.r.t. the parental gene, *meq*-W. These relative expression levels of *meq*-D (74.38 %, p = 0.0025), *meq*-O (26.94 %, p < 0.0001) and *meq*-R (38.05 %, p < 0.0001) were significantly lower than the expression levels of *meq*-W (**Figure 14B**, One-way ANOVA, n = 5).

7.6. *In vivo* characterization

The next step was to evaluate the effect of CPBD of *meq* on MD progression and disease outcome *in vivo*. One-day old chickens were infected with 5,000 PFU of the parental vvMDV strain, vΔIR$_L$; one of the mutant viruses, vΔIR$_L$-*meq*-D, vΔIR$_L$-*meq*-O or vΔIR$_L$-*meq*-R; or one of their respective revertants, vΔIR$_L$-*meq*-D-rev, vΔIR$_L$-*meq*-O-rev

---

[2] For promoter details see § 7.5.1. or the legend for **Figure 14C.**
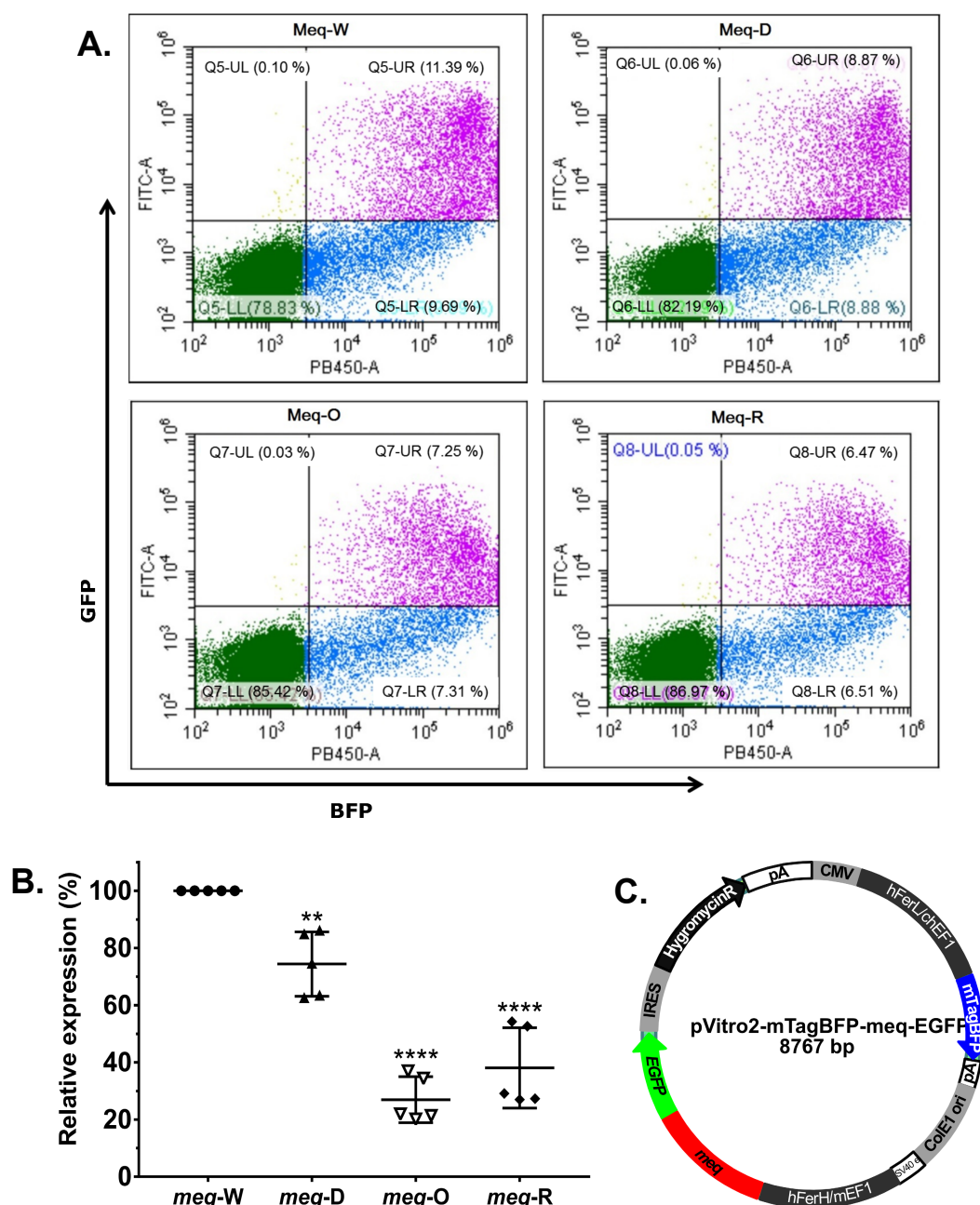
Figure 14. Flow cytometry based in vitro gene expression assay.

(A) Flow cytometric analysis of cells transfected with the denoted *meq* variant. **(B)** Comparison of relative expression of *meq*-D, *meq*-O and *meq*-R with that of *meq*-W (One-way ANOVA, n = 5, **p = 0.0025, ****p < 0.0001). **(C)** Schematic representation of the eukaryotic dual expression vector constructed for the assay. Different versions of the *meq* gene were fused with an *EGFP* gene at the 3'-end CMV: major immediate early enhancer of human cytomegalovirus; hFerL and hFerH: composite promoters and the light (L) and the heavy (H) subunits of the ubiquitously expressed human ferritin gene; to eliminate the iron regulation of the two promoters the 5'-UTRs of the promoters are respectively replaced with chimpanzee (chEF-1) and mouse (mEF-1) elongation factor 1 genes. pA: polyadenylation signal; pMB1 ori: a minimal *E. coli* origin of replication; Hygro: confers resistance to hygromycin B in both *E. coli* and mammalian cells; IRES: internal ribosome entry site of Foot and Mouth Disease virus.

or vΔIR$_L$-*meq*-R-rev. The infection experiment was performed in a blinded fashion. Naïve chickens (contact birds) were housed with the infected chickens to assess bird to bird viral transmission via the natural route of infection. The experiment was carried out for a duration of 91 days (§ 6.2.5.1.).

The ability of the viruses to replicate *in vivo* was assessed by determining the number of MDV genome copies present per 1×10^6 cells in the peripheral blood of the infected birds at weekly intervals (7, 14, 21 and 28 dpi) by qPCR (§ 6.2.5.2.). Blood samples were collected from the same 8 birds from each group for each time point. The number of MDV genome copies in the blood samples of the birds infected with either one of the mutant viruses (**Figure 15A**, p = 0.6576) or one of the revertant viruses (**Figure 15B**, p = 0.9792) were not significantly different compared to that of the parental virus, vΔIR$_L$ (**Figure 15A** and **B**, Kruskal Wallis test, df = 3).

The spread of the virus via the natural route was assessed by similarly determining the MDV genome copy number in peripheral blood samples of the contact birds. Blood samples were collected from eight birds from each contact group housed with different infected groups at weekly intervals between 21 and 42 dpi (on days 21, 28, 35 and 42 post infection). For logistical reasons, birds infected with vΔIR$_L$ and vΔIR$_L$-*meq*-O-rev were housed together in one room and those infected with vΔIR$_L$-*meq*-D-rev and vΔIR$_L$-*meq*-R-rev were housed together in another room. Both rooms had one group of contact animals each. Each of the groups infected with vΔIR$_L$-*meq*-D, vΔIR$_L$-*meq*-O and vΔIR$_L$-*meq*-R was housed in a separate room with one contact group each. The number of MDV genome copies present in the blood samples of the contact birds from all the groups were not significantly different in comparison to that of the contact birds from the vΔIR$_L$/vΔIR$_L$-*meq*-D-rev group at the respective time points (**Figure 15C**, p = 0.6319, Kruskal Wallis test, df = 4).

Birds were monitored daily for signs of MD. All birds infected with the parental virus, vΔIR$_L$ (n = 9), developed MD well before the end of the experiment with the last bird developing the disease 84 dpi. The time required for 50 % of the vΔIR$_L$ infected chickens to develop MD (MD$_{50}$) was 59 days (**Figure 16A** and **B**). At the end of the experiment (91 dpi), the revertant viruses, vΔIR$_L$-*meq*-D-rev, vΔIR$_L$-*meq*-O-rev and vΔIR$_L$-*meq*-R-rev; caused MD in 87.5 % (n = 24), 95.65 % (n = 23) and 96.00 % (n = 25) of the infected birds, respectively (**Figure 16D**). The time required for the revertant viruses to attain MD$_{50}$ was 48 dpi for vΔIR$_L$-*meq*-D-rev, 56 dpi for vΔIR$_L$-*meq*-O-rev, and 50 dpi for vΔIR$_L$-*meq*-R-rev (**Figure 16C**), while that for the mutant viruses was 67 dpi for vΔIR$_L$-*meq*-D, 71 dpi for vΔIR$_L$-*meq*-O, and 72 dpi vΔIR$_L$-*meq*-R (**Figure 16A**). At the end of the experiment, the three mutant viruses, vΔIR$_L$-*meq*-D, vΔIR$_L$-*meq*-O, and vΔIR$_L$-*meq*-R;
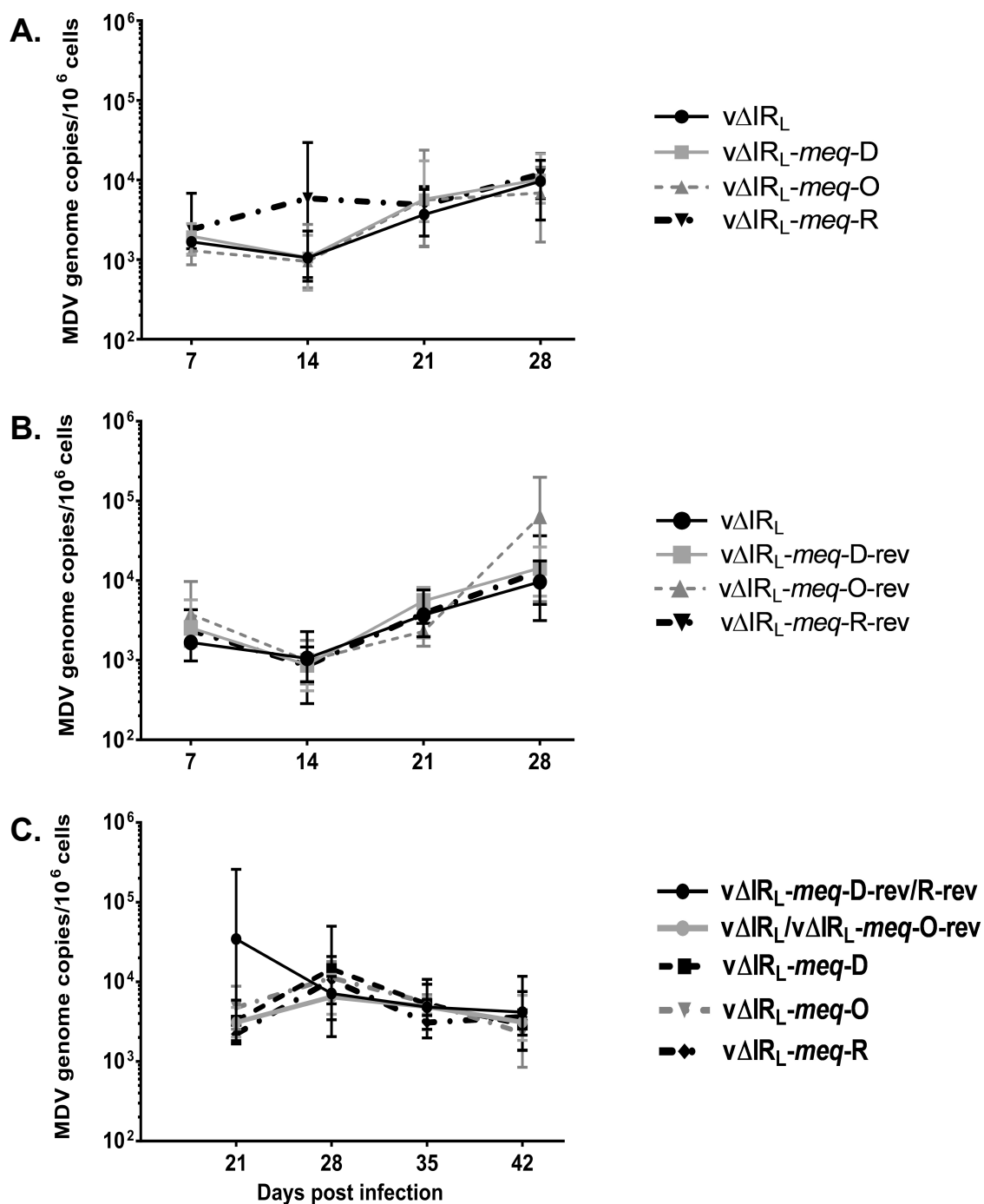
Figure 15. MDV replication *in vivo*.

Means of MDV genome copy numbers quantified by qPCR performed on DNA isolated from blood samples of 8 birds from each of the groups infected with **(A)** parental and mutant viruses, **(B)** parental and revertant viruses or their respective uninfected contact groups **(C)** which acquired the infection via the natural route of infection.

caused MD in 87.5 % (n = 24), 79.17 % (n = 24), and 78.26 % (n = 23) of the infected birds, respectively (**Figure 16B**).

The mutant viruses, thus, took on an average 70 days to attain MD_{50} in the inoculated birds which was ~17 days longer than the average time of 53.25 days taken by the revertant and parental viruses to cause MD_{50}. Moreover, the average MD incidence of 94.79 % in the birds infected with parental and revertant viruses was on an average 13.15 % higher than the 87.5 % (-7.29 %)[3], 79.17 % (-13.15 %) and 78.26 % (-16.53 %). MD incidence in the birds infected with vΔIR$_L$-*meq*-D, vΔIR$_L$-*meq*-O, and vΔIR$_L$-*meq*-R, respectively. However, the differences were not statistically significant.

The birds were euthanized as soon as the first signs of MD were evident and necropsies were carried out to examine the internal organs for lymphoma formation. Representative images of gross macroscopic tumorous lesions observed in affected birds during this experiment are shown in **Figure 1C** and **D**.

All birds infected with the parental virus, vΔIR$_L$, developed MD, out of these 88.89 % also had macroscopic tumours on their visceral organs. Compared to that only 61.90 % (54.17 %)[4] of the MD positive birds in the vΔIR$_L$-*meq*-D group had tumours. The tumour incidence in MD positive birds of the vΔIR$_L$-*meq*-O and vΔIR$_L$-*meq*-R infected groups was 84.20 % (66.67 %) and 61.10 % (47.83 %), respectively, (**Figure 16B**), while the MD positive birds infected with the revertant viruses had a tumour incidence of: 71.40 % (62.50 %) for vΔIR$_L$-*meq*-D-rev, 68.10 % (65.20 %) for vΔIR$_L$-*meq*-O-rev, and 83.30 % (80.00 %) for vΔIR$_L$-*meq*-R-rev (**Figure 16D**).

The MD and tumour incidences in the birds infected with both mutant and revertant viruses were not significantly different than the birds infected with parental vΔIR$_L$ (**Figure 16B**, Fisher's exact test, two-sided, p > 0.02, and **Figure 16D**, Fisher's exact test, two-sided, p > 0.02).

Among the contact birds, MD incidence was the highest for vΔIR$_L$-*meq*-O with 54.55 % showing MD signs while 49.99 % (27.27 %) of the MD positive birds also had tumours. It is worth mentioning again that most of the contact birds developed MD only in the last few days of the experiment and none of them reached MD_{50} by the end of the experiment, except for vΔIR$_L$-*meq*-O which reached MD_{50} on the final day (Day 91) of the experiment (**Figure 16E**). The contact group housed with vΔIR$_L$-*meq*-R infected birds

---

[3] Figures in parentheses indicate the difference between the average MD incidence of the revertant and parental viruses and that of the respective mutant viruses.

[4] Figures in parentheses indicate the percent tumour incidence w.r.t. the total number of birds infected with the indicated virus.
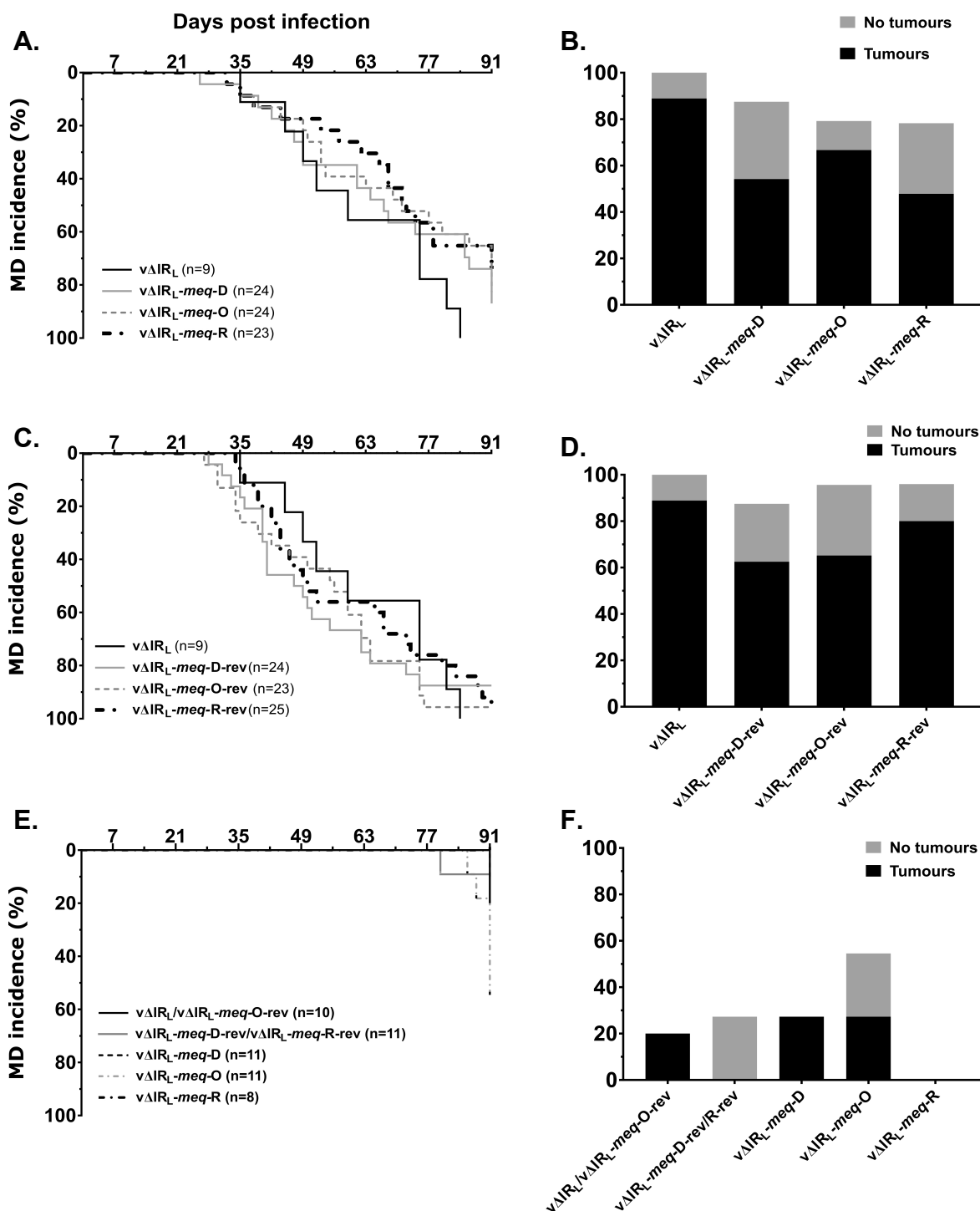
Figure 16. *In vivo* characterization of viruses with CPB altered *meq*.

Survival analysis of the chicken **(A)** infected with mutant viruses, **(C)** infected with revertant viruses, as compared to those infected with vΔIRL, and **(E)** contacts birds that were housed with the labelled infected groups.

Percent MD incidence at the end of the experiment (91 dpi) in the groups infected with the indicated **(B)** mutant, **(D)** revertant viruses, as compared to the MD incidence in the vΔIRL infected group; and **(F)** contact birds. Darker portion of the bars represent the percentage of MD positive birds in which macroscopic tumours were observed on one or more visceral organs during necropsy.

had the lowest MD as well as tumour incidence with none of the birds showing symptoms at the end of the experiment. The MD incidence in the contact groups of vΔIR$_L$-*meq*-D and vΔIR$_L$-*meq*-D-rev/R-rev was 27.27 % each and 100 % (27.27 %)[5] and 0 % of the MD positive birds, respectively, had tumours. The MD incidence for the contact group of vΔIR$_L$/vΔIR$_L$-*meq*-O-rev was 20.00 % with a 100 % (20.00 %) tumour incidence. The MD and tumour incidences in all the contact groups were not significantly different than those of the contact group housed with vΔIR$_L$/vΔIR$_L$-*meq*-O-rev (**Figure 16F**, Fisher's exact test, two-sided, p > 0.02).

---

[5] Figures in parentheses indicate the percent tumour incidence w.r.t. the total number of contact birds housed with the respective infected groups.

8. Discussion

Virus attenuation through CPBD has the following major advantages: (i) It preserves the immunogenicity of the encoded proteins as their amino acid composition is identical with the parental protein. (ii) The sheer number of introduced mutations makes reversion to parental sequence virtually impossible. (iii) The method is systematic and hence seemingly applicable to several viruses. (iv) The method is rapid, therefore, development of vaccines even against novel or emerging viral diseases is potentially possible in a relatively short time frame (7). Despite this apparent applicability of CPBD to many viruses, the method had only been applied to attenuate small RNA viruses so far with varying degrees of success (**Table 2**). Whether CPBD is suitable as an attenuation strategy for DNA viruses, including large DNA viruses like herpesviruses, remained to be verified. Therefore, the present study dealt with the application of CPBD to MDV, a highly contagious herpesvirus that causes T-cell lymphomas in chickens. Being a pathogen of chicken, MDV also provided a unique small animal model to study the effects of CPBD on a large DNA virus *in vivo.* Moreover, since MDV causes disease in 100 % of the infected birds, it is an ideal model to test vaccines.

The history of MD vaccination makes the virus an interesting candidate to study newer approaches to vaccine development. Since the introduction of the first vaccine in the late 1960s, MDV has continuously evolved towards increased virulence (from mMDV to vMDV to vvMDV to vv+ MDV) and has overcome two vaccine regimens in little over two decades (**Figure 6**) (87,91–95,113). To tackle the problem of vv+ MDV, the vaccine strain CVI988/Rispens was introduced in the 1990s. It remains the most effective vaccine strain against MD to date and all the vaccine strains that were developed later have been unable to surpass its efficacy (35,87,99,100). However, if MDV continues its course of vaccine evasion and continuous evolution towards increased virulence, a need for a newer and more potent vaccine would arise on short notice (100). Consequently, the search for better vaccines has been in the forefront of MDV research and exploring newer avenues towards vaccine development like CPBD might have the potential to provide us with a suitable candidate.

CPBD of the entire MDV genome is impossible at this point of time due to its huge size (~180 kb) and therefore must be applied to individual genes. At the same time, owing to the same large genome, MDV has a wide range of genes that could be targeted for CPBD. For the present study, *meq* – the major oncogene of MDV was selected. Despite being non-essential for viral growth *in vitro*, *meq* is a crucial immunogen and is also one of the few genes that are expressed during latency (108,110,112). Interestingly, CVI988/Rispens, the gold standard vaccine strain, expresses two functional

copies of *meq* and yet does not cause tumours (114), suggesting that attenuation of *meq* containing MDV strains can render them non-oncogenic.

With all this in mind and to assess if CPBD of the *meq* gene could be a viable strategy for MDV attenuation, the following three goals were set forth:

1. Generation and *in vitro* characterization of recombinant vvMDV strain RB-1B with CPB deoptimized *meq* gene,
2. To study the effects of CPBD on *meq* expression *in vitro,*
3. To study the effects of CPB deoptimized *meq* gene on MD progression and disease outcome *in vivo.*

It was assumed that MDV must have adapted to the CPB of its natural host, chicken. However, chicken CPB was still unknown; making its calculation the first task of this study. CPB is the arithmetic mean of the CPSs of all codon pairs in a given ORF, where CPS is defined as the natural logarithm of the ratio of the observed to the expected occurrences of a codon pair (**Figure 7**) (7). Therefore, an algorithm was developed to compute the CPS of each of the possible 3,721 codon pair combinations (61 × 61, excluding stop codons) based on 15,762 predicted, protein-encoding chicken ORFs (*Gallus gallus*, breed Red jungle fowl, line UCD001) (17). The calculations were performed exactly as described (7) and the CPSs obtained were used to determine the CPB of the said 15,762 chicken ORFs as well as 131 MDV ORFs (vv strain RB-1B).

CPB has been shown to exist in all the species studied so far and is species specific. In other words, CPB varies among different species (13,17–19). Phylogenetically related species have a highly similar CPB suggesting that CPB has co-evolved with the respective species (17). Consistent with these findings, the calculated chicken CPB score with an average of 0.0755 was highly similar to the average human CPB score of 0.07 (7). Despite this overall similarity, both the species differ in their preferences for codon pairs in their coding sequences to a certain extent. As seen in **Table 8** which compares the human and chicken CPSs of the 16 possible codon pair combinations that encode the amino acid pair alanine-alanine, the most overrepresented as well as the most underrepresented codon pairs were identical in both species while their preference for the remaining 14 codon pairs was found to be different.

The 131 MDV ORFs had an overall negative CPB score with an average of -0.0646 (**Figure 8**). Such differences, where a virus has a slightly lower CPB score than its host, have been reported in most studies listed in **Table 2** (7,32–34). Comparison of several human viruses suggests that the following two factors affect viral CPB: (i) Genome size: Viruses with smaller genomes tend to have highly similar CPB to their host while viruses with larger genomes have lower, usually negative, CPB scores.

(ii) Genetic material: DNA viruses have lower CPB scores compared to their RNA counterparts (17). This seems to hold true for the MDV and chicken virus-host pair as well. Thus, the lower CPB score of MDV can be attributed to the large size of its genome (~180 kb) and the fact that it is made up of dsDNA.

Using the calculated CPSs for the chicken ORFeome (17), the sequence of the MDV *meq* gene was recoded to obtain its CPB deoptimized (*meq*-D) and optimized (*meq*-O) versions, in which the number of underrepresented codon pairs was respectively maximized and minimized by reshuffling the constituent codons. A third recoded version of the gene, *meq*-R, had its codons reshuffled in a randomized fashion such that it had a similar overall CPB score compared to the parental gene (*meq*-W) (§ 6.2.1.3.). **Figure 8** demonstrates the distribution of the CPB scores of the recoded versions of the *meq* gene in comparison to *meq*-W.

To study the effects of the CPB altered *meq* genes on the replication and growth properties of MDV *in vitro*, mutant BACs of vvMDV strain RB-1B were generated by *en passant* mutagenesis (117) by replacing the wild-type *meq* gene (*meq*-W) with its CPB deoptimized (*meq*-D), optimized (*meq*-O) or randomized (*meq*-R) version. The parental BAC, ΔIR_L, had only one copy of the *meq* gene which ensured its complete replacement. Moreover, the reconstituted virus, $v\Delta IR_L$, has been shown to restore the deleted inverted repeat region within two passages after reconstitution (77). Therefore, the reconstituted viruses containing the recoded *meq* genes had two copies of the respective *meq* variant after two passages in cell culture just as the parental virus.

The ability of the reconstituted mutant viruses, $v\Delta IR_L$-*meq*-D, $v\Delta IR_L$-*meq*-O and $v\Delta IR_L$-*meq*-R, to replicate and infect neighbouring cells was assessed by plaque size assay. The mutant viruses formed plaques with similar sizes to those formed by the parental virus (**Figure 11A**). This was consistent with the non-essential nature of *meq* for MDV growth *in vitro* (114) and was further confirmed by the similar multi-step growth kinetics of the mutant and the parental viruses (**Figure 12A**). Additionally, to rule out unwarranted changes in the viral genome during mutagenesis, revertant BACs were produced from each of the mutants by restoring the parental *meq*-W gene. The reconstituted revertant viruses, $v\Delta IR_L$-*meq*-D-rev, $v\Delta IR_L$-*meq*-O-rev and $v\Delta IR_L$-*meq*-R-rev, indeed, were able to replicate *in vitro* similarly to the parental virus, $v\Delta IR_L$, as evident from their comparable plaque sizes (**Figure 11B**) and multi-step growth kinetics (**Figure 12B**), thus confirming the absence of unintended mutations elsewhere in the viral genome.

In vitro virus characterization proved that CPB altered *meq* did not have any effect on MDV replication. Although this was expected given the non-essential nature of *meq* for viral growth in cell culture, it did not answer the following two questions: (i) Are the CPB altered *meq* genes able to produce a protein? (ii) Does CPBD affect protein production and, if it does, to what extent? The ability of *meq*-D, *meq*-O and *meq*-R to produce a protein product was successfully tested by immunoblotting. Each of the three versions could produce a protein that was identical in size with the parental protein, Meq-W, in transiently transfected cells (**Figure 13**). But, to assess if and to what extent does CPB alteration affect the levels of protein production, a more sensitive assay was required. Therefore, a flow cytometry based assay using two fluorescent protein reporters was developed (§ 7.5.2.). An *EGFP* gene was fused with *meq* at the C-terminus (**Figure 14A**). Therefore, its fluorescence provided for a measurable marker for Meq protein production while the BFP signal served as a control for transfection efficiency and was used to normalize the EGFP signal. Using the normalized values, relative expression levels for each *meq* variant were determined as a percentage of *meq*-W expression. The relative expression levels for *meq*-D, the CPB deoptimized version of *meq*, were found to be significantly lower than the parental *meq*-W gene (**Figure 14B**). Thus, CPBD of the *meq* gene resulted in reduced protein production. This was consistent with the results of previous studies involving several viruses, such as poliovirus, influenza A virus, HRSV, VSV and PRRSV (7,18,29,31–33).

The other two versions, namely CPB optimized (*meq*-O) and randomized (*meq*-R), surprisingly also had lower expression levels than *meq*-W (**Figure 14B**). Coleman *et al.* have reported that CPB optimization of poliovirus results in higher levels of protein production compared to wild-type (7). The reduced gene expression of *meq*-O in the present study was inconsistent with these findings, however consistent with the results of a study on the L1 gene of VSV, where CPB optimization also caused a reduction in protein production (33).

Minimizing the number of underrepresented codons, i.e. CPB optimization, logically should not reduce protein production. However, it has been shown in *E. coli* that overrepresented codon pairs are translated slowly (137). The authors argue that since overrepresented codon pairs occur at structural boundaries of proteins, their slower translation might facilitate optimum protein folding. Nevertheless, this slower translation rate can lead to a *decrease* in protein production if such overrepresented codon pairs occur early in an ORF, possibly because of ribosome stalling and inhibition of attachment of new ribosomes (21). Swift translation initiation, thus, is necessary for efficient protein production, however, it is not sufficient and a high elongation rate plays an equally important role. Low expression, on the other hand, can result independently from either

a low initiation rate or a low elongation rate (138,139). Since *meq*-O and *meq*-R had respectively higher and similar CPB scores compared to *meq*-W, it is unlikely that they had lower elongation rates. A slower translation initiation rate due to the presence of overrepresented codon pairs at the beginning of these ORFs, however, seems to be a plausible explanation for their lower expression levels. The CPB scores of the first 20 codon pairs of *meq*-O (0.227) and *meq*-R (- 0.083) were, in fact, higher than that of *meq*-W (- 0.146). In other words, both *meq*-O and *meq*-R had overrepresented codon pairs at the 5'-end, hence, supporting the slow translation initiation theory behind their lower expression levels.

On the other hand, the CPB score of the first 20 codon pairs of *meq*-D (-0.356) was lower than that of *meq*-W (-0.146). Thus, the lower expression levels of *meq*-D should be because of its lower elongation rate owing to the maximized number of underrepresented codon pairs throughout its sequence, i.e. CPBD. It must be noted, however, that these speculations are based on studies performed in bacterial systems and further studies in eukaryotic systems are necessary for a better understanding of the underlying mechanism.

Clearly, the exact mechanism behind how CPBD leads to reduced protein production remains poorly understood. However, the following hypotheses have been put forth: (i) Certain underrepresented codon pairs are associated with missense, non-sense and frameshift errors in translation (21–23). (ii) Juxtaposition of aminoacyl-tRNAs on the A- and P-sites of ribosomes determine the peptide elongation rate (24), and hence might play an important role in shaping CPB, i.e. underrepresented codon pairs might bind with aminoacyl-tRNAs whose interactions slow down the rate of translation leading to reduced protein production (24). To summarize, the reduction in protein production by CPBD may be attributed to the following: immature chain termination, improper protein folding, and/or a low elongation rate.

The next question was whether CPBD of the *meq* gene affected MD progression and disease outcome *in vivo*. To answer this, one-day old chickens were infected with the parental virus or one of the mutants or revertants and were observed for MD signs over a period of 91 days. Quantitation of MDV genome copies in peripheral blood of the infected birds at weekly intervals showed that vΔIR$_L$-*meq*-D could replicate comparably to the parental virus vΔIR$_L$ *in vivo* (**Figure 15A**). However, the development of MD in birds infected with vΔIR$_L$-*meq*-D was slightly slower (time to MD$_{50}$ = 67 dpi) compared to vΔIR$_L$ (56 dpi) as well as its revertant, vΔIR$_L$-*meq*-D-rev (48 dpi) (**Figure 16A** and **C**). Moreover, while 100 % of the birds infected with vΔIR$_L$ developed MD, the MD incidence in vΔIR$_L$-*meq*-D infected birds was only 87.50 %. Similarly, the incidence of macroscopic

tumours was lower in vΔIR$_L$-*meq*-D infected birds with 54.17 % developing tumours compared to the 88.89 % of vΔIR$_L$ (**Figure 16B**). Thus, CPBD of the *meq* gene did not completely attenuate the oncogenicity of the virus, however, did cause a slight reduction.

Interestingly, birds infected with vΔIR$_L$-*meq*-O and vΔIR$_L$-*meq*-R also took longer to develop MD$_{50}$ at 71 dpi and 72 dpi (**Figure 16A**), respectively, while their initial replication was comparable to vΔIR$_L$ (**Figure 15A**). The overall MD incidence was 79.17 % for vΔIR$_L$-*meq*-O and 78.26 % for vΔIR$_L$-*meq*-R. The tumour incidence with 66.67 % for vΔIR$_L$-*meq*-O and 47.83 % for vΔIR$_L$-*meq*-R was also lower compared to the 88.89 % of the parental virus, vΔIR$_L$ (**Figure 16B**). These results also correspond with the reduced *in vitro* expression levels of the respective *meq* variants, *meq*-O and *meq*-R (**Figure 14B**). Interestingly, similar outcome has been reported in VSV, where CPB optimization of the L1 gene also lead to viral attenuation while the recombinant virus containing the CPB deoptimized version of the gene could not be reconstituted (33).

Initially, it was suggested that virus attenuation by CPBD is a direct consequence of reduced protein production by the recoded genes by a hitherto elusive mechanism (7,34). This theory, however, has recently been contested after the discovery that CPBD inadvertently increases the number of CpG and TpA dinucleotides in the ORFs. The alternative theory proposes that this inadvertent increase in dinucleotide frequencies makes the sequences susceptible to recognition by some largely uncharacterized innate immune mechanisms (17,25,26). CPBD of *meq* also caused such an inadvertent increase in the dinucleotide frequencies with a CpG and TpA abundance[6] of 1.12 and 1.26, compared to 0.71 and 0.77 of *meq*-W, respectively. Unsurprisingly, the CpG (0.44) and TpA (0.40) abundance was lower in *meq*-O while *meq*-R had similar CpG (0.73) and TpA (0.77) abundance as compared to *meq*-W. Despite these differences, all three mutant viruses had a similar effect *in vivo*, which incidentally corresponds to the *in vitro* gene expression levels of the respective *meq* versions. However, changes in the dinucleotide frequencies in *meq* could be too small to cause an observable phenotypic change, possibly owing to its considerably small size (~1 kb)[7] compared to the large MDV genome (~180 kb).

Even though CPBD of *meq* caused a slight reduction in the tumour incidence and lead to a somewhat delayed MD progression, the virus containing *meq*-D was not completely attenuated. On the other hand, merely possessing functional copies of *meq*

---

[6] Dinucleotide abundance was calculated as the ratio of the observed frequency of the dinucleotide to the product of the observed frequencies of its constituent nucleotides, as described previously (17).

[7] Size of vvMDV strain RB-1B *meq* gene = 1,020 bp

does not render the virus oncogenic, as is evident from the fact that the vaccine strain CVI988/Rispens possesses two functional copies of *meq* and does not cause tumours (113). An *in vitro* comparative study of the functional properties of *meq* genes from CVI988/Rispens and a vvMDV strain Md5 suggests that both of the Meq proteins of the vaccine strain are weaker transactivators than that of Md5 (114). Thus, alterations in *meq* may lead to an attenuated phenotype. However, since Meq-D preserved the amino acid composition and therefore the structural properties of Meq-W (vvMDV strain RB-1B), such a difference in its role as a transactivator can be ruled out. Moreover, the decrease in protein production *in vitro* caused by CPB alteration (**Figure 14B**) was also not sufficient to render the virus non-oncogenic. It can be argued that the structural properties of the Meq protein play a more important role in transformation as opposed to quantity of the produced protein.

Undoubtedly, CPBD of the *meq* gene *alone* was insufficient to render the virus non-oncogenic. A better understanding of the CPB phenomenon in the future would perhaps allow for further deoptimization of *meq*, possibly even rendering it non-oncogenic. However, since *meq* is only a small (yet important) gene of MDV, the level of attenuation can be increased by CPBD of additional genes, particularly those involved in virus replication. The magnitude of the initial cytolytic infection plays an important role in MDV tumorigenesis by ensuring the establishment of sufficient number of latently infected cells which, in turn, induce lymphomas (35,51). Therefore, reducing the initial level of replication by CPBD of essential genes might aid the reduction in oncogenicity caused by CPB deoptimized *meq*, perhaps even leading to complete attenuation of the virus. CPBD of the U_L30 gene, which encodes the catalytic subunit of the viral DNA polymerase and is essential for MDV replication, not only reduces viral replication but also causes delayed MD progression and lower mortality (140). Thus, although CPBD of a lone MDV gene seems to reduce its pathogenicity, it is not sufficient for complete attenuation of the virus, irrespective of whether the said gene plays a role in viral replication or host-control. Another possible solution could be to delete *meq*, considering the outcomes of the present study and its non-essential role in viral replication. This approach has indeed been studied, however, the candidate vaccines caused atrophy of the primary lymphoid organs, thymus and bursa of Fabricius, in the inoculated chickens. This atrophy was highly correlated to the magnitude of initial viral replication (141). When this drawback of the vaccine candidate was overcome by addition of a U_L5 helicase-primase subunit point mutation, the resulting recombinant had reduced vaccinal potential (142). The key is perhaps to find a balance between sufficiently reduced replication as well as reduced virulence. A synergistic effect of CPBD of *meq* and U_L30 might allow us to achieve this balance. To augment the effects of

CPB deoptimized *meq* and to achieve further reduction in virulence*,* other host control factors like pp38 and vIL-8 could also be targeted. At the same time, CPBD of one or more of the essential genes encoding the several surface-glycoproteins of MDV, namely, gB, gC, gD, gE, gI, and gH, could help in fine tuning the replication potential. CPBD of genes coding for envelope glycoproteins, indeed, has been proven to cause attenuation of influenza A virus, HIV-1, PRRSV and dengue virus (18,28–30,32,34).

In conclusion, although CPBD deoptimization of its major oncogene, *meq*, could not completely attenuate MDV, it reduced protein production *in vitro*, prolonged MD progression *in vivo* and reduced tumour incidence. Thus, CPBD can be a viable strategy for attenuating MDV, however, simultaneous CPB deoptimization of one or more additional genes will perhaps better attenuate the virus. Future studies should focus on simultaneously targeting multiple genes including those involved in replication as well as those involved in host control via immune evasion or pathogenesis. Further understanding of the exact mechanism(s) governing CPB, particularly the role of CPBD in virus attenuation, will not only help improve the recoding strategy but also the selection of target genes. The latter is especially important in case of viruses like MDV, where recoding of the entire genome is cumbersome due to their large genome size.

9. Summary

Codon pair bias deoptimization of a major oncogene as an attenuation strategy for Marek's disease herpesvirus

Degeneracy of the genetic code enables the encoding of most amino acids by more than one codon. Consequently, there is a vast number of ways in which a protein can be encoded. For example, a peptide with 300 amino acid residues can be encoded in 10^{151} different ways. However, in actual coding sequences, the usage of synonymous codons is biased, that is some codons occur more often than their synonymous counterparts, a phenomenon called codon bias. Similarly, but independently of codon bias, occurrence of codon pair combinations in open reading frames (ORFs) is also biased. That is, certain codon pairs appear in coding sequences more often than it would be expected based on the overall frequencies of their constituent codons. This phenomenon, dubbed codon pair bias (CPB), has been found in all the species studied so far and has a greater impact on translational efficiency than codon bias.

Virus attenuation by CPB deoptimization (CPBD) is achieved by reshuffling the existing synonymous codons in an ORF without changing the amino acid composition of the encoded protein. The reshuffling thus retains codon bias while simultaneously changing the CPB of the mutated genes. The mechanism behind CPBD mediated attenuation of viruses is currently explained by two competing, however poorly understood, theories: (i) CPBD increases the number of naturally underrepresented codon pair combinations, which results in inefficient translation leading to reduced protein production, and consequently virus attenuation. (ii) CPBD inadvertently increases the number of CpG and TpA dinucleotides in the sequence enabling the recoded viruses to be recognized by a yet to be identified innate immune mechanism.

Despite the unclear underlying mechanism, CPBD based attenuation strategy has been successfully employed to attenuate several unrelated RNA viruses, e.g. poliovirus, influenza A virus, dengue virus, etc. The resultant attenuated virus is not only antigenically identical to the parental pathogenic virus but the sheer number of introduced mutations also makes it genetically extremely stable. Moreover, being a systematic method, it is seemingly applicable to several viruses. However, its suitability for attenuating large double-stranded DNA viruses like herpesviruses remained to be studied. Therefore, the present study dealt with the application of CPBD to attenuate a very virulent strain RB-1B of Marek's disease virus (MDV), a highly contagious, lymphoproliferative and immunosuppressive herpesvirus that infects chickens.

Due to its large size, CPBD of the entire MDV genome was impossible. Hence, *meq,* the major oncogene, was selected as a target gene. Despite being non-essential for viral

growth *in vitro, meq* is a crucial immunogen as well as one of the few genes that are expressed during latency. CPB deoptimized *meq* showed lower protein production compared to the parental gene *in vitro.* A mutant virus, in which the parental *meq* was replaced with the CPB deoptimized version, showed comparable growth in cell culture, however, caused a slightly slower disease progression as well as decreased tumour incidence *in vivo.* These results indicate that although CPBD of the *meq* gene alone could not completely attenuate the virus, CPBD might be a suitable attenuation strategy for MDV.

10. Zusammenfassung

Codon-Paar-Deoptimierung eines Hauptonkogens als Strategie zur Attenuierung des Herpesvirus der Marekschen Krankheit

Die sogenannte Degeneration des genetischen Codes ermöglicht die Verschlüsselung der meisten Aminosäuren durch mehr als nur ein Codon. Somit gibt es sehr viele Wege ein Protein zu codieren. Als Beispiel kann ein Peptid bestehend aus 300 Aminosäuren auf 10^{151} verschiedene Arten codiert werden. Der Gebrauch von synonymen Codons bei der Proteincodierung ist jedoch nicht zufällig. Bestimmte Codons werden häufiger als ihre synonym gebrauchten Gegenstücke benutzt. Dieses Phänomen wird als Codon Bias bezeichnet. Unabhängig von dem Codon Bias ist auch das Vorhandensein der Codon-Paar-Kombinationen in den offenen Leserahmen (OLR; engl.: open reading frame, ORF) nicht zufällig. Das heißt, dass bestimmte Codon-Paar-Kombinationen häufiger vorkommen, als man basierend auf der Verwendungsfrequenz der einzelnen Codons, aus denen die Paare bestehen, annehmen würde. Dieses Codon-Paar-Bias (CPB) genannte Phänomen konnte bei allen bisher daraufhin untersuchten Arten nachgewiesen werden und hat einen größeren Einfluss auf die Effizienz der Translation als das Codon Bias.

Virus-Attenuierung durch CPB-Deoptimierung (CPBD) kann erreicht werden, indem die vorhandenen synonymen Codons in einem OLR neu angeordnet werden, ohne dabei die bestehende Aminosäurenkomposition zu verändern. Die Umstrukturierung verändert das CPB, während das Codon-Bias unverändert bleibt. Es gibt zurzeit zwei konkurrierende, jedoch noch wenig erforschte Theorien, die versuchen den Mechanismus zu erklären, welcher bei der CPBD zur Virus-Attenuierung führt: (i) CPBD erhöht die Zahl der Codon-Paar-Kombination, die naturgemäß vermieden werden, was eine ineffizientere Translation und somit eine verringerte Proteinproduktion zur Folge hat. So führt die CPBD zur Attenuierung des Virus. (ii) CPBD erhöht unbeabsichtigterweise die Anzahl der CpG und TpA Dinucleotide in der deoptimierten Sequenz, wodurch das umcodierte Virus von einem noch nicht identifizierten Mechanismus des angeborenen Immunsystem erkannt wird.

Trotz fehlender Einsicht in den Mechanismus sind bereits verschiedene, nicht verwandte RNA-Viren wie z.B. Poliovirus, Influenza-A-Virus und Denguevirus erfolgreich mit der an CPBD angelehnten Strategie abgeschwächt worden. Die resultierenden, attenuierten Viren sind nicht nur antigenisch identisch mit dem Wildtyp-Virus, sondern durch die Vielzahl der eingefügten Punktmutationen auch genetisch äußerst stabil. Als systematische Methode kann die CPBD vermutlich auf eine Vielzahl von Viren angewendet werden. Ob diese Strategie auch zur Abschwächung großer doppelsträngiger-

DNA-Viren angewendet werden kann, wurde bisher noch nicht untersucht. Aus diesem Grund beschäftigte sich die vorliegende Arbeit mit der Verwendung der CPBD zur Attenuierung eines sehr virulenten Stamms, nämlich des RB-1B-Stamms, des Virus der Marekschen Krankheit (MDV). MDV ist ein hochansteckendes, lymphoproliferatives, immunsuppressives Herpesvirus, welches Hühner infiziert.

Die CPBD des gesamten Genoms war bedingt durch seine Größe unmöglich. Deswegen wurde in dieser Arbeit *meq*, das Hauptonkogen von MDV, zum Gegenstand der Untersuchung gemacht. Obwohl *meq* nicht essenziell für das *in vitro* Wachstum des Virus ist, ist es ein wichtiges Immunogen und gehört zu den wenigen Genen, die während der Latenz exprimiert werden. Die CPBD von *meq* hatte im Vergleich zum Wildtyp-Gen eine geringere Proteinproduktion *in vitro* zur Folge. Eine Virusmutante, bei der *meq* durch eine CPB-deoptimierte Variante des Gens ersetzt wurde, zeigte ein mit dem Wildtyp vergleichbares Wachstum in Zellkultur. Dennoch verursachte die Virusmutante einen leicht verlangsamten Krankheitsverlauf sowie eine herabgesetzte Tumorinzidenz *in vivo*. Obwohl die CPBD von *meq* allein nicht zu einer vollständigen Attenuierung des Virus führte, deuten die erlangten Ergebnisse darauf hin, dass die CPBD eine geeignete Methode zur Attenuierung von MDV sein kann.

11. References

1.   Hussein IH, Chams N, Chams S, El Sayegh S, Badran R, Raad M, et al. Vaccines Through Centuries: Major Cornerstones of Global Health. Front Public Heal. 2015;3:269.

2.   Haaheim LR, Pattison JR. Virus Vaccines. In: Haaheim LR, Pattison JR, Whitley RJ, editors. A Practical Guide to Clinical Virology. 2nd Ed. John Wiley & Sons, Ltd., Chichester, England; 1989. p. 37–44.

3.   Vaccine Timeline [Internet]. Immunization Action Coaliton. 2016 [cited 2016 Sep 13]. Available from: http://www.immunize.org/timeline/

4.   Close RM, Pearson C, Cohn J. Vaccine-preventable disease and the under-utilization of immunizations in complex humanitarian emergencies. Vaccine. 2016;34(39):4649–55.

5.   Moi ML, Takasaki T, Kurane I. Human antibody response to dengue virus : implications for dengue vaccine design. Trop Med Health. 2016;44(1):1–6.

6.   Graham BS, Crowe JEJ. Immunization against viral diseases. In: Knipe DM, Howley PM, editors. Fields Virology. 5th Ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA; 2007. p. 487–538.

7.   Coleman JR, Papamichail D, Skiena S, Futcher B, Wimmer E, Mueller S. Virus Attenuation by Genome-Scale Changes in Codon Pair Bias. Science. 2008;320(5884):1784–7.

8.   Goldberg AL, Wittes RE. Genetic code: aspects of organization. Science. 1966;153(3734):420–4.

9.   Ikemura T. Codon usage and tRNA content in unicellular and multicellular organisms. Mol Biol Evol. 1985;2(1):13–34.

10.   Gutman GA, Hatfield GW. Nonrandom utilization of codon pairs in Escherichia coli. Proc Natl Acad Sci U S A. 1989;86(10):3699–703.

11.   Nogales A, Baker SF, Ortiz-Riano E, Dewhurst S, Topham DJ, Martinez-Sobrido L. Influenza A Virus Attenuation by Codon Deoptimization of the NS Gene for Vaccine Development. J Virol. 2014;88(18):10525–40.

12.   Dix DB, Thompson RC. Codon choice and gene expression: synonymous codons differ in translational accuracy. Proc Natl Acad Sci U S A. 1989;86(18):6888–92.

13.   Moura G, Pinheiro M, Silva R, Miranda I, Afreixo V, Dias G, et al. Comparative context analysis of codon pairs on an ORFeome scale. Genome Biol. 2005;6(3):R28.

14.   Zhou J, Liu WJ, Peng SW, Sun XY, Frazer I. Papillomavirus capsid protein expression level depends on the match between codon usage and tRNA availability. J Virol. 1999;73(6):4972–82.

References

15. Webster G, Teh AY-H, Ma JK-C. Synthetic gene design - The rationale for codon optimization and implications for molecular pharming in plants. Biotechnol Bioeng. 2016;1–34.

16. Mueller S, Papamichail D, Coleman JR, Skiena S, Wimmer E. Reduction of the rate of poliovirus protein synthesis through large-scale codon deoptimization causes attenuation of viral virulence by lowering specific infectivity. J Virol. 2006;80(19):9687–96.

17. Kunec D, Osterrieder N. Codon Pair Bias Is a Direct Consequence of Dinucleotide Bias. Cell Rep. 2016;14(1):55–67.

18. Mueller S, Coleman JR, Papamichail D, Ward CB, Nimnual A, Futcher B, et al. Live Attenuated Influenza Vaccines by Computer-Aided Rational Design. Nat Biotechnol. 2010;28(7):723–6.

19. Moura GR, Pinheiro M, Freitas A, Oliveira JL, Frommlet JC, Carreto L, et al. Species-specific codon context rules unveil non-neutrality effects of synonymous mutations. PLoS One. 2011;6(10):e26817.

20. Chevance FFV, Le Guyon S, Hughes KT. The Effects of Codon Context on In Vivo Translation Speed. PLoS Genet. 2014;10(6):e1004392.

21. Irwin B, Heck JD, Hatfield GW. Codon pair utilization biases influence translational elongation step times. J Biol Chem. 1995;270(39):22801–6.

22. Precup J, Parker J. Missense misreading of asparagine codons as a function of codon identity and context. J Biol Chem. 1987;262(23):11351–5.

23. Shah AA, Giddings MC, Parvaz JB, Gesteland RF, Atkins JF, Ivanov IP. Computational identification of putative programmed translational frameshift sites. Bioinformatics. 2002;18(8):1046–53.

24. Buchan JR, Aucott LS, Stansfield I. tRNA properties help shape codon pair preferences in open reading frames. Nucleic Acids Res. 2006;34(3):1015–27.

25. Greenbaum BD, Rabadan R, Levine AJ. Patterns of oligonucleotide sequences in viral and host cell RNA identify mediators of the host innate immune system. PLoS One. 2009;4(6):e5969.

26. Atkinson NJ, Witteveldt J, Evans DJ, Simmonds P. The influence of CpG and UpA dinucleotide frequencies on RNA virus replication and characterization of the innate cellular pathways underlying virus attenuation and enhanced replication. Nucleic Acids Res. 2014;42(7):4527–45.

27. Coleman JR, Papamichail D, Yano M, Del Mar García-Suárez M, Pirofski LA. Designed reduction of Streptococcus pneumoniae pathogenicity via synthetic changes in virulence factor codon-pair bias. J Infect Dis. 2011;203(9):1264–73.

28. Yang C, Skiena S, Futcher B, Mueller S, Wimmer E. Deliberate reduction of hemagglutinin and neuraminidase expression of influenza virus leads to an ultraprotective live vaccine in mice. Proc Natl Acad Sci U S A. 2013;110(23):9481–6.

29. Broadbent AJ, Santos CP, Anafu A, Wimmer E, Mueller S, Subbarao K. Evaluation of the attenuation, immunogenicity, and efficacy of a live virus vaccine generated by codon-pair bias de-optimization of the 2009 pandemic H1N1 influenza virus, in ferrets. Vaccine. 2016;34(4):563–70.

30. Martrus G, Nevot M, Andres C, Clotet B, Martinez MA. Changes in codon-pair bias of human immunodeficiency virus type 1 have profound effects on virus replication in cell culture. Retrovirology. 2013;10(1):78.

31. Le Nouën C, Brock LG, Luongo C, McCarty T, Yang L, Mehedi M, et al. Attenuation of human respiratory syncytial virus by genome-scale codon-pair deoptimization. Proc Natl Acad Sci U S A. 2014;111(36):13169–74.

32. Ni YY, Zhao Z, Opriessnig T, Subramaniam S, Zhou L, Cao D, et al. Computer-aided codon-pairs deoptimization of the major envelope GP5 gene attenuates porcine reproductive and respiratory syndrome virus. Virology. 2014;450–451:132–9.

33. Wang B, Yang C, Tekes G, Mueller S, Paul A, Whelan SPJ. Recoding of the Vesicular Stomatitis Virus L Gene by Computer- Aided Design Provides a Live , Attenuated Vaccine Candidate. MBio. 2015;6(2):1–10.

34. Shen SH, Stauft CB, Gorbatsevych O, Song Y, Ward CB, Yurovsky A, et al. Large-scale recoding of an arbovirus genome to rebalance its insect versus mammalian preference. Proc Natl Acad Sci U S A. 2015;112(15):4749–54.

35. Osterrieder N, Kamil JP, Schumacher D, Tischer BK, Trapp S. Marek's disease virus: from miasma to model. Nat Rev Microbiol. 2006;4(4):283–94.

36. Churchill AE, Biggs PM. Agent of Marek's disease in tissue culture. Nature. 1967;215:528–30.

37. Churchill AE, Biggs PM. Herpes-Type Virus Isolated in Cell Culture From Tumors of Chickens With Marek's Disease. II. Studies In Vivo. J Natl Cancer Inst. 1968;41(4):951–6.

38. Churchill AE. Herpes-Type Virus Isolated in Cell Culture From Tumors of Chickens With Marek's Disease. I. Studies In Cell Culture. J Natl Cancer Inst. 1968;41(4):951–6.

39. Fukuchi K, Tanaka A, Schierman LW, Witter RL, Nonoyama M. The structure of Marek disease virus DNA - the presence of unique expansion in nonpathogenic viral DNA. Proc Natl Acad Sci U S A. 1985;82(3):751–4.

40. Tulman ER, Afonso CL, Lu Z, Zsak L, Rock DL, Kutish GF. The genome of a very virulent Marek's disease virus. J Virol. 2000;74(17):7980–8.

41. Lee LF, Wu P, Sui D, Ren D, Kamil J, Kung HJ, et al. The complete unique long sequence and the overall genomic organization of the GA strain of Marek's disease virus. Proc Natl Acad Sci U S A. 2000;97(11):6091–6.

42. Osterrieder K, Vautherot J-F. The genome content of Marek's Disease-like viruses. In: Davison F, Nair V, editors. Marek's Disease: An Evolving Problem. Elsevier Academic Press, London, UK; 2004. p. 17–31.

43. Davison AJ, Eberle R, Ehlers B, Hayward GS, McGeoch DJ, Minson AC, et al. The order Herpesvirales. Arch Virol. 2009;154(1):171–7.

44. Morrow C, Fehler F. Marek's disease: a worldwide problem. In: Davison F, Nair V, editors. Marek's Disease: An Evolving Problem. London, UK: Elsevier Academic Press; 2004. p. 49–61.

45. Zelnik V. Diagnosis of Marek's disease. In: Davison F, Nair V, editors. Marek's Disease: An Evolving Problem. Elsevier Academic Press, London, UK; 2004. p. 156–67.

46. Pappenheimer AM, Dunn LC, Cone V. Studies on fowl paralysis (neurolymphomatosis gallinarum): I. Clinical features and pathology. J Exp Med. 1929;49(1):63–86.

47. Payne LN. Pathological responses to infection. In: Davison F, Nair V, editors. Marek's Disease. Elsevier Academic Press, London, UK; 2004. p. 78–97.

48. Nazerian K. Marek's disease lymphoma of chicken and its causative herpesvirus. Biochim Biophys Acta. 1979;560(3):375–95.

49. Cebrian J, Kaschka-Dierich C, Berthelot N, Sheldrick P. Inverted repeat nucleotide sequences in the genomes of Marek disease virus and the herpesvirus of the turkey. Proc Natl Acad Sci U S A. 1982;79(2):555–8.

50. Carrozza JH, Fredrickson TN, Prince RP, Luginbuhl RE. Role of Desquamated Epithelial Cells in Transmission of Marek's Disease. Avian Dis. 1973;17(4):767–81.

51. Baigent SJ, Davison F. Marek's disease virus: biology and life cycle. In: Davison F, Nair V, editors. Marek's Disease An Evolving Problem. Elsevier Academic Press, London, UK; 2004. p. 62–76.

52. Baigent SJ, Ross LJ, Davison TF. Differential susceptibility to Marek's disease is associated with differences in numbers, but not phenotype or location, of pp38 lymphocytes. J Gen Virol. 1998;79:2795–802.

53. Baigent SJ, Davison TF. Development and composition of lymphoid lesions in the spleens of Marek's disease virus-infected chickens: Association with virus spread and the pathogenesis of Marek's disease. Avian Pathol. 1999;28(3):287–300.

54. Calnek BW, Schat KA, Ross LJN, Chen C-LH. Further Characterization of Marek â€™ S Disease Virus-Infected Lymphocytes . II. in Vitro Infection. Int J Cancer. 1984;33:399–406.

55. Virgin S. Pathogenesis of Viral Infection. In: Knipe DM, Howley PM, editors. Fields Virology. 5th Ed. Lippincott Williams & Wilkins, Philadelphia, PA, USA; 2007. p. 328–88.

56. Adldinger HK, Calnek BW. Pathogenesis of Marek's Disease: Early Distribution of Virus and Viral Antigens in Infected Chicken. J Natl Cancer Inst. 1973;50:1287–98.

57. Delecluse H-J, Hammerschmidt W. Status of Marek's Disease Virus in Established Lymphoma Cell Lines: Herpesvirus Integration Is Common. J Virol. 1993;67(1):82–92.

58. Kaufer BB, Jarosinski KW, Osterrieder N. Herpesvirus telomeric repeats facilitate genomic integration into host telomeres and mobilization of viral DNA during reactivation. J Exp Med. 2011;208(3):605–15.

59. Greco A, Fester N, Engel AT, Kaufer BB. Role of the short telomeric repeat region in Marek's disease virus replication, genomic integration, and lymphomagenesis. J Virol. 2014;88(24):14138–47.

60. Morissette G, Flamand L. Herpesviruses and chromosomal integration. J Virol. 2010;84(23):12100–9.

61. Prusty BK, Gulve N, Rasa S, Murovska M, Hernandez PC, Ablashi D V. Possible chromosomal and germline integration of human herpesvirus 7. J Gen Virol. 2017;98(2):266–74.

62. Cantello JL, Parcells MS, Anderson AS, Morgan RW. Marek's disease virus latency-associated transcripts belong to a family of spliced RNAs that are antisense to the ICP4 homolog gene. J Virol. 1997;71(2):1353–61.

63. Cantello JL, Anderson AS, Morgan RW. Identification of latency-associated transcripts that map antisense to the ICP4 homolog gene of Marek's disease virus. J Virol. 1994;68(10):6280–90.

64. Li D-S, Pastorek J, Zelnik V, Smith GD, Ross LJN. Identification of novel transcripts complementary to the Marek's disease virus homologue of the ICP4 gene of herpes simplex virus. J Gen Virol. 1994;75(7):1713–22.

65. Kawamura M, Hayashi M, Furuichi T, Nonoyama M, Isogai E, Namioka S. The inhibitory effects of oligonucleotides, complementary to Marek's disease virus mRNA transcribed from the BamHI-H region, on the proliferation of transformed lymphoblastoid cells, MDCC-MSB1. J Gen Virol. 1991;72(5):1105–11.

66. Davison F, Kaiser P. Immunity to Marek's disease. In: Davison F, Nair V, editors. Marek's Disease: An Evolving Problem. Elsevier Academic Press, London, UK; 2004. p. 126–41.

67. Buscaglia C, Calnek BW. Maintenance of Marek's Disease Herpesvirus Latency in vitro by a Factor Found in Conditioned Medium. J Gen Virol. 1988;69:2809–18.

68. Pappenheimer AM, Dunn LC, Cone V. Studies on Fowl Paralysis (Neurolymphomatosis gallinarium). I. Clinical Features and Pathology. J Exp Med. 1929;49(1):63–86.

69. Pappenheimer AM, Dunn LC, Seidlin SM. Studies on foul paralysis (Neurolymphomatosis gallinarum). II. Transmission Experiments. J Exp Med. 1929;49(1):87–102.

70. Burgess SC, Davison TF. Identification of the neoplastically transformed cells in Marek's disease herpesvirus-induced lymphomas: recognition by the monoclonal antibody AV37. J Virol. 2002;76(14):7276–92.

71. Delecluse HJ, Schüller S, Hammerschmidt W. Latent Marek's disease virus can be activated from its chromosomally integrated state in herpesvirus-transformed lymphoma cells. EMBO J. 1993;12(8):3277–86.

72. Burgess SC, Young JR, Baaten BJG, Hunt L, Ross LNJ, Parcells MS, et al. Marek's disease is a natural model for lymphomas overexpressing Hodgkin's disease antigen (CD30). Proc Natl Acad Sci. 2004;101(38):13879–84.

73. Ross NLJ. T-cell transformation by Marek's disease virus. Trends Microbiol. 1999;7(1):22–9.

74. Nair V, Kung H-J. Marek's disease virus oncogenicity: molecular mechanisms. In: Davison F, Nair V, editors. Marek's Disease: An Evolving Problem. Elsevier Academic Press, London, UK; 2004. p. 32–48.

75. Peng Q, Zeng M, Bhuiyan ZA, Ubukata E, Tanaka A, Nonoyama M, et al. Isolation and Characterization of Marek's Disease Virus (MDV) cDNAs Mapping to the BamHI-I 2, BamHI-Q 2, and BamHI-L Fragments of the MDV Genome from Lymphoblastoid Cells Transformed and Persistently Infected with MDV. Virology. 1995;213(2):590–9.

76. Xie Q, Anderson AS, Morgan RW. Marek's Disease Virus (MDV) ICP4, pp38, and meq Genes Are Involved in the Maintenance of Transformation of MDCC-MSB1 MDV-Transformed Lymphoblastoid Cells. J Virol. 1996;70(2):1125–31.

77. Engel AT, Selvaraj RK, Kamil JP, Osterrieder N, Kaufer BB. Marek's disease viral interleukin-8 promotes lymphoma formation through targeted recruitment of B cells and CD4+ CD25+ T cells. J Virol. 2012;86(16):8536–45.

78. Liu J-L, Ye Y, Lee LF, Kung H-J. Transforming potential of the herpesvirus oncoprotein MEQ: morphological transformation, serum-independent growth, and inhibition of apoptosis. J Virol. 1998;72(1):388–95.

79. Jones D, Lee L, Liu JL, Kung HJ, Tillotson JK. Marek disease virus encodes a basic-leucine zipper gene resembling the fos/jun oncogenes that is highly expressed in lymphoblastoid tumors. Proc Natl Acad Sci U S A. 1992;89(9):4042–6.

80. Liu JL, Ye Y, Qian Z, Qian Y, Templeton DJ, Lee LF, et al. Functional interactions between herpesvirus oncoprotein MEQ and cell cycle regulator CDK2. J Virol. 1999;73(5):4208–19.

81. Cheson BD. What is new in lymphoma? CA A Cancer J Clin. 2004;54:260–72.

82. Okada T, Takagi M, Murata S, Onuma M, Ohashi K. Identification and characterization of a novel spliced form of the meq transcript in lymphoblastoid cell lines derived from Marek's disease tumours. J Gen Virol. 2007;88(8):2111–20.

83. Jarosinski KW, Tischer BK, Trapp S, Osterrieder N. Marek's disease virus: lytic replication, oncogenesis and control. Expert Rev Vaccines. 2006;5(6):761–72.

84. Levy AM, Gilad O, Xia L, Izumiya Y, Choi J, Tsalenko A, et al. Marek's disease virus Meq transforms chicken cells via the v-Jun transcriptional cascade: a converging transforming pathway for avian oncoviruses. Proc Natl Acad Sci U S A. 2005;102(41):14831–6.

85. Zhao Y, Kurian D, Xu H, Petherbridge L, Smith LP, Hunt L, et al. Interaction of Marek's disease virus oncoprotein Meq with heat-shock protein 70 in lymphoid tumour cells. J Gen Virol. 2009;90(9):2201–8.

86. Lupiani B, Lee LF, Cui X, Gimeno I, Anderson A, Morgan RW, et al. Marek's disease virus-encoded Meq gene is involved in transformation of lymphocytes but is dispensable for replication. Proc Natl Acad Sci U S A. 2004;101(32):11815–20.

87. Davison F, Nair V. Use of Marek's disease vaccines: could they be driving the virus to increasing virulence? Expert Rev Vaccines. 2005;4(1):77–88.

88. Churchill AE, Payne LN, Chubb RC. Immunization against Marek's Disease using a Live Attenuated Virus. Nature. 1969;221(5182):744–7.

89. Churchill AE, Chubb RC, Baxendale W. The attenuation, with loss of oncogenicity, of the herpes-type virus of Marek's disease (strain HPRS-16) on passage in cell culture. J Gen Virol. 1969;4(4):557–64.

90. Okazaki W, Purchase HG, Burmester BR. Protection against Marek's disease by vaccination with a herpesvirus of turkeys. Avian Dis. 1970;14(2):413–29.

91. Witter RL. Characteristics of Marek's Disease Viruses Isolated from Vaccinated Commercial Chicken Flocks: Association of Viral Pathotype with Lymphoma Frequency. Avian Dis. 1983;27(1):113–32.

92. Calnek BW, Schat KA, Peckham MC, Fabricant J. Field Trials with a Bivalent Vaccine (HVT and SB-1) against Marek's Disease. Avian Dis. 1983;27(3):844–9.

93. Witter RL. Increased virulence of Marek's disease virus field isolates. Avian Dis. 1997;41(1):149–63.

94. Witter RL. Control Strategies for Marek's Disease: A Perspective for the Future. Poult Sci. 1998;77(8):1197–203.

References

95.   Witter RL. The changing landscape of Marek's disease. Avian Pathol. 1998;27(sup1):S46–53.

96.   Gandon S, Mackinnon MJ, Nee S, Read AF. Imperfect vaccines and the evolution of pathogen virulence. Nature. 2001;414(6865):751–6.

97.   Kupferschmidt K. Risk of "leaky" vaccines debated. Science. 2015;349(6247):461–2.

98.   Read AF, Baigent SJ, Powers C, Kgosana LB, Blackwell L, Smith LP, et al. Imperfect vaccination can enhance the transmission of highly virulent pathogens. PLoS Biol. 2015;13(7):1–18.

99.   Bublot M, Sharma JM. Vaccination against Marek's disease. In: Davison F, Nair V, editors. Marek's Disease: An envolving problem. Elsevier Academic Press, London, UK; 2004. p. 168–85.

100.   Gimeno IM. Marek's disease vaccines: A solution for today but a worry for tomorrow? Vaccine. 2008;26(SUPPL. 3):C31–41.

101.   Rispens BH, Vloten H Van, Mastenbroek N, Maas HJL, Schat A. Control of Marek's Disease in the Netherlands. I. Isolation of an Avirulent Marek's Disease Virus (Strain CVI 988) and Its Use in Laboratory Vaccination Trials. Avian Dis. 1972;16(1):108–25.

102.   Schumacher D, Tischer BK, Fuchs W, Osterrieder N. Reconstitution of Marek's Disease Virus Serotype 1 (MDV-1) from DNA Cloned as a Bacterial Artificial Chromosome and Characterization of a Glycoprotein B-Negative MDV-1 Mutant. J Virol. 2000;74(23):11088–98.

103.   Tischer BK, Schumacher D, Beer M, Teifke JP, Osterrieder K, Wink K, et al. A DNA vaccine containing an infectious Marek's disease virus genome can confer protection against tumorigenic Marek's disease in chickens. J Gen Virol. 2002;83:2367–76.

104.   Lee LE, Witter RL, Reddy SM, Wu P, Yanagida N, Yoshida S. Protection and synergism by recombinant fowl pox vaccines expressing multiple genes from Marek's disease virus. Avian Dis. 2003;47(3):549–58.

105.   Nazerian K, Lee LF, Yanagida N, Ogawa R. Protection against Marek's disease by a fowlpox virus recombinant expressing the glycoprotein B of Marek's disease virus. J Virol. 1992;66(3):1409–13.

106.   Nazerian AK, Witter RL, Lee LF, Yanagida N. Protection and Synergism by Recombinant Fowl Pox Vaccines Expressing Genes from Marek's Disease Virus. Avian Dis. 1996;40(2):368–76.

107.   Heine HG, Foord AJ, Young PL, Hooper PT, Lehrbach PR, Boyle DB. Recombinant fowlpox virus vaccines against Australian virulent Marek's disease virus: gene sequence analysis and comparison of vaccine efficacy in specific pathogen free and production chickens. Virus Res. 1997;50:23–33.

108. Haq K, Schat KA, Sharif S. Immunity to Marek's disease: Where are we now? Dev Comp Immunol. 2013;41:439–46.

109. Witter RL. Attenuated Revertant Serotype 1 Marek's Disease Viruses: Safety and Protective Efficacy. Avian Dis. 1991;35(4):877–91.

110. Alcami A, Koszinowski UH. Viral mechanisms of immune evasion. Mol Med Today. 2000;6(9):365–72.

111. Kung HJ, Xia L, Brunovskis P, Li D, Liu J-L, Lee LF. Meq: an MDV-specific bZIP transactivator with transforming properties. Curr Top Microbiol Immunol. 2001;255:245–60.

112. Schat KA, Xing Z. Specific and nonspecific immune responses to Marek's disease virus. Dev Comp Immunol [Internet]. 2000;24(2–3):201–21. Available from: www.elsevier.com/locate/devcompimm

113. Witter RL, Lee LF, Fadly AM. Characteristics of CVI988/Rispens and R2/23, Two Prototype Vaccine Strains of Serotype 1 Marek's Disease Virus. Avian Dis. 1995;39(2):269–84.

114. Ajithdoss DK, Reddy SM, Suchodolski PF, Lee LF, Kung HJ, Lupiani B. In vitro characterization of the Meq proteins of Marek's disease virus vaccine strain CVI988. Virus Res. 2009;142(1–2):57–67.

115. GraphPad Prism version 7.02 for Windows [Internet]. GraphPad Software, La Jolla, CA, USA. Available from: www.graphpad.com

116. pDRAW32 [Internet]. AcaClone Software (Kjeld Olesen); Available from: http://www.acaclone.com/

117. Tischer BK, Von Einem J, Kaufer B, Osterrieder N. Two-step Red-mediated recombination for versatile high-efficiency markerless DNA manipulation in Escherichia coli. Biotechniques. 2006;40(2):191–7.

118. Schippers T. Functional characterization of the potential immune evasion proteins pUL49.5 and p012 of Marek's disease virus (MDV). Dr. rer. nat. [Dissertation] [Internet]. Berlin: Freie Universität Berlin; 2015. Available from: http://www.diss.fu-berlin.de/diss/receive/FUDISS_thesis_000000098929

119. Jarosinski KW, Osterrieder N, Nair VK, Schat KA. Attenuation of Marek's Disease Virus by Deletion of Open Reading Frame RLORF4 but Not RLORF5a. J Virol. 2005;79(18):11647–59.

120. Suchodolski PF, Izumiya Y, Lupiani B, Ajithdoss DK, Lee LF, Kung HJ, et al. Both homo and heterodimers of Marek's disease virus encoded Meq protein contribute to transformation of lymphocytes in chickens. Virology. 2010;399(2):312–21.

121. Tischer BK, Smith GA, Osterrieder N. En passant mutagenesis: a two step markerless red recombination system. Methods Mol Biol. 2010;634:421–30.

References

122. Kirkpatrick S, Gelatt CDJ, Vecchi MP. Optimization by Simulated Annealing. Science. 1982;220(4598):671–80.

123. Parcells MS, Arumugaswami V, Prigge JT, Pandya K, Dienglewicz RL. Marek's disease virus reactivation from latency: changes in gene expression at the origin of replication. Poult Sci. 2003;82(6):893–8.

124. Jarosinski KW, Schat KA. Multiple alternative splicing to exons II and III of viral interleukin-8 (vIL-8) in the Marek's disease virus genome: The importance of vIL-8 exon I. Virus Genes. 2007;34(1):9–22.

125. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003;31(13):3406–15.

126. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7(1):539.

127. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, et al. Analysis Tool Web Services from the EMBL-EBI. Nucleic Acids Res. 2013;41(Web Server issue):597–600.

128. Sambrook J, Russel D. Molecular Cloning: a laboratory manual. 3rd. Ed. Argentine J, Irwin N, editors. Cold Spring Harbor Laboratory Press, N.Y.; 2001.

129. NEB Tm Calculator v1.9.4 [Internet]. New England Biolabs. Available from: http://tmcalculator.neb.com

130. Lee EC, Yu D, Martinez de Velasco J, Tessarollo L, Swing D a, Court DL, et al. A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics. 2001;73(1):56–65.

131. Petherbridge L, Brown AC, Susan J, Howes K, Sacco MA, Nair VK, et al. Oncogenicity of Virulent Marek's Disease Virus Cloned as Bacterial Artificial Chromosomes. J Virol. 2004;78(23):13376–80.

132. Brune W, Messerle M, Koszinowski UH. Forward with BACs - New tools for herpesvirus genomics. Trends Genet. 2000;16(6):254–9.

133. Schat KA, Purchase H. Cell-culture methods. In: Swayne DE, Glisson JR, Jacwood MW, Pearson JE, Reed WM, editors. A Laboratory Manual for the Isolation and Identification of Avian Pathogens. 4th. Ed. American Association of Avian Pathologists, Kennett Square, PA, USA; 1998. p. 223–34.

134. Longo PA, Kavran JM, Kim M, Leahy DJ. Transient Mammalian Cell Transfection wtih Polyethylenimine. Methods Enzymol. 2013;529:227–40.

135. Schumacher D, Tischer BK, Trapp S, Osterrieder N. The Protein Encoded by the US3 Orthologue of Marek's Disease Virus Is Required for Efficient De-Envelopment of Perinuclear Virions and Involved in Actin Stress Fiber Breakdown. J Virol. 2005;79(7):3987–97.

136. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671–5.

137. Hatfield GW, Roth DA. Optimizing scaleup yield for protein production: Computationally Optimized DNA Assembly (CODA) and Translation Engineering™. In: El-Gewely MR, editor. Biotechnology Annual Review, Vol 13. Burlington: Elsevier Science; 2007. p. 27–42.

138. Quax TEF, Claassens NJ, Söll D, van der Oost J. Codon Bias as a Means to Fine-Tune Gene Expression. Mol Cell. 2015;59(2):149–61.

139. Chu D, Kazana E, Bellanger N, Singh T, Tuite MF, Von Der Haar T. Translation elongation can control translation initiation on eukaryotic mRNAs. EMBO J. 2014;33(1):21–34.

140. Eschke K, Osterrieder N, Kunec D. Attenuation of a very virulent Marek's disease herpesvirus by codon pair deoptimization. Paper presented at: In: 11th International Symposium on Marek's Disease and Avian Herpesviruses, 2016 July 6-9; Tours, France.

141. Dunn JR, Silva RF. Ability of MEQ-Deleted MDV Vaccine Candidates to Adversely Affect Lymphoid Organs and Chicken Weight Gain. Avian Dis. 2012;56(3):494–500.

142. Hildebrandt E, Dunn JR, Cheng HH. Addition of a UL5 helicase-primase subunit point mutation eliminates bursal–thymic atrophy of Marek's disease virus ∆Meq recombinant virus but reduces vaccinal protection. Avian Pathol. 2015;44(4):254–8.

12. Supplementary information

12.1. Sequences

The splice donor site, D2 (cacctacGTaagga**)**, present in the *meq* ORF was preserved while recoding of the sequence and is underlined.

12.1.1. *meq*-W

(Parental, CPB score= -0.057)

ATGTCTCAGGAGCCAGAGCCGGGCGCTATGCCCTACAGTCCCGCTGACGATCCGTCCCCCTCGATCT
TTCTCTCGGGTCGACTTCGAGACGGAAAAAAAGGAAAGTCACGACATCCCCAACAGCCCCTCCAAAC
ACCCCTTCCCTGACGGCCTATCTGAGGAGGAGAAACAGAAGCTGGAAAGGAGGAGAAAAAGGAATCGT
GACGCCGCTCGGAGAAGACGCAGGAAGCAGACGGACTATGTAGACAAACTCCATGAAGCATGTGAAGA
GCTGCAGAGGGCCAATGAA**cacctacGTaagga**AATTCGAGATCTAAGGACTGAGTGCACGTCCCTGC
GTGTACAGTTGGCTTGTCATGAGCCAGTTTGCCCTATGGCGGTACCCCTAACGGTGACCCTTGGACTG
CTTACCACCCCGCACGATCCCGTTCCTGAACCTCCCATTTGCACTCCTCCACCTCCCTCACCGGATGA
ACCTAACGCTCCACATTGCTCCGGTTCCCAACCTCCTATCTGTACCCCCCCTCCTCCCGATACGGAGG
AACTTTGCGCCCAGCTCTGCTCGACCCCACCACCTCCCATCTCTACTCCCCATATTATCTACGCTCCG
GGGCCTTCCCCCCTCCAACCTCCTATCTGTACCCCCCCTCCTCCCGATGCGGAGGAGCTTTGCGCCCA
GCTCTGCTCGACCCCACCACCTCCCATCTGTACTCCCCATTCCCTCTTCTGCCCTCCCCAGCCTCCAT
CTCCGGAGGGCATCTTCCCTGCATTGTGTCCTGTTACCGAGCCGTGTACCCCTCCATCGCCGGGGACG
GTTTACGCTCAGCTTTGTCCTGTTGGCCAGGCTCCCCTTTTTACCCCATCTCCCCCACATCCGGCTCC
GGAGCCGGAGAGGCTTTATGCTCGTCTTACCGAGGATCCCGAACAGGATTCCTTGTATTCGGGCCAGA
TTTATATTCAGTTTCCCTCGGATACTCAGTCTACGGTCTGGTGGTTTCCAGGTGACGGGAGACCCTGA

12.1.2. *meq*-D

(CPB deoptimized, CPB = -0.442)

ATGTCTCAGGAACCCGAGCCAGGCGCTATGCCTTATTCGCCAGCCGACGATCCGTCTCCACTCGATCT
GTCACTCGGGTCTACGTCTAGGCGAAAAAAAAGAAAGTCTCACGACATTCCGAATTCGCCTTCGAAAC
ATCCATTCCCTGACGGGCTATCCGAGGAGGAGAAACAGAAACTCGAGCGTAGACGGAAACGGAATAGG
GACGCTGCCAGGCGTAGGCGTAGAAAACAGACCGATTACGTCGATAAGCTCCACGAGGCTTGCGAAGA
GCTTCAGAGGGCTAACGAA**cacctacGTaagga**AATTCGCGATCTTAGAACCGAGTGTACGTCCCTTA
GGGTGCAACTCGCATGCCACGAACCCGTATGCCCTATGGCCGTTCCGCTTACCGTTACTCTCGGACTG
CTTACGACTCCGCATGATCCCGTTCCGGAACCCCCTATCTGTACTCCGCCACCCCCTTCCCCCGACGA
GCCTAACGCTCCGCATTGTTCCGGGTCCCAGCCCCCTATCTGTACCCCTCCCCCTCCCGATACCGAGG
AGCTTTGCGCTCAGTTGTGCTCTACCCCTCCCCCTCCGATTTCGACTCCGCATATTATCTATGCGCCA
GGCCCATCGCCACTCCAGCCCCCTATCTGTACGCCTCCCCCTCCCGACGCCGAAGAGCTTTGCGCACA
GTTGTGTTCGACTCCCCCTCCCCCTATCTGTACGCCACATAGTCTGTTTTGCCCACCTCAGCCTCCGT
CTCCCGAGGGGATCTTTCCCGCTCTATGCCCCGTTACCGAGCCTTGTACTCCCCCTAGTCCCGGTACC

GTATACGCTCAGTTGTGCCCCGTAGGCCAGGCGCCACTGTTTACGCCTAGCCCACCGCATCCCGCTCC

CGAACCCGAGAGACTTTACGCTAGGCTTACCGAGGATCCCGAACAGGATTCCCTATATTCCGGCCAGA

TTTATATCCAATTCCCTTCCGATACCCAGTCGACCGTTTGGTGGTTCCCAGGCGACGGTAGGCCTTGA

12.1.3. *meq*-O

(CPB optimized, CPB score = 0.283)

ATGTCGCAGGAGCCAGAACCAGGTGCCATGCCCTACTCTCCAGCAGATGACCCCTCGCCCTTGGATCT

TTCTCTTGGCTCCACCTCCAGGAGGAAAAAAGAAAATCTCATGACATCCCAAACAGCCCCTCCAAAC

ATCCTTTTCCAGATGGCCTTTCAGAGGAGGAGAAGCAGAAGCTGGAGAGGAGGAGGAAAAGAAACAGA

GATGCGGCGCGGCGGCGCAGGAGGAAACAAACTGACTATGTAGACAAACTACATGAAGCTTGTGAAGA

ACTACAAAGAGCAAATGAA**cacctacGTaagga**AATTAGAGATCTACGAACGGAGTGCACCTCTCTTC

GTGTTCAGCTGGCCTGCCATGAACCTGTCTGTCCCATGGCTGTTCCTTTGACTGTAACTCTTGGGCTT

CTCACCACCCCCCATGATCCTGTTCCTGAGCCCCCCATCTGCACGCCGCCGCCGCCCTCTCCAGATGA

GCCCAATGCTCCCCACTGCAGTGGCTCCCAGCCCCCCATCTGTACTCCTCCTCCTCCAGACACGGAGG

AGCTGTGTGCCCAGCTCTGCTCCACCCCACCTCCTCCCATCTCCACCCCCCACATTATTTATGCTCCT

GGGCCCTCGCCGCTGCAGCCCCCCATCTGCACCCCGCCGCCGCCGGATGCTGAGGAGCTCTGTGCTCA

GCTCTGCTCCACCCCACCTCCTCCCATCTGCACCCCCCACTCGCTCTTCTGTCCTCCTCAGCCTCCTT

CTCCCGAAGGAATTTTTCCTGCTCTTTGTCCTGTAACGGAGCCCTGCACCCCGCCCTCGCCGGGCACT

GTTTATGCTCAGCTTTGTCCTGTTGGGCAGGCCCCGCTCTTCACCCCATCTCCTCCTCATCCTGCTCC

AGAGCCAGAGAGGCTTTATGCTCGTTTGACTGAAGATCCAGAGCAGGACTCGCTCTACAGTGGGCAGA

TCTACATTCAGTTCCCCTCGGACACGCAGTCCACGGTGTGGTGGTTTCCTGGTGATGGCCGTCCCTGA

12.1.4. *meq*-R

(CPB preserved, random reshuffling, CPB score = -0.058)

ATGTCCCAGGAGCCGGAGCCTGGCGCTATGCCATATTCACCAGCCGATGATCCGTCGCCTCTTGACTT

GTCGCTCGGTTCCACCTCGAGGCGGAAGAAAAGGAAATCTCACGATATCCCCAATTCCCCATCTAAAC

ACCCTTTTCCCGACGGTCTAAGTGAGGAGGAGAAACAGAAGCTTGAGAGAAGGAGAAAAAGGAACAGA

GATGCCGCCAGACGGAGGCGCAGGAAACAAACTGATTATGTGGACAAACTACATGAGGCTTGTGAAGA

ATTGCAGCGAGCTAATGAG**cacctacGTaagga**AATTCGTGATCTCCGTACCGAGTGCACTAGCCTGA

GAGTACAACTTGCATGTCATGAGCCCGTTTGCCCCATGGCGGTACCCTTGACTGTTACCCTGGGGCTC

CTCACTACGCCCCACGATCCAGTTCCCGAACCCCCAATCTGCACCCCTCCTCCACCCAGTCCCGATGA

GCCCAACGCTCCCCATTGCTCTGGATCGCAGCCCCCGATTTGCACGCCTCCTCCACCCGATACCGAGG

AACTCTGCGCTCAGCTTTGTTCCACTCCTCCTCCCCCTATTTCCACCCCTCATATCATCTACGCACCT

GGGCCGTCGCCTCTTCAGCCCCCCATTTGTACTCCACCTCCGCCTGACGCCGAAGAGCTCTGTGCGCA

GCTATGCTCTACCCCCCCACCGCCTATTTGTACCCCTCATTCCCTGTTCTGTCCTCCACAGCCCCCTT

CGCCCGAGGGCATCTTCCCCGCTCTTTGCCCGGTTACGGAGCCCTGCACCCCGCCCTCCCCGGGGACC

GTCTATGCTCAGCTTTGTCCAGTTGGCCAGGCTCCTCTCTTCACGCCATCCCCTCCCCATCCAGCTCC

```
CGAACCTGAAAGGCTGTACGCTAGGCTTACGGAGGATCCTGAACAGGACTCTCTTTACTCTGGCCAGA

TCTATATCCAGTTTCCGTCGGACACCCAGTCTACGGTATGGTGGTTTCCTGGGGACGGCCGTCCGTGA
```

12.2. Sequence alignment

```
meq-D    ATGTCTCAGGAACCCGAGCCAGGCGCTATGCCTTATTCGCCAGCCGACGATCCGTCTCCA
meq-R    ATGTCCCAGGAGCCGGAGCCTGGCGCTATGCCATATTCACCAGCCGATGATCCGTCGCCT
meq-W    ATGTCTCAGGAGCCAGAGCCGGGCGCTATGCCCTACAGTCCCGCTGACGATCCGTCCCCC
meq-O    ATGTCGCAGGAGCCAGAACCAGGTGCCATGCCCTACTCTCCAGCAGATGACCCCTCGCCC
         ***** ***** ** ** ** ** ** ***** **    ** ** ** ** ** ** **


meq-D    CTCGATCTGTCACTCGGGTCTACGTCTAGGCGAAAAAAAAGAAAGTCTCACGACATTCCG
meq-R    CTTGACTTGTCGCTCGGTTCCACCTCGAGGCGGAAGAAAAGGAAATCTCACGATATCCCC
meq-W    CTCGATCTTTCTCTCGGGTCGACTTCGAGACGGAAAAAAAGGAAAGTCACGACATCCCC
meq-O    TTGGATCTTTCTCTTGGCTCCACCTCCAGGAGGAAAAAAAGAAAATCTCATGACATCCCA
          * **   * ** ** ** ** ** ** **   * ** ***** **   *** ** ** **


meq-D    AATTCGCCTTCGAAACATCCATTCCCTGACGGGCTATCCGAGGAGGAGAAACAGAAACTC
meq-R    AATTCCCCATCTAAACACCCTTTTCCCGACGGTCTAAGTGAGGAGGAGAAACAGAAGCTT
meq-W    AACAGCCCCTCCAAACACCCCTTCCCTGACGGCCTATCTGAGGAGGAGAAACAGAAGCTG
meq-O    AACAGCCCCTCCAAACATCCTTTTTCCAGATGGCCTTTCAGAGGAGGAGAAGCAGAAGCTG
         **      ** ** ***** ** ** ** ** ** ** ********** ***** **


meq-D    GAGCGTAGACGGAAACGGAATAGGGACGCTGCCAGGCGTAGGCGTAGAAAACAGACCGAT
meq-R    GAGAGAAGGAGAAAAAGGAACAGAGATGCCGCCAGACGGAGGCGCAGGAAACAAACTGAT
meq-W    GAAAGGAGGAGAAAAAGGAATCGTGACGCCGCTCGGAGAAGACGCAGGAAGCAGACGGAC
meq-O    GAGAGGAGGAGGAAAAGAAACAGAGATGCGGCGCGGCGGCGCAGGAGGAAACAAACTGAC
         **   * **   * *** * ** * ** ** **   *   *   * ** ** ** ** ** **


meq-D    TACGTCGATAAGCTCCACGAGGCTTGCGAAGAGCTTCAGAGGGCTAACGAACACCTACGT
meq-R    TATGTGGACAAACTACATGAGGCTTGTGAAGAATTGCAGCGAGCTAATGAGCACCTACGT
meq-W    TATGTAGACAAACTCCATGAAGCATGTGAAGAGCTGCAGAGGGCCAATGAACACCTACGT
meq-O    TATGTAGACAAACTACATGAAGCTTGTGAAGAACTACAAAGAGCAAATGAACACCTACGT
         ** ** ** ** ** ** ** ** ** ** **   * **   * ** ** ** *********


meq-D    AAGGAAATTCGCGATCTTAGAACCGAGTGTACGTCCCTTAGGGTGCAACTCGCATGCCAC
meq-R    AAGGAAATTCGTGATCTCCGTACCGAGTGCACTAGCCTGAGAGTACAACTTGCATGTCAT
meq-W    AAGGAAATTCGAGATCTAAGGACTGAGTGCACGTCCCTGCGTGTACAGTTGGCTTGTCAT
meq-O    AAGGAAATTAGAGATCTACGAACGGAGTGCACCTCTCTTCGTGTTCAGCTGGCCTGCCAT
         ********* * *****   * ** ***** **   ** * ** ** *  * ** ** **


meq-D    GAACCCGTATGCCCTATGGCCGTTCCGCTTACCGTTACTCTCGGACTGCTTACGACTCCG
meq-R    GAGCCCGTTTGCCCCATGGCGGTACCCTTGACTGTTACCCTGGGGCTCCTCACTACGCCC
meq-W    GAGCCAGTTTGCCCTATGGCGGTACCCCTAACGGTGACCCTTGGACTGCTTACCACCCCG
meq-O    GAACCTGTCTGTCCCATGGCTGTTCCTTTGACTGTAACTCTTGGGCTTCTCACCACCCCC
         ** ** ** ** ** ** ***** ** **   * ** ** ** ** ** ** ** ** **


meq-D    CATGATCCCGTTCCGGAACCCCCTATCTGTACTCCGCCACCCCCTTCCCCCGACGAGCCT
meq-R    CACGATCCAGTTCCCGAACCCCCAATCTGCACCCCTCCTCCACCCAGTCCCGATGAGCCC
meq-W    CACGATCCCGTTCCTGAACCTCCCATTTGCACTCCTCCACCTCCCTCACCGGATGAACCT
meq-O    CATGATCCTGTTCCTGAGCCCCCCATCTGCACGCCGCCGCCGCCCTCTCCAGATGAGCCC
         ** ***** ***** ** ** ** ** ** ** ** ** ** **    ** ** ** **


meq-D    AACGCTCCGCATTGTTCCGGGTCCCAGCCCCCTATCTGTACCCCTCCCCCTCCCGATACC
meq-R    AACGCTCCCCATTGCTCTGGATCGCAGCCCCCGATTTGCACGCCTCCTCCACCCGATACC
meq-W    AACGCTCCACATTGCTCCGGTTCCCAACCTCCTATCTGTACCCCCCCTCCTCCCGATACG
meq-O    AATGCTCCCCACTGCAGTGGCTCCCAGCCCCCCATCTGTACTCCTCCTCCTCCAGACACG
         ** ***** ** **       ** ** ** ** ** ** ** ** ** ** ** ** **
```

```
meq-D    GAGGAGCTTTGCGCTCAGTTGTGCTCTACCCCTCCCCCTCCGATTTCGACTCCGCATATT
meq-R    GAGGAACTCTGCGCTCAGCTTTGTTCCACTCCTCCTCCCCCTATTTCCACCCCTCATATC
meq-W    GAGGAACTTTGCGCCCAGCTCTGCTCGACCCCACCACCTCCCATCTCTACTCCCCATATT
meq-O    GAGGAGCTGTGTGCCCAGCTCTGCTCCACCCCACCTCCTCCCATCTCCACCCCCCACATT
         ***** ** ** ** *** * ** ** ** ** ** ** ** ** ** ** ** ** **

meq-D    ATCTATGCGCCAGGCCCATCGCCACTCCAGCCCCCTATCTGTACGCCTCCCCCTCCCGAC
meq-R    ATCTACGCACCTGGGCCGTCGCCTCTTCAGCCCCCCATTTGTACTCCACCTCCGCCTGAC
meq-W    ATCTACGCTCCGGGGCCTTCCCCCCTCCAACCTCCTATCTGTACCCCCCCTCCTCCCGAT
meq-O    ATTTATGCTCCTGGGCCCTCGCCGCTGCAGCCCCCCATCTGCACCCCGCCGCCGCCGGAT
         ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** ** **

meq-D    GCCGAAGAGCTTTGCGCACAGTTGTGTTCGACTCCCCCTCCCCCTATCTGTACGCCACAT
meq-R    GCCGAAGAGCTCTGTGCGCAGCTATGCTCTACCCCCCCACCGCCTATTTGTACCCCTCAT
meq-W    GCGGAGGAGCTTTGCGCCCAGCTCTGCTCGACCCCACCACCTCCCATCTGTACTCCCCAT
meq-O    GCTGAGGAGCTCTGTGCTCAGCTCTGCTCCACCCCACCTCCTCCCATCTGCACCCCCCAC
         ** ** ***** ** ** *** * ** ** ** ** ** ** ** ** ** ** ** **

meq-D    AGTCTGTTTTGCCCACCTCAGCCTCCGTCTCCCGAGGGGATCTTTCCCGCTCTATGCCCC
meq-R    TCCCTGTTCTGTCCTCCACAGCCCCCTTCGCCCGAGGGCATCTTCCCCGCTCTTTGCCCG
meq-W    TCCCTCTTCTGCCCTCCCCAGCCTCCATCTCCGGAGGGCATCTTCCCTGCATTGTGTCCT
meq-O    TCGCTCTTCTGTCCTCCTCAGCCTCCTTCTCCCGAAGGAATTTTTCCTGCTCTTTGTCCT
             ** ** ** ** ** ***** ** ** ** ** ** ** ** ** ** *  * ** **

meq-D    GTTACCGAGCCTTGTACTCCCCCTAGTCCCGGTACCGTATACGCTCAGTTGTGCCCCGTA
meq-R    GTTACGGAGCCCTGCACCCCGCCCTCCCCGGGGACCGTCTATGCTCAGCTTTGTCCAGTT
meq-W    GTTACCGAGCCGTGTACCCCTCCATCGCCGGGGACGGTTTACGCTCAGCTTTGTCCTGTT
meq-O    GTAACGGAGCCCTGCACCCCGCCCTCGCCGGGCACTGTTTATGCTCAGCTTTGTCCTGTT
         ** ** ***** ** ** ** **       ** ** ** ** ** ****** * ** ** **

meq-D    GGCCAGGCGCCACTGTTTACGCCTAGCCCACCGCATCCCGCTCCCGAACCCGAGAGACTT
meq-R    GGCCAGGCTCCTCTCTTCACGCCATCCCCTCCCCATCCAGCTCCCGAACCTGAAAGGCTG
meq-W    GGCCAGGCTCCCCTTTTTACCCCATCTCCCCACATCCGGCTCCGGAGCCGGAGAGGCTT
meq-O    GGGCAGGCCCCGCTCTTCACCCCATCTCCTCCTCATCCTGCTCCAGAGCCAGAGAGGCTT
         ** ***** ** ** ** ** **       ** ** ***** ***** ** ** ** ** **

meq-D    TACGCTAGGCTTACCGAGGATCCCGAACAGGATTCCCTATATTCCGGCCAGATTTATATC
meq-R    TACGCTAGGCTTACGGAGGATCCTGAACAGGACTCTCTTTACTCTGGCCAGATCTATATC
meq-W    TATGCTCGTCTTACCGAGGATCCCGAACAGGATTCCTTGTATTCGGGCCAGATTTATATT
meq-O    TATGCTCGTTTGACTGAAGATCCAGAGCAGGACTCGCTCTACAGTGGGCAGATCTACATT
         ** *** *   * ** ** ***** ** ***** **   * **    ** ***** ** **

meq-D    CAATTCCCTTCCGATACCCAGTCGACCGTTTGGTGGTTCCCAGGCGACGGTAGGCCTTGA
meq-R    CAGTTTCCGTCGGACACCCAGTCTACGGTATGGTGGTTTCCTGGGGACGGCCGTCCGTGA
meq-W    CAGTTTCCCTCGGATACTCAGTCTACGGTCTGGTGGTTTCCAGGTGACGGGAGACCCTGA
meq-O    CAGTTCCCCTCGGACACGCAGTCCACGGTGTGGTGGTTTCCTGGTGATGGCCGTCCCTGA
         ** ** ** ** ** ** ** ***** ** ** ******* ** ** ** ** * ** ***
```

13. List of publications

13.1. Research article

Khedkar PH, Osterrieder N, Kunec D. Codon pair bias deoptimization of the major oncogene meq of a very virulent Marek's disease virus. J Gen Virol. 2018;(Epub ahead of print):doi: 10.1099/jgv.0.001136

13.2. Conference proceedings

Khedkar PH, Osterrieder N, Kunec D. Codon pair deoptimization of a major oncogene as an attenuation strategy for Marek's disease herpesvirus. Poster presented at: 11th International Symposium on Marek's Disease and Avian Herpesviruses, 2016 July 6-9; Tours, France.

Khedkar PH, Osterrieder N, Kunec D. Codon pair deoptimization of a major oncogene as an attenuation strategy for Marek's disease herpesvirus. Poster presented at: 26th Annual Meeting of the Society for Virology, 2016 April 6-9; Münster, Germany.

14. Acknowledgment

15. Selbständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen Anspruch genommen habe.

Berlin, den 27. August 2018

Pratik Hemant Khedkar