



Optimizing a Human Papillomavirus Type 16 L1-Based Chimaeric Gene for Expression in Plants

Inga I. Hitzeroth¹, Aleyo Chabeda¹, Mark P. Whitehead¹, Marcus Graf² and Edward P. Rybicki^{1,3*}

¹ Biopharming Research Unit, Department of Molecular and Cell Biology, University of Cape Town, Rondebosch, South Africa, ² Thermo Fisher Scientific GENEART GmbH, Regensburg, Germany, ³ Faculty of Health Sciences, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Observatory, South Africa

OPEN ACCESS

Edited by:

Johannes Felix Buyel,
Fraunhofer-Institut für
Molekularbiologie und Angewandte
Oekologie IME, Germany

Reviewed by:

Markus Sack,
RWTH Aachen Universität, Germany
Ashokkumar Balasubramaniam,
Madurai Kamaraj University, India
Udo Conrad,
Leibniz-Institut für Pflanzengenetik und
Kulturpflanzenforschung (IPK),
Germany

*Correspondence:

Edward P. Rybicki
ed.rybicki@uct.ac.za

Specialty section:

This article was submitted to
Process and Industrial Biotechnology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 30 April 2018

Accepted: 26 June 2018

Published: 16 July 2018

Citation:

Hitzeroth II, Chabeda A,
Whitehead MP, Graf M and Rybicki EP
(2018) Optimizing a Human
Papillomavirus Type 16 L1-Based
Chimaeric Gene for Expression in
Plants.
Front. Bioeng. Biotechnol. 6:101.
doi: 10.3389/fbioe.2018.00101

Human papillomaviruses (HPVs) are the causative agents of cervical cancer, the fourth most prevalent cancer in women worldwide. The major capsid protein L1 self-assembles into virus-like particles (VLPs), even in the absence of the minor L2 protein: such VLPs have successfully been used as prophylactic vaccines. There remains a need, however, to develop cheaper vaccines that protect against a wider range of HPV types. The use of all or parts of the L2 minor capsid protein can potentially address this issue, as it has sequence regions conserved across several HPV types, which can elicit a wider spectrum of cross-neutralizing antibodies. Production of HPV VLPs in plants is a viable option to reduce costs; the use of a L1/L2 chimera which has previously elicited a cross-protective immune response is an option to broaden cross-protection. The objective of this study was to investigate the effect of codon optimization and of increasing the G+C content of synthetic L1/L2 genes on protein expression in plants. Additionally, we replaced varying portions of the 5' region of the L1 gene with the wild type (*wt*) viral sequence to determine the effect of several negative regulatory elements on expression. We showed that GC-rich genes resulted in a 10-fold increase of mRNA levels and 3-fold higher accumulation of proteins. However, the highest increase of expression was achieved with a high GC-content human codon-optimized gene, which resulted in a 100-fold increase in mRNA levels and 8- to 9-fold increase in protein levels. Changing the 5' end of the L1 gene back to its *wt* sequence decreased mRNA and protein expression. Our results suggest that the negative elements in the 5' end of L1 are inadvertently destroyed by changing the codon usage, which enhances protein expression.

Keywords: HPV-16, L1/L2chimera, plant expression, codon usage, L1, regulatory elements

INTRODUCTION

Cervical cancer is the fourth most common cancer among women worldwide (Ferlay et al., 2015), and the causal association between Human papillomavirus (HPV) infection and cervical cancer has been well described (zur Hausen et al., 1981; zur Hausen, 1996). Three multivalent HPV prophylactic vaccines based on L1 virus-like proteins (VLPs) have been licensed, and are highly effective in the prevention of vaccine-type infections and associated disease (Schiller et al., 2008; Joura et al., 2015; Toh et al., 2015; Signorelli et al., 2017). CervarixTM (GlaxoSmithKline) contains

L1 VLPs from types 16 and 18; Gardasil[®] (Merck & Co., Inc.) contains L1 VLPs of low-risk genital wart types 6 and 11 and high cancer-risk types 16 and 18; and Gardasil[®] 9 (Merck & Co., Inc.), the most recently approved nonavalent vaccine, targets HPV-6/11/16/18, and an additional five HPV types (HPV-31/33/45/52/58). Despite the success of these vaccines, the cervical cancer burden remains high. Additionally, these vaccines are expensive and show type-restrictive prophylactic efficacy.

To address the need for vaccines that will protect against more than one HPV type, the L2 minor capsid protein has been investigated as a vaccine candidate (Karanam et al., 2009). Sequences within the HPV-16 L2 N-terminal region, especially amino acids (aa) 1–120, contain broadly cross-neutralizing epitopes that can neutralize a broad range of mucosal and cutaneous HPVs (Pastrana et al., 2005; Alphas et al., 2008). However, in the context of the native virion this L2 sequence is not prominently displayed: while residues 1–88 are accessible to antibodies during infection, there are only 12–36 molecules per natural virion (Guan et al., 2017). Additionally, although the L2 N-terminal region contains broadly cross-neutralizing epitopes, L2 is also subdominant to L1 as an antigen, and the use of L1+L2 VLPs in vaccination does not confer more cross-protection in animals compared to use of L1 VLPs only (Roden et al., 2000). Accordingly, increasing the density of these sequences at the surface of the particles in a different context than the immunodominant L1 pentamer would almost certainly significantly increase their immunogenicity.

The HPV-16 L2 residues 108–120 (LVEETSFDIAGAP; L₂_{108–120}) form a common neutralizing epitope across HPV-16, – 6, – 11, and – 18 (Kawana et al., 2001). We have previously used this epitope in a number of chimaeric constructs for the production of a potential multivalent vaccine candidate, with the L₂_{108–120} replacing L1 residues in a number of selected surface-exposed regions that do not impact L1 conformational epitopes, in both insect cells and in plants (Varsani et al., 2003a; McGrath et al., 2013; Pineo et al., 2013). A consistent finding was that replacement of part of the h4 helix with the L2 peptide resulted in the best immunogenicity.

Plant expression systems present a cost-effective alternative to conventional vaccine production due to their scalability, rapid production and low risk of contamination (Fischer et al., 2004; Rybicki, 2010; Merlin et al., 2014). Estimates suggest that generic production cost of goods could be reduced by 50% compared to conventional production (Nandi et al., 2016). Various HPV proteins have been expressed in plants (reviewed in Rybicki, 2014). Plant-derived HPV L1 proteins self-assemble into higher-order structures, and are both immunogenic and show protective efficacy as vaccines in animal models (Kohl et al., 2006). Transient plant expression systems are particularly useful for the rapid production of antigens (Rybicki, 2010), and significantly higher L1 protein levels have been obtained in comparison to stable nuclear transformation (Varsani et al., 2003b; Giorgi et al., 2010). Low yields of recombinant protein have often been reported using plant expression systems; however, methods such as the use of strong plant promoters (Twyman et al., 2003; Obembe et al., 2011), codon optimization (Biemelt et al., 2003; Maclean et al., 2007), co-expression of silencing suppressors (Takeda et al., 2002;

Voinnet et al., 2003), and subcellular targeting have been used to increase protein yields (Maclean et al., 2007; Twyman et al., 2013).

Our previous investigation of HPV-16 L1 transient expression in plants determined that use of a human codon-optimized gene, and targeting protein accumulation to the chloroplast rather than the cytosol, ER, or apoplast, resulted in highest accumulation of HPV-16 L1 protein; in contrast, a plant codon-optimized gene was not expressed at meaningful levels (Maclean et al., 2007). Similar results for independently made plant and human codon-optimized HPV-16 L1 genes were obtained by Biemelt et al. (2003). Given the counter-intuitive results in both investigations, it is clear that further insights into optimization of protein expression are necessary.

In our previous work, the *wt* HPV-16 L1 had a GC content of 38% and the plant codon-optimized gene 35%, whereas the human codon-optimized L1 (*hL1*) had a GC content of 63%. This led us to hypothesize that expression levels in plants are determined at least in part by GC content. This is similar to the situation for mammalian cells, where GC-rich genes are better expressed than AT-rich genes (Kudla et al., 2006). Another factor affecting mammalian cell expression of HPV L1 genes specifically are negative control elements acting at the transcriptional level (Collier et al., 2002; Rollman et al., 2004; Johansson and Schwartz, 2013). It is thought that these negative elements control L1 expression during epithelial cell differentiation, as it is only produced at later stages of the virus life cycle and at a late stage in differentiation. Codon optimization changes the sequence of these elements and thereby significantly impacts their functions.

We report here on the use of *Agrobacterium*-mediated transient expression in *Nicotiana benthamiana* of seven HPV-16-derived genes—six synthetic and one *wt* L1/L₂_{108–120}-encoding the single L1/L2 chimera of interest as a candidate vaccine, in order to investigate the impact of codon alteration and overall GC content on the accumulation of the protein. We investigated if differences in expression were at the transcriptional level, as well as exploring whether destruction of known negative regulatory elements are involved in determining protein expression level, by replacing parts of the 5' region of the L1/L2 chimera gene with *wt* L1 DNA sequence.

METHODS

Synthesis of the L1/L2 and wt/L1/L2 Chimeras

The L1/L2 chimaeric gene *ChiF/SAF* (Genbank number: AY177679) (Varsani et al., 2003a; McGrath et al., 2013) was used as a starting point for sequence modification (Figure 1A). Sequences were generated using GeneOptimizer[®] (Life Technologies, USA), a multi-parameter gene optimization software tool which allowed a balance of codon choice (human or tobacco), and of GC or CpG dinucleotide content. Briefly, by using a sliding combination window and applying different emphasis to certain gene optimization parameters (in this case preferred dicot codon usage and a certain GC content), the described algorithm allowed us to identify DNA sequences

showing the best balance between a given GC content and a preferred dicot codon choice, as assessed by the software. In the case of the back-translated (BT) *SAF2* sequence, only tobacco-preferred codons were used for back-translating the amino-acid sequence. The resulting sequences were then assembled from oligonucleotides, cloned and sequence verified (GeneArt, Regensburg, Germany) (Table 1).

To investigate if the increase in mRNA and protein levels was due to removal of negative elements found in the 5' end of the HPV16 *L1* DNA, 5 chimeras were created where the 5' end of the gene was replaced with *wt* sequence. In these *wt/L1/L2* chimeras the following sequences were replaced using assembly PCR: 1–66, 1–147, 1–251, 1–429, and 1–620; they were called 66, 147, 251, 429, and 620 (Figure 1C). The *wt/L1/L2* chimeras were created using *wt L1* as template for the 5' PCR and the 3' end was created using human codon-optimized *L1/L2* as template using primers listed in Table 2. The middle primers had overlapping sequencing to allow amplification of whole gene in a second PCR reaction.

Subcloning of Genes Into Plant Expression Vectors

Two binary *Agrobacterium* non-replicative plant expression vectors were used to compare HPV chimera expression: these were pTRAc, which targets the expressed protein to the cytoplasm, and pTRAc-*rbcS1-cTP* which targets the protein to the stroma in chloroplasts via the chloroplast-transit peptide sequence of the potato *rbcS1* gene (where *rbcS1* is the ribulose biphosphate carboxylase small chain 1) (vectors kindly provided by Prof. Rainer Fischer, Fraunhofer Institute for Molecular Biology and Applied Ecology, Germany) (Maclean et al., 2007). The genes were excised with 5' *Bsp*HI or *Mlu*I and 3' *Xho*I and directionally cloned into *Afl*III, *Xho*I sites in pTRAc or *Mlu*I and *Xho*I in the chloroplast targeting vector.

DH5- α chemically competent *E. coli* cells (E.cloniTM, Lucigen, USA) were transformed with the plasmid constructs and recombinants selected on ampicillin plates (100 μ g/mL). Recombinant clones were screened by colony PCR, using pTRA vector-specific primers (Fwd pTRAc Primer 5'-CATTTCATTGGAGAGGACACG-3' and RVS pTRAc Primer 5'-GAACTACTCACATTATTCTGG-3') and recombinant genes were verified by pyrosequencing.

Agrobacterium-Mediated Transient Expression

Agrobacterium-mediated protein expression was performed as described by Maclean et al. (2007). Electrocompetent *A. tumefaciens* GV3101::pMP90RK were transformed with pTRA constructs, plated on 50 μ g/mL carbenicillin, 50 μ g/mL rifampicin, and 30 μ g/mL kanamycin plates and successful transformation confirmed by colony PCR. Chimeras were co-expressed with or without *Agrobacterium* LBA4404 (pBIN-NSs) containing the NSs silencing suppressor gene of Tomato spotted wilt virus (TSWV) (Takeda et al., 2002). Recombinant *Agrobacterium* were grown in induction medium and the *Agrobacterium* suspension was either injected- (small scale—a few leaves) or vacuum-infiltrated (large scale—whole plants) into

the abaxial air spaces of 6–8 week old *N. benthamiana* leaves. The plants were grown at 22°C under 16/8 h light/dark cycles and samples harvested 1–10 days post-infiltration (dpi).

Extraction of Protein and RNA From Plants

To screen leaf tissue for protein expression, five leaf discs (5 mm diameter, ~0.05 g wet plant mass) were ground in liquid nitrogen, and incubated in 650 μ l of high-salt phosphate buffer (0.5 M NaCl). The supernatant was clarified by centrifugation for 20 min (13,000 rpm, desktop centrifuge, 4°C) and L1/L2 protein was detected by western blotting. For RNA extractions RNeasy Plant Mini Kit (Qiagen, Netherlands) was used and RNA was extracted as per manufacturer's instructions. Four leaf discs were homogenized in 450 μ L extraction buffer, samples were clarified, and eluted into 30 μ L. 1/1,000 dilution was used in qRT/PCR.

ELISA Quantification of L1, L1/L2, and wt/L1/L2 Chimera Yields

Plant-expressed L1/L2 chimeras were quantified by capture ELISA using a modified polyvinyl alcohol (PVA)-blocking ELISA method (Studentsov et al., 2002), as described in Pineo et al. (2013). Briefly, a 96-well Nunc Maxisorp microtitre plate (Thermo Fisher, USA) was coated with 1:2,000 CamVir1 (Abcam, UK; a mouse anti HPV-16 L1 MAb) overnight at 4°C, washed and blocked with PVA buffer. Diluted plant cell extract was added to the wells and incubated for 1 h at 37°C, followed by a washing step and the addition of rabbit anti-HPV-16 polyclonal serum (1:1,000) overnight at 4°C. After washing, swine anti-rabbit horseradish peroxidase (HRP) conjugate (1:5,000; DAKO, Denmark) was added to wells, plates were incubation for 30 min at 37°C and the proteins were detected with OPD substrate (DAKO). Plates were developed in the dark, the reaction was stopped with 0.5M H₂SO₄ and the absorbance was detected at 490 nm. Total soluble protein (TSP) was determined using a Lowry protein assay (BioRad, USA) as per the manufacturer's instructions, with a bovine plasma IgG standard (BioRad, USA) to normalize the ELISA data.

Protein Detection and Quantitation by Western Blot

Samples were incubated at 95°C for 5 min in loading buffer, separated on 10% SDS-PAGE, then either stained with Coomassie blue or transferred onto a nitrocellulose membrane using the Trans-Blot[®] SD Semi-Dry Transfer Cell (Bio-rad, USA) for western blot analysis. L1 proteins were detected CamVir1 (1:10,000; Abcam, UK), a monoclonal antibody against HPV16 (<http://www.abcam.com/hpv16-l1-antibody-camvir-1-ab69.html>). MAbs were detected with secondary goat-anti-mouse-alkaline phosphatase conjugate (1:10,000; Sigma Aldrich, USA) and blots developed with NBT/BCIP substrate (Roche, Switzerland). Proteins were measured by semiquantitative analysis by measuring the density of the band on a western blot or Coomassie stained bands in comparison to a known protein concentration standard and purified L1 that had previously been quantitated (Pineo et al., 2013), using GeneTools software (SYNGENE, UK) on scanned images.

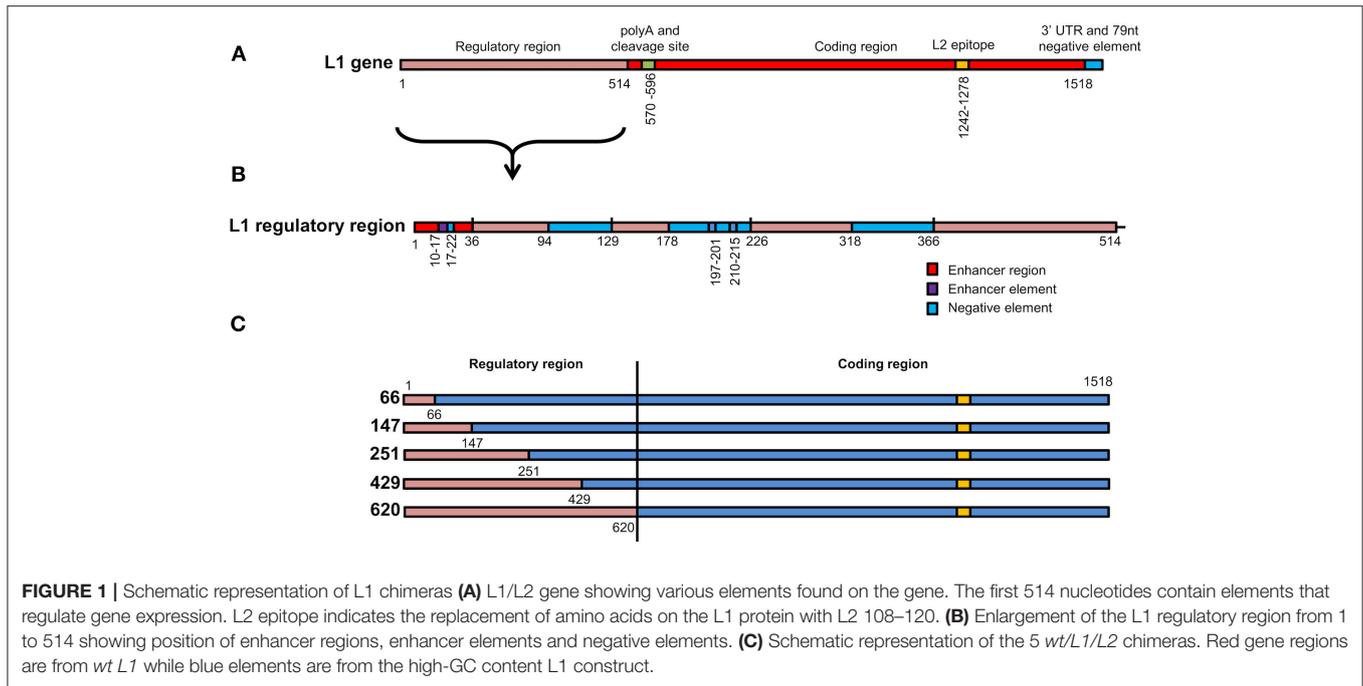


TABLE 1 | Summary of genes with varying GC content.

Name	GC content	Details	L1/L2 _{108–120} Chimera
<i>Chi F</i>	38.27% (WT)	Wild type gene – HPV codon usage	Yes
<i>SAF 2</i>	34.64% (BT)	Tobacco codon optimization (TCO)	Yes
<i>SAF 3</i>	50.88%	TCO with elevated GC content	Yes
<i>SAF 4</i>	61.01%	TCO with elevated GC content	Yes
<i>SAF 5</i>	43.06% (CpG)	TCO with elevated CpG dinucleotides	Yes
<i>SAF</i>	62.52%	Human codon optimized	Yes
<i>hL1</i>	62.20% (hL1)	Human codon optimized L1 gene	No

Quantitative RT-PCR

RT-PCR reactions were performed on RNA extracted from leaves using a SensiMix One-Step real-time RT-PCR kit (Quantace, UK) and a Rotor-Gene RG-3000A real-time PCR machine (Qiagen). As a positive control L1 RNA was *in vitro*-transcribed using Ribomax™ Large Scale RNA Production System—T7 kit (Promega, USA). RT-PCR reactions (25 μL) contained *in vitro*-transcribed L1 RNA or 10 μL of the RNA extracted from leaves, 50 mM MgCl₂ and 50 pmol of forward and reverse primers to amplify a 100 bp DNA fragment. The forward and reverse primer sequences are given in **Table 2**. The reaction profile used was as follows: 49°C, 30 min; 95°C, 10 min; 45 cycles of 95°C, 15 s, 54°C, 15 s, and 72°C, 15 s. All real-time RT-PCR data was analyzed using the Rotor-gene 6, Version 6.0 (Build 27) software (Corbett Research).

RESULTS

Modification of L1/L2 Chimera GC Content

To determine the impact of GC content on protein expression and mRNA stability, HPV 16 L1/L2_{108–120} chimera coding was

changed to give genes with varying GC contents (**Table 1**). The *wt ChiF* has 38% GC content and the native HPV codon usage was preserved; in the *SAF2* back-translated gene the GC content was kept similar to the *wt* type gene, but tobacco (*N. benthamiana*) preferred codons were used. In *SAF3* and *SAF4* the GC content was 50 and 60% respectively, and the codons used were the ones preferred by tobacco plants. *SAF* was made using human codon usage, with 63% GC content. Additionally, *SAF5* was made with 43% GC content and elevated CpG dinucleotide content with the idea of determining the possible effect of methylation on protein expression. The HPV-16 L1 gene was human codon optimized for a gene with 62% GC content (*hL1*).

To investigate whether the change in mRNA and protein levels was due to negative elements found in the 5' end of the HPV-16 L1 DNA (**Figure 1B**), 5 chimeras were created from *SAF* with the first 66, 147, 251, 429, or 620 5' nucleotides replaced with *wt L1* sequence from *ChiF*; chimeras were called *wt/L1/L2*. Schematic design of the 5 *wt/L1/L2* chimeras is shown in **Figure 1C**. These genes were also expressed in plants and mRNA and protein levels measured.

TABLE 2 | Primers used in plasmid construction.

Primer name	Primer sequence	5' Primer binding site	Product size (kb)
ChiF F	5'-ACTGCAGACGTTATGACATAC-3'	1,170	99
ChiF R	5'-TTCCACTAATGTGCCTCCTG-3'	1,229	
SAF-2 F	5'-ACCAGGTGGAACCTCTTGTG-3'	1,298	97
SAF-2 R	5'-AGTGGATCTTCCTTTGGAGC-3'	1,355	
SAF-3 F	5'-CGATCTCCAGTTCATCTTCC-3'	1,181	100
SAF3 R	5'-TTCCAATCCTCAAGGATAGTG-3'	1,240	
SAF-4 F	5'-AAGAGTACGACCTCCAGTTC-3'	1,195	96
SAF-4 R	5'-AAGGATGGTGGAGTTCATGG-3'	1,251	
SAF-5 F	5'-TACTCTTCAAGCTAATAAGTCC-3'	727	105
SAF-5 R	5'-AAGAATCGCCGTATGGTTCG-3'	790	
SAF F	5'-ACTTCAAGGAGTACCTGAGG-3'	1,152	103
SAF R	5'-TGTGGATGTAGGTCATCACG-3'	1,215	
SAF F66	5'-AGTATCTAAGGTTGTAAGCACCGATGAGTACGTGG-3'		
Chi F R66	5'-ACGTACTCATCGGTGCTTACAACCTTAGATACTGG-3'		
SAF F147	5'-AGTTGGACATCCCTATTTCCCATCAAGAAGC-3'		
ChiFR147	5'-CTTCTTGATGGGGAATAGGGATGTCCAACCTGC-3'		
SAF F251	5'-TGACCCCAATAAGTTTGGCTTCCCGACACCAGC-3'		
ChiF R251	5'-TGGTGTCCGGGAAGCCAAACTTATTGGGGTCAGG-3'		
SAF F429	5'-CAGGTGTGGATAACAGAGAATGCATCAG-3'		
ChiF R429	5'-CTGATGCATTCTCTGTTATCCACACTGCATTTGC-3'		
SAF F620	5'-GTTGATACTGGCTTGGTCCATGGACTTACCACC-3'		
ChiF R 620	5'-GTGGTGAAGTCCATGGCACCAAGCCAGTATCAACC-3'		

Comparative Expression of *hL1* and the L1/L2 Chimaeric Genes in Plants

L1/L2 chimeras and hL1 were transiently expressed in 8-week old *N. benthamiana* plants and proteins were extracted at 3, 5, 7, and 10 dpi. Accumulation levels of L1/L2 and hL1 protein targeted to the cytoplasm or chloroplast with or without co-expression of the silencing suppressor NSs, were compared by western blots at 5 dpi (Figure 2A) and quantitative ELISA (Figure 2B). Clear bands in western blots below the expected L1 chimera size of 55 kDa are almost certainly products of proteolysis, which we have reported previously for plant-produced HPV-16 and HPV-11 L1s (Varsani et al., 2003b; Kohl et al., 2007). All experiments were repeated at least 3 times and ELISA results are presented as mean averages. Co-infiltration of plants with NSs did not enhance protein accumulation at 5 dpi. Targeting the protein to the chloroplast increased protein accumulation levels 3 to 10-fold relative to cytoplasmic levels (Figure 2B). Higher accumulation levels were seen for all the chimeras with GC contents of 50% and above when targeted either to the cytoplasm or to the chloroplast. The yield of L1/L2 protein ranged from 1.4 mg/kg fresh weight for ChiF to 128 mg/kg for hL1 when targeted to the chloroplast, whereas expression levels in the cytoplasm were 1.2 and 4.8 mg/kg for ChiF and hL1 respectively. L1 accumulated to between 0.1 and 0.9% of TSP.

Protein expression levels also varied with the GC content of the chimeras: the *wt ChiF* and the back-translated *SAF2*, which had GC contents of 38 and 35% respectively, hardly expressed at all. The *SAF5* gene with 43% GC content and high levels of

CpG was also minimally expressed. *SAF3* with 50% GC content produced L1/L2 levels up to 53 mg/kg; *SAF4* with 60% GC content expressed up to 85 mg/kg. The two genes that used mammalian codon preference had the highest accumulation levels: *SAF* with 63% GC content expressed protein up to 100 mg/kg and *hL1*—HPV-16 L1 with a GC content of 62%—expressed the highest level of any construct, with up to 128 mg/kg (Figure 2B).

Comparative mRNA Levels in Plants

To investigate whether the difference in protein expression levels were due to translation or transcription, as it has been postulated that the RNA secondary structure might have an effect on translation and RNA half-life, we used the ratio of specific mRNA levels to plant 18S RNA as determined by quantitative RT-PCR to normalize for extraction variation. Addition of the NSs did not increase mRNA levels (data not shown), while expression from constructs targeting the protein to the chloroplast increased mRNA levels 2.5–8.8 times compared to targeting to the cytoplasm (Figure 2C). This increase could possibly come from changes in secondary structure of the mRNA with the chloroplast targeting peptide. In plants infiltrated with the L1/L2 the genes with GC content of 60% and above, targeting the resulting protein to the chloroplast showed a 5 to 100-fold increase in mRNA levels. Interestingly both the back-translated *SAF2* gene (35% GC) and *SAF5* with the high CpG level expressed higher mRNA levels than *SAF3* with 50% GC content. The latter produced low mRNA levels, but protein levels were high when

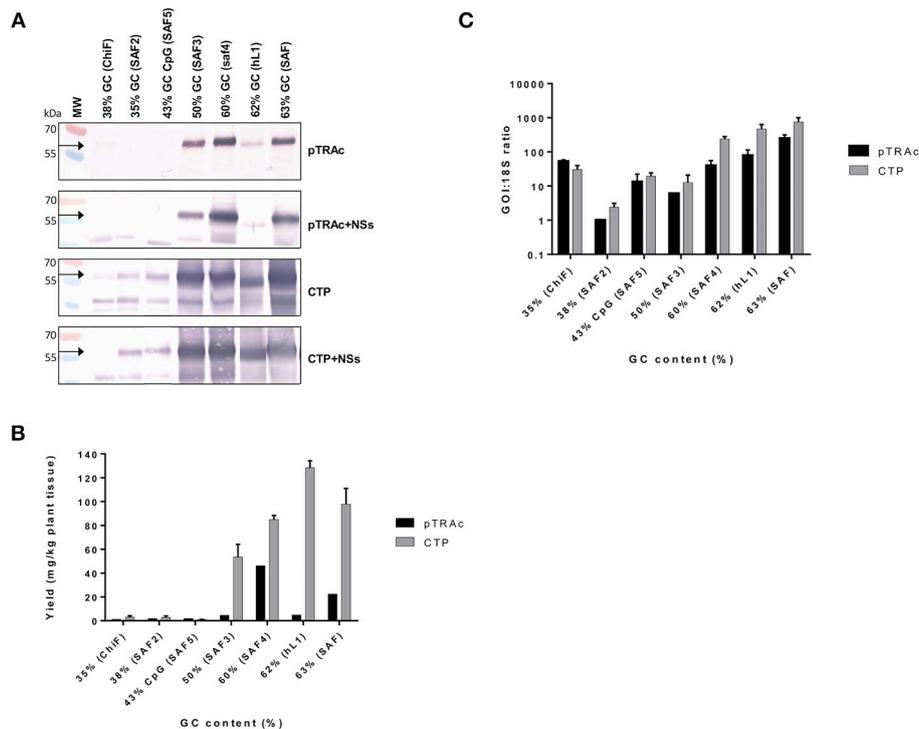


FIGURE 2 | Comparative expression of L1/L2 and hL1 with different GC contents using pTRAc and pTRAc rbcS1-CTP (CTP). Protein was expressed with or without NSs and extracted 5dpi. **(A)** Western blot detection of L1/L2 and hL1. MW, molecular weight marker, arrow indicates position of L1 and L1/L2 proteins. Equal volumes of sample were loaded. **(B)** Protein yield as mg/kg fresh weight determined by capture ELISA. Results represent 3 biological repeats, error bars indicate \pm SEM. **(C)** Comparison of RNA transcription levels of *L1* and chimeras targeted to different subcellular locations. mRNA levels were measured as ratio of 18S RNA to normalize for extraction variation.

targeted to the chloroplast. This could confirm the theory that the protein is protected from degradation by targeting it to the chloroplast. Overall the mRNA levels did not show the same pattern as protein accumulation.

Creation and Comparative mRNA and Protein Levels of *wt/L1/L2* Chimeras in Plants

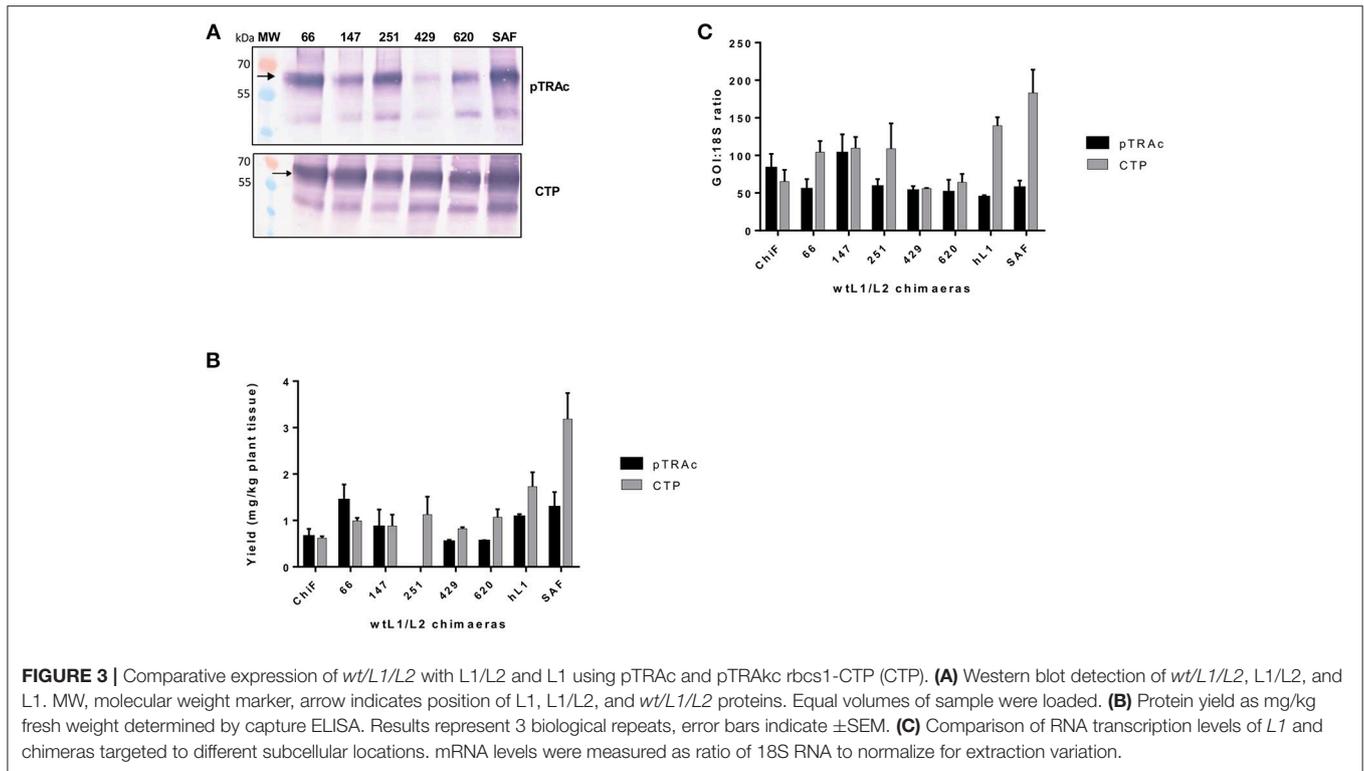
Previous studies have shown that negative elements in the 5' end of HPV-16 *L1* DNA were responsible for reduced expression of the gene in mammalian and insect cells (Figure 1B). To investigate the influence of these negative elements on protein expression, 5 chimeras were created from *SAF* with the first 66, 147, 251, 429, or 620 5' nucleotides being replaced with *wt L1* sequence from *ChiF*; chimeras were called *wt/L1/L2*. Schematic design of the 5 *wt/L1/L2* chimeras is shown in Figure 1C. These genes were also expressed in plants and mRNA and protein levels measured. Protein accumulation was assessed by western blot and ELISA (Figures 3A,B). Interestingly, when the proteins were targeted to the chloroplast, the variation in *wt/L1/L2* accumulation was similar for the first three *wt/L1/L2* chimeras (66, 147, and 251). When *wt/L1/L2* was targeted to the cytoplasm, however, constructs 429 and 620 showed markedly less protein accumulation, which is evident in the western blot and ELISA

(Figures 3A,B). In summary, there was not a profound effect on protein accumulation by replacing the 5' regulatory regions of the *L1* gene with *wt* sequence but replacing the protein from nt 429 onwards results in decrease of expression.

Comparing the wildtype *ChiF* mRNA levels to the mRNA levels of *wt/L1/L2* chimeras, it was interesting to note that the two chimeras 429 and 620 had similar mRNA levels when proteins were targeted to the chloroplast and cytoplasm (Figure 3C). Overall mRNA levels were 1.6 times lower for these two chimeras and protein expression was also lower than all the other chimeras again showing a region of the *L1* that has an influence on protein expression levels.

DISCUSSION

In previous work we and others determined that using human codon optimization—which increased the GC content of the HPV *L1* gene to above 60%—increased HPV-16 *L1* protein expression in plants (Biemelt et al., 2003; Maclean et al., 2007). Here we wished to systematically explore the effect of GC content on gene expression in plants by creating genes encoding the same HPV *L1/L2* chimera with 6 different GC contents. Additionally, as there is an enormous body of work that describes that negative elements in the *L1* gene affect protein expression in the natural



cycle of the virus (Schwartz, 1998; Collier et al., 2002; Rollman et al., 2004; Zhao and Schwartz, 2008), we further wished to investigate if the HPV L1 negative elements also function in plants to limit L1 expression.

The *wt ChiF* has 38% GC content and the native HPV codon usage was preserved. If the most preferred dicot codons are used to back translate any given gene this results in sequence with a very low GC content, usually significantly below 45%. *SAF2* is the back-translated gene where tobacco (*N. benthamiana*) preferred codons were used, resulting in a gene with 35% GC content. In order to be able to balance the use of preferred dicot codons and step by step increasing the GC content of the resulting DNA sequence, the multi-parameter gene optimization approach described in detail by Raab et al. was used (Raab et al., 2010). By using this sliding combination window and applying different emphasis to certain gene optimization parameters (in this case preferred dicot codon usage and a certain GC content), optimal DNA sequences showing the best balance between a given GC content and a preferred dicot codon choice were created. *SAF3* and *SAF4* were created in this way, resulting in genes with GC content of 50% and 60% respectively, but using codons preferred by tobacco plants. The positive control genes *SAF* and HPV-16 L1 (*hL1*) were made using human codon usage, resulting in genes with 63 and 62% GC content, respectively.

Lastly, *SAF5* was created which had a 43% GC content and elevated CpG dinucleotide content. This was done as previously researchers have shown that aggregation of CpG dinucleotides positively influenced expression in mammalian cells. We therefore speculated that a similar mechanism might

positively increase expression in plants as well (Bauer et al., 2010; Krinner et al., 2014). While this was plausible as a hypothesis, the resolution of the experimental approach was not sufficient to infer a putative effect of CpG islands.

Transient Expression of hL1, L1/L2, and *wt/L1/L2* in Plants

In the present study we confirmed our previous results that L1-based protein accumulation was enhanced when the protein was targeted to the chloroplast: this was true for the L1 protein, and the L1/L2 chimera. We further determined that human codon-optimized genes with the highest GC content showed the highest expression levels. A plant codon-optimized gene with a similar GC content of 60% showed approximately 66% of the expression level of the hL1. In previous codon-optimization studies on papillomavirus gene expression in mammalian cells, L1, E5, and E7 genes all showed increased mRNA translation efficiency and thus higher protein expression when human codon-optimized (Leder et al., 2001; Liu et al., 2002; Disbrow et al., 2003). When a HPV-16 L1-E7 hybrid gene was created for DNA vaccination, there was increased expression and thus immune response when the gene was human codon-optimized (Cheung et al., 2004). Further, Collier et al. (2002) created a mutant L1 cDNA which was codon-optimized to increase the GC content of the AU-rich HPV-16 genome: they demonstrated high levels of both L1 mRNA and protein (Collier et al., 2002).

In this study we also measured mRNA levels in plants infiltrated with the genes with varied GC content. We observed a 170-fold increase in mRNA levels when comparing the *wt* gene

to the human codon-optimized genes. The *SAF3* plant codon-optimized gene with GC content of 50% showed only a 10-fold increase in mRNA levels; the *SAF4* gene with 60% GC content showed a 50-fold increase (Figure 2C). Protein expression levels also increased with increasing GC content (Figures 2A,B). This confirms that changing the codon composition of *L1* from an AT-rich to GC-rich results in increase of accumulation of *L1* mRNA, and that the effect on protein expression is thus mainly at the transcriptional level. Interestingly, we also found an increase in mRNA in the proteins targeted to the chloroplast. This was a bit surprising as it is thought that targeting the protein to the chloroplast could potentially increase accumulation by sequestering the protein away and thereby protecting it from degradation (Maclean et al., 2007). Our results indicate that the mRNA of the chloroplast targeted proteins is more stable, which might be inadvertently due to the addition of the signal sequence coding region giving the mRNA more stability.

The increase in mRNA levels may be due to increase in *L1* mRNA half-life (Mori et al., 2006); additionally, poor expression of the *wt* gene could be due to the presence of inhibitory RNA elements. *L1* late gene expression is strongly influenced by post-transcriptional gene regulation (Zheng and Baker, 2006); therefore viral elements and cellular RNA binding factors are important in regulating the HPV-16 genes (Mole et al., 2006). It is thought that codon-optimization affects RNA inhibitory sequences and thus influences the relevant binding protein interactions. Therefore, it could be argued that the increased translation efficiency is solely due to disruption of these protein binding sites on the mRNAs with little effect actually coming from the codon optimization (Zhao and Schwartz, 2008).

Mapping of Negative Elements on *L1* Gene

The first 514 nucleotides (nt) of the *wt L1* coding region contain multiple RNA sequences that inhibit gene expression (Figures 1A,B). These sequences reduce mRNA levels as well as inhibiting translation. Some sequences are thought to be splicing silencers elements and binding of heterogeneous nuclear ribonucleoprotein (hnRNP) A1 to these sites inhibits transcription (Schwartz, 2013). As we wanted to investigate if gene expression in plants was enhanced due to change in some of the negative elements described, we created 5 *wt/L1/L2* chimeras where we changed the 5' end of the human codon optimized *L1/L2* gene back to the *wt* sequence (Figure 1C). Changing the first 147 nt or more in the *L1* gene back to its *wt* sequence decreased protein accumulation 2-fold (Figure 3A). Overall, we found that changing small parts of the 5' end of the *L1* gene decreased protein accumulation in plants. This suggests that these negative sequences on the *L1* gene may also be effective in plants. This was surprising as the *L1* 5' and 3' UTR regions dramatically influence expression levels, but as the 5' and 3' UTR were unchanged in our constructs, there are probably elements on *L1* influencing expression levels and these are also effective in plant cells. While additional experiments designed to determine whether increased translation efficiency is solely due to disruption of negative regulatory protein binding sites on the mRNAs by codon optimization would be useful, unfortunately this did not fall into the scope of this work.

To determine if post-transcriptional gene silencing played a role in *L1* mRNA levels and thus the expressed protein levels, in all experiments the silencing suppressor protein NSs from the TSWV was expressed together with the *L1* proteins. NSs is thought to interfere with steps that generate the dsRNA, thereby inhibiting silencing (Takeda et al., 2002). However, we did not observe an increase in protein or mRNA levels, indicating that gene silencing did not play a role in controlling protein expression in this system. In a recent study by Jackson et al., codon optimization, RNA instability motifs, blocking of sRNA binding sites and randomization of non-rare codons were investigated to provide rules for efficient transgene expression in plants and reduce gene silencing (Jackson et al., 2014). The authors showed that eliminated on such motifs yielded up to 935-fold increase in gene expression. Therefore, these additional rules could be used to increase gene expression of HPV-*L1* in plants.

Overall, we determined that constructs based on the main HPV-16 capsid protein gene *L1* were not well expressed in plants if the *wt* nucleotide sequence was used. We showed that mRNA levels increased 10-fold when the GC content was increased from 35 to 43% or higher, but protein expression only increased once the GC content was above 50%. Changing the first 620 nt on the 5' end of *L1* back to *wt* sequence showed a 50% reduction of protein expression. This indicates that the negative elements found in the 5' end of the *L1* gene do play a significant role in *L1* protein expression. We were not able to elucidate the effect of each element, but determined that this region plays an important role in protein expression in plants. While it is clear that in this study changing codon usage quite dramatically affected expression levels of our chosen gene, it cannot be claimed that this is the only reason for the differences. For instance, mRNA splicing and nuclear export were not studied, and therefore we cannot generally exclude that aberrant splicing, inefficient transcript processing or nuclear export also affected the results.

This work shows that optimization of protein expression in plants remains a challenge, but that the challenge is surmountable. In general, our research group has found that GC content above 50% is necessary for adequate protein expression of HPV and other proteins, such as H5N1 influenza virus haemagglutinin, in plants (Maclean et al., 2007; Pereira et al., 2009; Mortimer et al., 2012; Pineo et al., 2013). Our findings are also of significant interest in the continuing search for a second generation of affordable HPV vaccines.

AUTHOR CONTRIBUTIONS

IH and AC wrote the manuscript. MW performed the experiments, AC compiled figures. MG designed and synthesized the genes. IH and ER designed and conceptualized the study.

ACKNOWLEDGMENTS

We thank Paul Kennedy for his invaluable help in writing this document and Prof. Rainer Fischer for pTRA cloning vectors. This work is based on research supported in part

by the National Research Foundation of South Africa. MW was supported with funding from the Beit Trust, Carnegie Corporation of New York, the Poliomyelitis Research

Foundation (PRF) and by UCT postgraduate funding. AC was supported with funding from the Carnegie Corporation of New York.

REFERENCES

- Alphs, H. H., Gambhira, R., Karanam, B., Roberts, J. N., Jagu, S., Schiller, J. T., et al. (2008). Protection against heterologous human papillomavirus challenge by a synthetic lipopeptide vaccine containing a broadly cross-neutralizing epitope of L2. *Proc. Natl. Acad. Sci. U.S.A.* 105, 5850–5855. doi: 10.1073/pnas.0800868105
- Bauer, A. P., Leikam, D., Krinner, S., Notka, F., Ludwig, C., Langst, G., et al. (2010). The impact of intragenic CpG content on gene expression. *Nucleic Acids Res.* 38, 3891–3908. doi: 10.1093/nar/gkq115
- Biemelt, S., Sonnwald, U., Galmbacher, P., Willmitzer, L., and Muller, M. (2003). Production of human papillomavirus type 16 virus-like particles in transgenic plants. *J. Virol.* 77, 9211–9220. doi: 10.1128/JVI.77.17.9211-9220.2003
- Cheung, Y. K., Cheng, S. C., Sin, F. W., and Xie, Y. (2004). Plasmid encoding papillomavirus Type 16 (HPV16) DNA constructed with codon optimization improved the immunogenicity against HPV infection. *Vaccine* 23, 629–638. doi: 10.1016/j.vaccine.2004.07.010
- Collier, B., Oberg, D., Zhao, X., and Schwartz, S. (2002). Specific inactivation of inhibitory sequences in the 5' end of the human papillomavirus type 16 L1 open reading frame results in production of high levels of L1 protein in human epithelial cells. *J. Virol.* 76, 2739–2752. doi: 10.1128/JVI.76.6.2739-2752.2002
- Disbrow, G. L., Sunitha, I., Baker, C. C., Hanover, J., and Schlegel, R. (2003). Codon optimization of the HPV-16 E5 gene enhances protein expression. *Virology* 311, 105–114. doi: 10.1016/S0042-6822(03)00129-6
- Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., et al. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* 136, E359–E386. doi: 10.1002/ijc.29210
- Fischer, R., Stoger, E., Schillberg, S., Christou, P., and Twyman, R. M. (2004). Plant-based production of biopharmaceuticals. *Curr. Opin. Plant Biol.* 7, 152–158. doi: 10.1016/j.pbi.2004.01.007
- Giorgi, C., Franconi, R., and Rybicki, E. P. (2010). Human papillomavirus vaccines in plants. *Expert Rev. Vaccines* 9, 913–924. doi: 10.1586/erv.10.84
- Guan, J., Bywaters, S. M., Brendle, S. A., Ashley, R. E., Makhov, A. M., Conway, J. F., et al. (2017). Cryoelectron microscopy maps of human papillomavirus 16 reveal L2 densities and heparin binding site. *Structure* 25, 253–263. doi: 10.1016/j.str.2016.12.001
- Jackson, M. A., Sternes, P. R., Mudge, S. R., Graham, M. W., and Birch, R. G. (2014). Design rules for efficient transgene expression in plants. *Plant Biotechnol. J.* 12, 925–933. doi: 10.1111/pbi.12197
- Johansson, C., and Schwartz, S. (2013). Regulation of human papillomavirus gene expression by splicing and polyadenylation. *Nat. Rev. Microbiol.* 11, 239–251. doi: 10.1038/nrmicro2984
- Joura, E. A., Giuliano, A. R., Iversen, O.-E., Bouchard, C., Mao, C., Mehlsen, J., et al. (2015). A 9-valent HPV vaccine against infection and intraepithelial neoplasia in women. *N. Engl. J. Med.* 372, 711–723. doi: 10.1056/NEJMoa1405044
- Karanam, B., Jagu, S., Huh, W. K., and Roden, R. B. (2009). Developing vaccines against minor capsid antigen L2 to prevent papillomavirus infection. *Immunol. Cell Biol.* 87, 287–299. doi: 10.1038/icb.2009.13
- Kawana, K., Kawana, Y., Yoshikawa, H., Taketani, Y., Yoshiike, K., and Kanda, T. (2001). Nasal immunization of mice with peptide having a cross-neutralization epitope on minor capsid protein L2 of human papillomavirus type 16 elicit systemic and mucosal antibodies. *Vaccine* 19, 1496–1502. doi: 10.1016/S0264-610X(00)00367-4
- Kohl, T., Hitzeroth, I. I., Stewart, D., Varsani, A., Govan, V. A., Christensen, N. D., et al. (2006). Plant-produced cottontail rabbit papillomavirus L1 protein protects against tumor challenge: a proof-of-concept study. *Clin. Vaccine Immunol.* 13, 845–853. doi: 10.1128/CI.00072-06
- Kohl, T. O., Hitzeroth, I. I., Christensen, N. D., and Rybicki, E. P. (2007). Expression of HPV-11 L1 protein in transgenic *Arabidopsis thaliana* and *Nicotiana tabacum*. *BMC Biotechnol.* 7:56. doi: 10.1186/1472-6750-7-56
- Krinner, S., Heitzer, A. P., Diermeier, S. D., Obermeier, I., Langst, G., and Wagner, R. (2014). CpG domains downstream of TSSs promote high levels of gene expression. *Nucleic Acids Res.* 42, 3551–3564. doi: 10.1093/nar/gkt1358
- Kudla, G., Lipinski, L., Caffin, F., Helwak, A., and Zylicz, M. (2006). High guanine and cytosine content increases mRNA levels in mammalian cells. *PLoS Biol.* 4:e180. doi: 10.1371/journal.pbio.0040180
- Leder, C., Kleinschmidt, J. A., Wieth, C., and Muller, M. (2001). Enhancement of capsid gene expression: preparing the human papillomavirus type 16 major structural gene L1 for DNA vaccination purposes. *J. Virol.* 75, 9201–9209. doi: 10.1128/JVI.75.19.9201-9209.2001
- Liu, W. J., Gao, F., Zhao, K. N., Zhao, W., Fernando, G. J., Thomas, R., et al. (2002). Codon modified human papillomavirus type 16 E7 DNA vaccine enhances cytotoxic T-lymphocyte induction and anti-tumour activity. *Virology* 301, 43–52. doi: 10.1006/viro.2002.1584
- Maclelean, J., Koekemoer, M., Olivier, A. J., Stewart, D., Hitzeroth, I., Rademacher, T., et al. (2007). Optimization of human papillomavirus type 16 (HPV-16) L1 expression in plants: comparison of the suitability of different HPV-16 L1 gene variants and different cell-compartment localization. *J. Gen. Virol.* 88, 1460–1469. doi: 10.1099/vir.0.82718-0
- McGrath, M., De Villiers, G. K., Shephard, E., Hitzeroth, I., and Rybicki, E. P. (2013). Development of human papillomavirus chimaeric L1/L2 candidate vaccines. *Arch. Virol.* 158, 2079–2088. doi: 10.1007/s00705-013-1713-8
- Merlin, M., Gecchele, E., Capaldi, S., Pezzotti, M., and Avesani, L. (2014). Comparative evaluation of recombinant protein production in different biofactories: the green perspective. *Biomed Res. Int.* 2014:136419. doi: 10.1155/2014/136419
- Mole, S., Veerapraditsin, T., McPhillips, M. G., and Graham, S. V. (2006). Regulation of splicing-associated SR proteins by HPV-16. *Biochem. Soc. Trans.* 34, 1145–1147. doi: 10.1042/BST0341145
- Mori, S., Ozaki, S., Yasugi, T., Yoshikawa, H., Taketani, Y., and Kanda, T. (2006). Inhibitory cis-element-mediated decay of human papillomavirus type 16 L1-transcript in undifferentiated cells. *Mol. Cell Biochem.* 288, 47–57. doi: 10.1007/s11010-006-9117-7
- Mortimer, E., Maclean, J. M., Mbewana, S., Buys, A., Williamson, A. L., Hitzeroth, I. I., et al. (2012). Setting up a platform for plant-based influenza virus vaccine production in South Africa. *BMC Biotechnol.* 12:14. doi: 10.1186/1472-6750-12-14
- Nandi, S., Kwong, A. T., Holtz, B. R., Erwin, R. L., Marcel, S., and McDonald, K. A. (2016). Techno-economic analysis of a transient plant-based platform for monoclonal antibody production. *MAbs* 8, 1456–1466. doi: 10.1080/19420862.2016.1227901
- Obembe, O. O., Popoola, J. O., Leelavathi, S., and Reddy, S. V. (2011). Advances in plant molecular farming. *Biotechnol. Adv.* 29, 210–222. doi: 10.1016/j.biotechadv.2010.11.004
- Pastrana, D. V., Gambhira, R., Buck, C. B., Pang, Y. Y., Thompson, C. D., Culp, T. D., et al. (2005). Cross-neutralization of cutaneous and mucosal Papillomavirus types with anti-sera to the amino terminus of L2. *Virology* 337, 365–372. doi: 10.1016/j.viro.2005.04.011
- Pereira, R., Hitzeroth, I. I., and Rybicki, E. P. (2009). Insights into the role and function of L2, the minor capsid protein of papillomaviruses. *Arch. Virol.* 154, 187–197. doi: 10.1007/s00705-009-0310-3
- Pineo, C. B., Hitzeroth, I. I., and Rybicki, E. P. (2013). Immunogenic assessment of plant-produced human papillomavirus type 16 L1/L2 chimaeras. *Plant Biotechnol. J.* 11, 964–975. doi: 10.1111/pbi.12089
- Raab, D., Graf, M., Notka, F., Schodl, T., and Wagner, R. (2010). The GeneOptimizer Algorithm: using a sliding window approach to cope with the vast sequence space in multiparameter DNA sequence optimization. *Syst. Synth. Biol.* 4, 215–225. doi: 10.1007/s11693-010-9062-3
- Roden, R. B., Yutzy, W. H. T., Fallon, R., Inglis, S., Lowy, D. R., and Schiller, J. T. (2000). Minor capsid protein of human genital papillomaviruses

- contains subdominant, cross-neutralizing epitopes. *Virology* 270, 254–257. doi: 10.1006/viro.2000.0272
- Rollman, E., Arnheim, L., Collier, B., Oberg, D., Hall, H., Klingstrom, J., et al. (2004). HPV-16 L1 genes with inactivated negative RNA elements induce potent immune responses. *Virology* 322, 182–189. doi: 10.1016/j.viro.2004.02.002
- Rybicki, E. P. (2010). Plant-made vaccines for humans and animals. *Plant Biotechnol. J.* 8, 620–637. doi: 10.1111/j.1467-7652.2010.00507.x
- Rybicki, E. P. (2014). Plant-based vaccines against viruses. *Viol. J.* 11:205. doi: 10.1186/s12985-014-0205-0
- Schiller, J. T., Castellsague, X., Villa, L. L., and Hildesheim, A. (2008). An update of prophylactic human papillomavirus L1 virus-like particle vaccine clinical trial results. *Vaccine* 26(Suppl. 10), K53–K61. doi: 10.1016/j.vaccine.2008.06.002
- Schwartz, S. (1998). Cis-acting negative RNA elements on papillomavirus late mRNAs. *Semin. Virol.* 8, 291–300. doi: 10.1006/smvy.1997.0131
- Schwartz, S. (2013). Papillomavirus transcripts and posttranscriptional regulation. *Virology* 445, 187–196. doi: 10.1016/j.viro.2013.04.034
- Signorelli, C., Odone, A., Ciorba, V., Cella, P., Audisio, R. A., Lombardi, A., et al. (2017). Human papillomavirus 9-valent vaccine for cancer prevention: a systematic review of the available evidence. *Epidemiol. Infect.* 145, 1962–1982. doi: 10.1017/S0950268817000747
- Studentsov, Y. Y., Schiffman, M., Strickler, H. D., Ho, G. Y., Pang, Y. Y., Schiller, J., et al. (2002). Enhanced enzyme-linked immunosorbent assay for detection of antibodies to virus-like particles of human papillomavirus. *J. Clin. Microbiol.* 40, 1755–1760. doi: 10.1128/JCM.40.5.1755-1760.2002
- Takeda, A., Sugiyama, K., Nagano, H., Mori, M., Kaido, M., Mise, K., et al. (2002). Identification of a novel RNA silencing suppressor, NSs protein of Tomato spotted wilt virus. *FEBS Lett.* 532, 75–79. doi: 10.1016/S0014-5793(02)03632-3
- Toh, Z. Q., Licciardi, P. V., Fong, J., Garland, S. M., Tabrizi, S. N., Russell, F. M., et al. (2015). Reduced dose human papillomavirus vaccination: an update of the current state-of-the-art. *Vaccine* 33, 5042–5050. doi: 10.1016/j.vaccine.2015.07.102
- Twyman, R. M., Schillberg, S., and Fischer, R. (2013). Optimizing the yield of recombinant pharmaceutical proteins in plants. *Curr. Pharm. Des.* 19, 5486–5494. doi: 10.2174/1381612811319310004
- Twyman, R. M., Stoger, E., Schillberg, S., Christou, P., and Fischer, R. (2003). Molecular farming in plants: host systems and expression technology. *Trends Biotechnol.* 21, 570–578. doi: 10.1016/j.tibtech.2003.10.002
- Varsani, A., Williamson, A. L., De Villiers, D., Becker, I., Christensen, N. D., and Rybicki, E. P. (2003a). Chimeric human papillomavirus type 16 (HPV-16) L1 particles presenting the common neutralizing epitope for the L2 minor capsid protein of HPV-6 and HPV-16. *J. Virol.* 77, 8386–8393. doi: 10.1128/JVI.77.15.8386-8393.2003
- Varsani, A., Williamson, A. L., Rose, R. C., Jaffer, M., and Rybicki, E. P. (2003b). Expression of Human papillomavirus type 16 major capsid protein in transgenic *Nicotiana tabacum* cv. XanthiArch. *Viol.* 148, 1771–1786. doi: 10.1007/s00705-003-0119-4
- Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003). An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *Plant J.* 33, 949–956. doi: 10.1046/j.1365-313X.2003.01676.x
- Zhao, X., and Schwartz, S. (2008). Inhibition of HPV-16 L1 expression from L1 cDNAs correlates with the presence of hnRNP A1 binding sites in the L1 coding region. *Virus Genes* 36, 45–53. doi: 10.1007/s11262-007-0174-0
- Zheng, Z. M., and Baker, C. C. (2006). Papillomavirus genome structure, expression, and post-transcriptional regulation. *Front. Biosci.* 11, 2286–2302. doi: 10.2741/1971
- zur Hausen, H. (1996). Papillomavirus infections — a major cause of human cancers. *Biochim. Biophys. Acta Rev. Cancer* 1288, F55–F78. doi: 10.1016/0304-419X(96)00020-0
- zur Hausen, H., De Villiers, E. M., and Gissmann, L. (1981). Papillomavirus infections and human genital cancer. *Gynecol. Oncol.* 12, S124–S128. doi: 10.1016/0090-8258(81)90067-6

Conflict of Interest Statement: MG declares a competing financial interest: Genent AG performs gene design optimization as a free service with the genes that it sells.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Hitzeroth, Chabeda, Whitehead, Graf and Rybicki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.