universität wien

# MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

## Genetic manipulation and recombinant protein expression in *Histomonas meleagridis*


verfasst von / submitted by

Julia Loni Schellhammer, BSc


angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien, 2023 / Vienna, 2023

| | |
|---|---|
| Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet: | UA 066 834 |
| Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet: | Masterstudium Molekulare Biologie |
| Betreut von / Supervisor: | Univ.-Prof. Dr.med.vet. Michael Hess, Dipl.ECPVS |

# Abstract

The protozoan parasite *Histomonas meleagridis* is the causative agent of blackhead disease in poultry. Currently, no modifications have been made to its DNA. Therefore, the aim of this work was to establish a transfection protocol to introduce foreign DNA into *H. meleagridis* cells, laying the foundation for future CRISPR/Cas 9 modifications.

Six different plasmids carrying the green fluorescent protein (GFP) reporter gene were designed and constructed. These plasmids included three different promoter or terminator sequences and three different GFP genes. Three different methods were tested to transfect the plasmids into parasite cells: (1) conventional electroporation with an *E. coli* Pulser™ device; (2) electroporation with the Neon Transfection System; and (3) lipofection.

To assess the success of the transformation protocol, a read-out system was employed. This system involved detecting GFP protein in histomonad cells through fluorescence microscopy and Western blot analysis. In addition, a specific PCR reaction targeting the GFP gene was applied to determine whether the plasmid had entered the parasite cell.

Unfortunately, none of the experiments yielded any detection of the GFP gene or GFP protein in *H. meleagridis* cells, suggesting that the transfection attempts were unsuccessful. To evaluate the effectiveness of the transfection systems and plasmid constructs, transfections were performed on *Trichomonas gallinae,* another avian protozoan parasite, another avian protozoan parasite, and on hepatocellular carcinoma cell line of male leghorn chicken (LMH). Only the LMH cells produced a GFP signal after transfection indicating a deficiency to establish transfections in protozoan parasites.

Based on these results, it is obvious that further variations of transfection protocols need to be explored and evaluated before successful genomic manipulation of *H. meleagridis* can be achieved.

## German abstract- deutsche Zusammenfassung

*Histomonas meleagridis*, der Erreger der Schwarzkopfkrankheit beim Geflügel, ist ein einzelliger Parasit und bis zum jetzigen Zeitpunkt wurde noch keine Modifikation an seinem DNA -Material vorgenommen. Ziel dieser Arbeit war die Etablierung eines Transfektionsprotokolls zur Einführung fremder DNA in *H. meleagridis*-Zellen und damit die Schaffung einer Grundlage für zukünftige CRISPR/Cas 9-Modifikationen.

Sechs verschiedene Plasmide, mit dem Reportergen für das grün fluoreszierende Protein (GFP), wurden entworfen und konstruiert. Insgesamt wurden drei verschiedene Promotor- und Terminator- Sequenzen und drei verschiedene GFP-Gene getestet. Für die Transfektion des Plasmids in die Parasitenzellen wurden drei verschiedene Verfahren getestet: (1) herkömmliche Elektroporation mit einem *E. coli* Pulser TM -Gerät; (2) Elektroporation mit dem Neon-Transfektionssystem; und (3) Lipofektion.

Zur Überprüfung des Erfolges des Transformationsprotokolls, wurde ein Testsystem eingesetzt. Dieses System basierte auf der Detektion des GFP-Proteins in den Histomonas-Zellen mittels Fluoreszenzmikroskopie und Western-Blot-Analyse. Zusätzlich wurde eine spezifische PCR-Reaktion für das GFP-Gen durchgeführt, um festzustellen, ob das Plasmid in die Parasitenzelle gelangt war.

Keines der Experimente führte zur Detektion des GFP-Proteins oder des GFP-Gens in den *H. meleagridis*-Zellen, was darauf hindeutet, dass die Transfektionsversuche nicht erfolgreich waren. Um die Wirksamkeit der angewandten Transfektionssysteme und Plasmidkostrukte zu bewerten, wurden Transfektionen in *Trichomonas gallina*e, einem weiteren aviären Protozoen, und in der hepatozellulären Karzinom-Zelllinie von männlichen Leghorn-Hühnern (LHM) durchgeführt. Doch nur die Kontroll-LMH-Zellen produzierten ein GFP-Signal, was die Probleme Protozoenzellen zu transfizieren unterstereicht. Basierend auf diesen Ergebnissen ist offensichtlich, dass weitere Variationen der Transfektionsprotokolle erforscht und bewertet werden müssen, bevor eine erfolgreiche genomische Manipulation von *H. meleagridis* erreicht werden kann.

# Index

# List of abbreviation

| | |
|---|---|
| **Cas** | CRISPR-associated |
| **CRISPR** | Clustered Regularly Interspaced Short Palindromic Repeats |
| **DNA** | Desoxyribonucleic acid |
| **DNase 1** | deoxyribonuclease 1 |
| **GFP** | Green fluorescent protein |
| **IHC** | Immunohistochemistry |
| **LB** | Luria Broth |
| **PCR** | Polymerase chain reaction |
| **SDS-PAGE** | sodium dodecyl sulfate polyacrylamide gel electrophoresis |

# 1 Introduction

*Histomonas meleagridis* is the causative agent of blackhead disease also known as histomonosis in poultry [1,2]. The infection manifests as an enterohepatitis with pathological lesions in the caeca and liver of turkeys, whereas in chickens they are often limited to the caeca [2,3]. Due to severity of the disease in turkeys, often entire flocks have to be killed, which results in substantial compromise of animal welfare and high economic losses [2]. For a century histomonosis was well controlled with anti-histomonal products used for prophylaxis and therapy, however, changes in the drug legislation in Europe and the USA in the last two decades, caused a re-emergence of the disease. As a result, research on the parasite is gaining importance again [4,5]. Even though *H. meleagridis* has been already described in 1893 [6], very little is known about this pathogen on the genomic level. Several molecular studies on *H. meleagrids* were published, encompassing phylogenetic analyses, genotyping, different omic- studies and detection and diagnostics. Table 1 lists the most important ones. Until now *H. meleagridis* has not been manipulated at the genetic level. The complete genome has only recently been published, allowing the identification of specific DNA regions [7] and with that providing the basis for genetic engineering in *H. meleagridis* for the first time.

**Table 1: List of molecular studies on *H. meleagridis***

| Study outcome | Technique | Title | Year of publication [reference] |
|---|---|---|---|
| **Phylogenetic / Genotyping studies** | | | |
| Close relationship between *H. meleagridis*, and *Dientamoeba fragilis* | rRNA analyses | Phylogenetic Position of the Trichomonad Parasite of Turkeys, *Histomonas meleagridis* (Smith) Tyzzer, Inferred from Small Subunit rRNA Sequence | 2001 [8] |
| Three different *H. meleagridis* genotypes were identified | C-profiling | Genotyping of *Histomonas meleagridis* isolates based on Internal Transcribed Spacer-1 sequences | 2006 [9] |
| Relationship between *H. meleagridis*, and *Dientamoeba fragilis* | DNA Sequence analyses | Partial sequence of the beta-tubulin of *Histomonas meleagridis* and the activity of benzimidazoles against *H. meleagridis* in vitro | 2009 [10] |
| Classification of *H. meleagridis;* class: Tritrichomonadea, order: Tritrichomonadida, family: Dientamoebidae | rRNA analyses | Critical Taxonomic Revision of Parabasalids with Description of one New Genus and three New Species | 2010 [11] |
| Close relationship of *H. meleagridis* with *Tritrichomonas foetus* | DNA sequence analyses | Systematic Position of *Histomonas meleagridis* Based on Four Protein Genes | 2010 [12] |
| Identification of various *H. meleagridis* strains | rRNA analyses | Molecular Characterization of *Histomonas meleagridis* and Other Parabasalids in the United States Using the 5.8S, ITS-1, and ITS-2 rRNA Regions | 2011 [13] |
| *H. meleagridis* was separated in two different genotypes | Sequence analyses | Multi-Locus Typing of *Histomonas meleagridis* Isolates Demonstrates the Existence of Two Different Genotypes | 2014 [14] |
| Monoxenic vaccine candidate as an efficacious prophylactic strategy | qPCR and IHC | Experimental reproduction of histomonosis caused by *Histomonas meleagridis* genotype 2 in turkeys can be prevented by oral vaccination of day-old birds with a monoxenic genotype 1 vaccine candidate | 2022 [15] |
| **Molecular studies** | | | |
| Identification of three proteins involve in the hydrogenosomal carbon metabolism | Sequence analyses | First molecular characterisation of hydrogenosomes in the protozoan parasite *Histomonas meleagridis* | 2008 [16] |
| Identification of antigenic proteins of *H. meleagridis* | cDNA library | Identification and molecular characterization of numerous *Histomonas meleagridis* proteins using a cDNA library | 2009 [17] |
| Identification of upstream sequences of 2 β-tubulin genes | Splinkerette PCR | Identification of Gene Expression Elements in *Histomonas meleagridis* Using Splinkerette PCR, a Variation of Ligated Adaptor PCR | 2012 [18] |

| | | | |
|---|---|---|---|
| Identification of 81 genes coding for putative hydrogenosomal proteins and codon usage frequency | cDNA library | A genomic analysis of *histomonas meleagridis* through sequencing of a cDNA library | 2013 [19] |
| **Omics** | | | |
| Identification of proteins specific for the virulent and attenuated strain | two-dimensional electrophoresis | Unravelling the differences: comparative proteomic analysis of a clonal virulent and an attenuated *Histomonas meleagridis* strain | 2017 [20] |
| Identification of proteins specific for the virulent and attenuated strain | two-dimensional electrophoresis / MALDI-TOF/TOF | An Alliance of Gel-Based and Gel-Free Proteomic Techniques Displays Substantial Insight into the Proteome of a Virulent and an Attenuated *Histomonas meleagridis* Strain | 2018 [21] |
| Exoproteomic differences between the virulent and the attenuated strain | 1D-in-gel-zymography / micro-LC-ESI-MS/MS | Molecular characterization of *Histomonas meleagridis* exoproteome with emphasis on protease secretion and parasite-bacteria interaction | 2019 [22] |
| Complete genome sequence of two closely related *H. meleagridis* strains | Illumina / nanopore sequencing | Complete genomes of the eukaryotic poultry parasite *Histomonas meleagridis*: linking sequence analysis with virulence / attenuation | 2021 [7] |
| Identification of surface proteins in the virulent and attenuated strains | nanoHPLC-Orbitrap MS/MS analysis | Comparative surfaceome analysis of clonal *Histomonas meleagridis* strains with different patholgenicity reveals strain-dependent profiles | 2022 [23] |

There are many different tools for manipulating genetic material including restriction enzymes, DNA sequencing, and DNA cloning [24]. A key components of molecular cloning are plasmid vectors; so-called vehicles used to introduce recombinant DNA into a host cell [25]. Plasmids are extrachromosomal DNA elements that can replicate independently of chromosomal DNA and are often used for gene transfer. They have a number of advantages: a simple structure, easy to isolate and modify *in vitro* [26]. A simple vector consists of three major components: an origin of replication (ORI) in *Escherichia coli*, a selection marker and a multiple cloning site (MCS). In general, the ORI and the selection marker aid the replication of the plasmid in *E. coli*. The MSC facilitates an easy cloning of the desired DNA. There are three classes of commonly used plasmid vectors: cloning vectors, expression vectors and reporter vectors. Cloning vectors are used to amplify a DNA of interest. Expression vectors are used for large scale protein production and with the system of a reporter vector it is possible to analyse the promoter dynamics, by cloning a promoter of interest with an additional reporter gene into the MCS [25].

An often-used reporter gene is for example the green fluorescent protein (GFP). Originally isolated from a jellyfish *Aequorea vctoria* [27], nowadays there are many different GFP variants, isolated from different organisms or genetic modifications [28]. Advantages of using GFP proteins are that they do not require an additional substrate or ATP for a fluorescence signal. Furthermore, due to their small size GFP can be attached to other proteins without affecting their function serving as quantitative marker for gene expression [27–29].

For plasmid construction there are several established cloning methods. One method that allows the joining of multiple DNA fragments in a single step, into a single plasmid, is the Gibson assembly method [30,31]. This homology-based cloning method requires the design of specific primers with a homologous overhang to the desired location in the construct. All fragments are combined within one reaction using three enzymes simultaneously, an exonuclease (5'-3'), a DNA polymerase and a ligase. In a single tube reaction, first, a 5'exonuclease digests up to 200 nucleotides creating a 3' single-stranded overhang. Once the overhang is exposed, the homologous regions on DNA fragments can join, and the repair mechanism starts with a polymerase synthesizing the exposed DNA and filling-in the gaps. At last the ligase closes all the gaps [30–32]. With this method, up to 15 fragments can be simultaneously joined without having specific restriction sites [33]. Gibson assembly is a

frequently used method in CRISPR/Cas experiments. It allows to insert large and GC-rich gene sequences into a plasmid [34].

The CRISPR/Cas system is a relatively new innovation for precise gene editing by introducing genomic DNA breaks at specific sites. CRISPR that stands for Clustered Regularly Interspaced Short Palindromic Repeats and was first found in *E. coli* in 1987 [35]. Later, this locus was detected in 84% of archaea and 45% of bacteria species. The CRISPR gene needs additionally CRISPR-associated (Cas) genes for its full activity. This system serves in bacteria and archaea as an antiviral defence mechanism. The simplified version of CRISPR/Cas9 system is the most commonly used CRISPR system for introducing precise gene modification in eukaryotic cells [35,36]. It requires a guide RNA to direct the Cas9 nuclease to an exact location in the genome. The guide RNA is part of the riboprotein complex with the Cas9 nuclease. As soon as the RNA binds to the genomic DNA, Cas9 is activated, leading to a double-strand break in the DNA at the specific location [37,38]. The generated double strand break can be repaired in two ways, either by non-homologous end joining which is the typical repair of double-strand DNA (dsDNA) breaks or by homology-directed repair via homologous recombination. In non-homologous end joining double-strand-break ends are ligated directly. However, this process is an error prone process, because during this process some bases are lost, which leads to a random mutation [37–40]. The second, the homology directed repair mechanism is less error-prone and uses homologous donor DNA to repair dsDNA damage. The advantage of this method is that by providing a DNA repair template exact changes in the gene or insertion of completely new genes is possible. Oligonucleotides of up to 200 bases length, single-stranded DNA in the form of recombinant adeno-associated virus (rAAV) and circular or linear plasmid DNA are options for homologues repair templates [41]. Plasmids are the most frequently used ones [38,39]. Yet, in order to be able to alter the genome at all, the above mentioned riboprotein complex of Cas, the guide RNA together with the repair templates have to enter the cell [42].

Transfection is the process of deliberately introducing foreign DNA or RNA into eukaryotic cells [43]. There are different transfection methods for this purpose. It is generally divided into chemical and physical methods. In chemical transfection, the DNA is packaged in a complex and taken up by the cell via endocytosis. This method depends on many factors, such as the pH value, the structure of the cell membrane or the nucleic acid/ chemical ratio

[42]. This method offers advantages, for example low cytotoxicity, no mutagenesis, no additional carrier DNA, and no size limitation on the packaged nucleic acid. Lipofection is a chemical transfection method, in which the nucleic acid (e.g., plasmid) is packaged in lipid pots, vesicle that can bind to the cell membrane. This allows the nucleic acid to be taken up by the cell and processed further [42,43]. In the physical method, the nucleic acid is introduced into the cell directly via physical process using a tool, an electrical impulse, an ultrasound or laser. As a result, the cell can be damaged, which usually leads to a relatively high mortality of the cells. One of the most common methods is an electroporation. In electroporation, cells are exposed to the electric field of sufficient amplitude, in form of short electrical impulses which causes the cell plasma membrane to become permeable or porous allowing the nucleotide acids to be absorbed through these "holes". With this method, many cells can be transfected simultaneously. There is an optimal setting for electroporation for each organism. But generally, as the voltage increases, the transfection rate increases, while the cell survival rate decreases [42–44].

# 2 Aims

In order to establish genetic manipulation method for *H. meleagridis*, the objectives of this work were to:

(1) construct various plasmids with a promoter and terminator region specific for *H. meleagridis* and a GFP gene as a reporter gene;

(2) test various transfection methods for *H. meleagridis* using the plasmid constructs,

(3) establish a screening/detection method of the GFP reporter signal in transfected cells and

(4) establish the basis for CRISPR/Cas 9 experiments for *H. meleagridis* based upon 1-3 .

# 3 Material and methods

## 3.1 Organisms and cell line

***Histomonas meleagridis***

The virulent *H. meleagridis* strain, *H. meleagridis*/Turkey/Austria/2922-C6/04 grown as monoxenic mono-eukaryotic culture with the *E. coli DH5 α* strain was used in all experiments [45]. The cultures were incubated under anaerobic conditions at 40°C in medium M199 (Gibco, Thermo Fisher Scientific, USA) supplemented with 15% FBS (Gibco, Thermo Fisher Scientific, USA) and 0.25% rice starch. The cells were routinely passaged every 2 to 3 days. Only passage numbers between 12 and 35 were used to ensure that the cells still correspond to the virulent cell type. Cells that had grown exponentially over 48 hours were used for each experiment.

***Trichomonas gallinae***

Xenic *T. gallinae* culture, *T. gallinae*/Budgerigar/Austria/5895-C1/06 [46] with the wild-type bacterial flora was incubated in the same medium and under the same conditions as *H. meleagridis* descried above. The cells were passaged every 2 to 3 days and exponentially grown (48h) cells between passage number 18 and 30 were used for experiments.

***Leghorn Male Hepatoma* (LMH) cell-line**

LMH cells [47] were commercially sourced (Thermo Fisher Scientific, USA). The cells were grown in RPMI 1640 medium (Gibco, Thermo Fisher Scientific, USA) supplemented with 10% FBS (Gibco, Thermo Fisher Scientific, USA) and 0.5% antibiotic-antimycotic solution (Gibco, Thermo Fisher Scientific, USA) [47]. The cells were incubated at 37°C with 5% $CO_2$. The culture was passaged every 2 to 3 days. Cells between passage number 18 and 22 were used for experiments.

## 3.2 Plasmids

Nine different plasmids were tested in the experiments. Three of them were commercially sourced. The other six plasmids were designed based on the previously reported transformation constructs of *Trichomonas vaginalis* [48]. Five out of this six plasmids were built with Gibson assembly method using the NEBuilder HiFi DNA Assembly Cloning Kit (New England

BioLabs Inc., USA) and the remaining plasmid was synthesized by a plasmid construction service (Eurofins Genomics, GmbH, Germany) (Table 2).

**Table 2. Overview of tested plasmids.**

| Plasmid name | Commercial | Gibson assembly | Synthesized |
|---|---|---|---|
| pUC19_*g6464_AcGFP*opt | No | **Yes** | No |
| pBlue_*g6464_AcGFP*opt | No | **Yes** | No |
| pBlue_Tricho_*eGFP*opt | No | No | **Yes** |
| pBlue_*g7554_eGFP*opt | No | **Yes** | No |
| pBlue_Tricho_*AcGFP* | No | **Yes** | No |
| pBlue_*g7554_AcGFP* | No | **Yes** | No |
| pMaxGFP | **Yes** | No | No |
| pGFPuv | **Yes** | No | No |
| p*AcGFP*1-C1 | **Yes** | No | No |

### 3.2.1 Commercial plasmids

Three commercial plasmids available in the laboratory were tested. The pMaxGFP plasmid (Lonza Ltd., Switzerland) that has an early cytomegalovirus (PCMV IE) promoter and a chimeric intron in front of a GFP gene [49]. The pGFPuv plasmid (Clontech, France) has a lacZ promoter in front of the GFPuv gene [45,50]. Finally, the p*AcGFP*1-C1 plasmid, which has a cytomegalovirus (CMV) promoter in front of the *AcGFP*1-C1 gene [51]. All plasmids were transformed into NEB 5-α chemically competent *E. coli* cells (New England Biolabs Inc., USA) by chemical transformation implementing heat shock according to the manufacturer's protocol. All the commercial plasmids served as controls. Plasmid maps are given in Annex 1.

### 3.2.2 Plasmid and insert design

Six different plasmids were designed using the NEBuilder assembly tool (https://nebuilder.neb.com/#!/). The plasmid and the insert design were based on the previously reported transformation construct of *T. vaginalis* [48]. pUC19 (New England BioLabs Inc., USA) and pBluescript II sk (+) (Agilent Technologies, Inc., United States) served as plasmid backbones.

The insert contained a promoter region, the GFP gene and a terminator region. The promoter sequence was carefully selected to cover all TATA-box options in front of the target gene (Figure 1). The target gene with the hypothetical promoter sequences were chosen based on the transcriptome analysis data of *H. meleagridis* [52], with the ideal criteria of being expressed in every cell and throughout the entire cell cycle. In addition, the first approx. 80 base pairs (bp) of the selected gene was fused in-frame to the reporter genes, as reported previously for *T. vaginalis* [48].

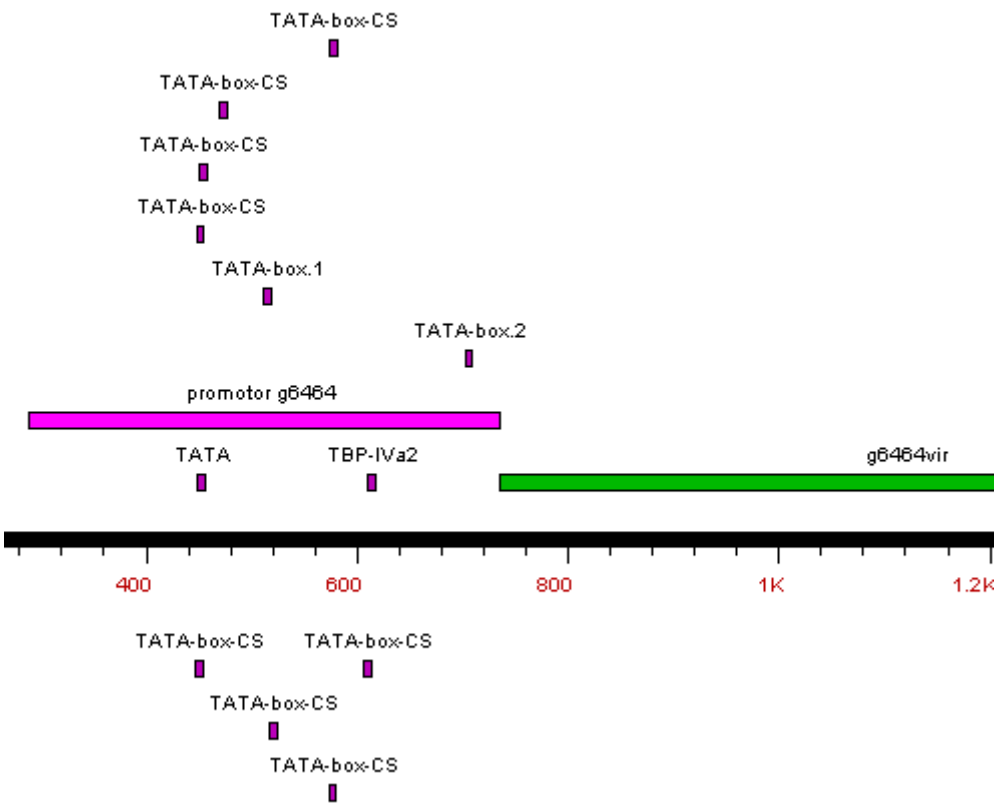



**Figure 1. Schematic representation of the putative g6464 promoter as example for the promotor sequence design.** The putative promoter region of the gene *g6464* (GenBank: WSYJ01000142; locus_tag="GO595_006457"; product="40S ribosomal protein S4, X isoform") is labelled as pink rectangular, and contains the 335 bp of the non-coding sequence upstream of the *g6464* gene. The TATA boxes within promoter sequence are labelled as pink boxes. TATA boxes belong to the group of Transcription Factor Binding Motifs. The coding sequence (cds) of *g6464* is labelled as green rectangular.

The plasmids pUC19_*g6464_AcGFPopt* (Figure 2a) and pBlueScript_ *g6464_AcGFPopt* (Figure 2b) differed only in their plasmid backbone. The insert in both plasmids consisted of

the 335 bp region upstream the coding sequence (cds) of *g6464* gene (GenBank: WSYJ01000142; locus_tag="GO595_006457"; product="40S ribosomal protein S4, X isoform"), which was used as a putative promoter sequence (p-*g6464*). Followed by an 84 bp 5'-cds of the g6464 gene, cloned in-frame with *H. meleagridis* codon optimised *AcGFP* reporter gene (*AcGFPopt*). The 520 bp downstream cds of the *g6464* gene was used as a putative terminator sequence (t-*g6464*) (Figure 2c). The pEX_*AcGFP* plasmid (Poultry Clinic, Veterinary University Vienna, unpublished data) was used as amplification template for the *H. meleagridis*-codon optimised, synthetic *AcGFP* reporter gene *AcGFP* (derived from *Aequorea coerulescens* [51]) The complete insert sequence is given in Annex 2.

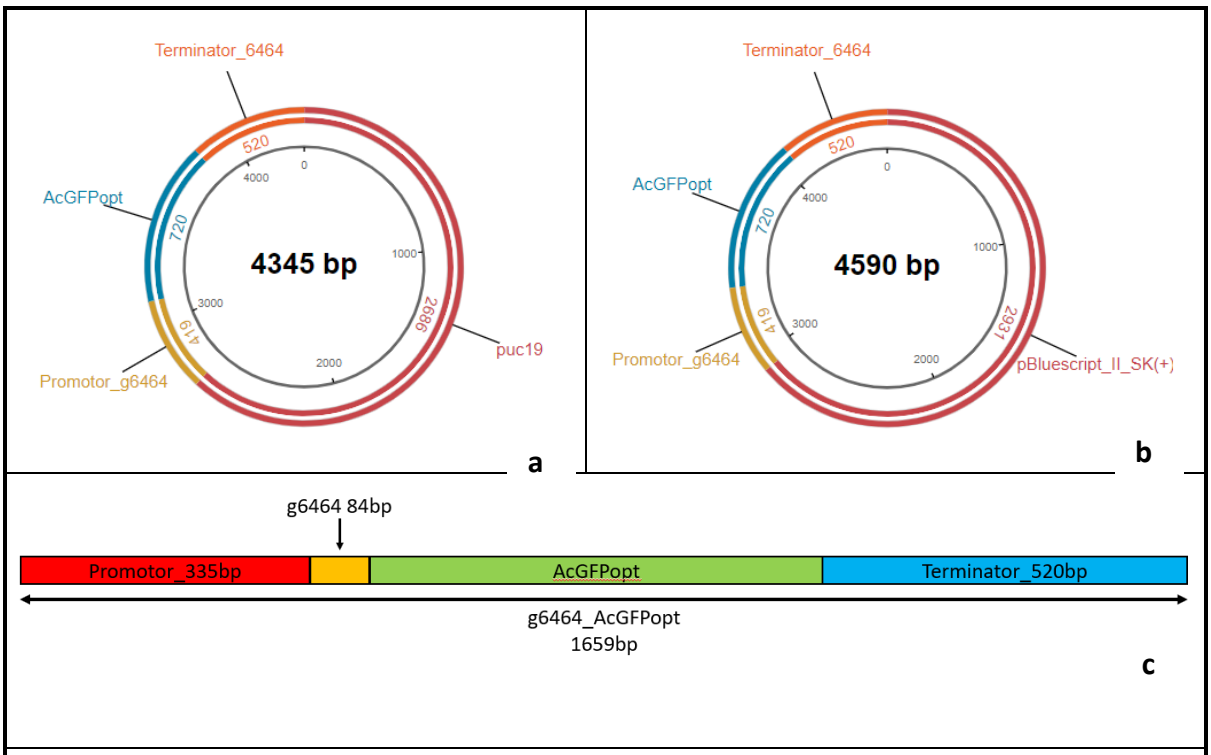



**Figure 2. Schematic representation of the plasmids pUC19_*g6464_AcGFPopt* and pBlue_*g6464_AcGFPopt* and the corresponding inserts. (a)** Plasmid map of pUC19_g6464_*AcGFP*opt plasmid. **(b)** Plasmid map of pBlue_g6464_*AcGFP*opt **(c)** The g6464_Ac*eGFP*opt plasmid insert consists of the 335 bases long promoter matching upstream region of the *g6464* gene (GenBank: WSYJ01000142; locus_tag="GO595_006457"; product="40S ribosomal protein S4, X isoform"), followed by the first 83bp of the gene *g6464* and the codon optimised *AcGFP* gene (*AcGFP*opt). The 520 bp long region after the open reading frame of the *g6464* served as terminator.

The plasmid pBlue_*g7445_eGFPopt* consisted of the pBlueScript II SK (+) as backbone (Figure 3a) with the insert containing a 1603 bp upstream cds of *g7445* gene (GenBank: WSYJ01000152; locus_tag="GO595_007545"; product="hypothetical protein "), which

represents a putative promoter region. Followed by an 83 bp long 5'-cds of the *g7554* gene cloned in-frame with the *H. meleagridis*-codon optimized *eGFP* (enhanced GFP) reporter gene (*eGFPopt*). The insert ends with the 450 bp downstream cds of the *g7554* gene, which served as the terminator sequence. (Figure 3b) The pBlue_Tricho_*eGFPopt* plasmid (Figure 5) was used as amplification template for the *eGFPopt*. This single point mutant enhanced GFP was explored as an alternative to the *AcGFP* protein. The complete insert sequence is given in Annex 3.

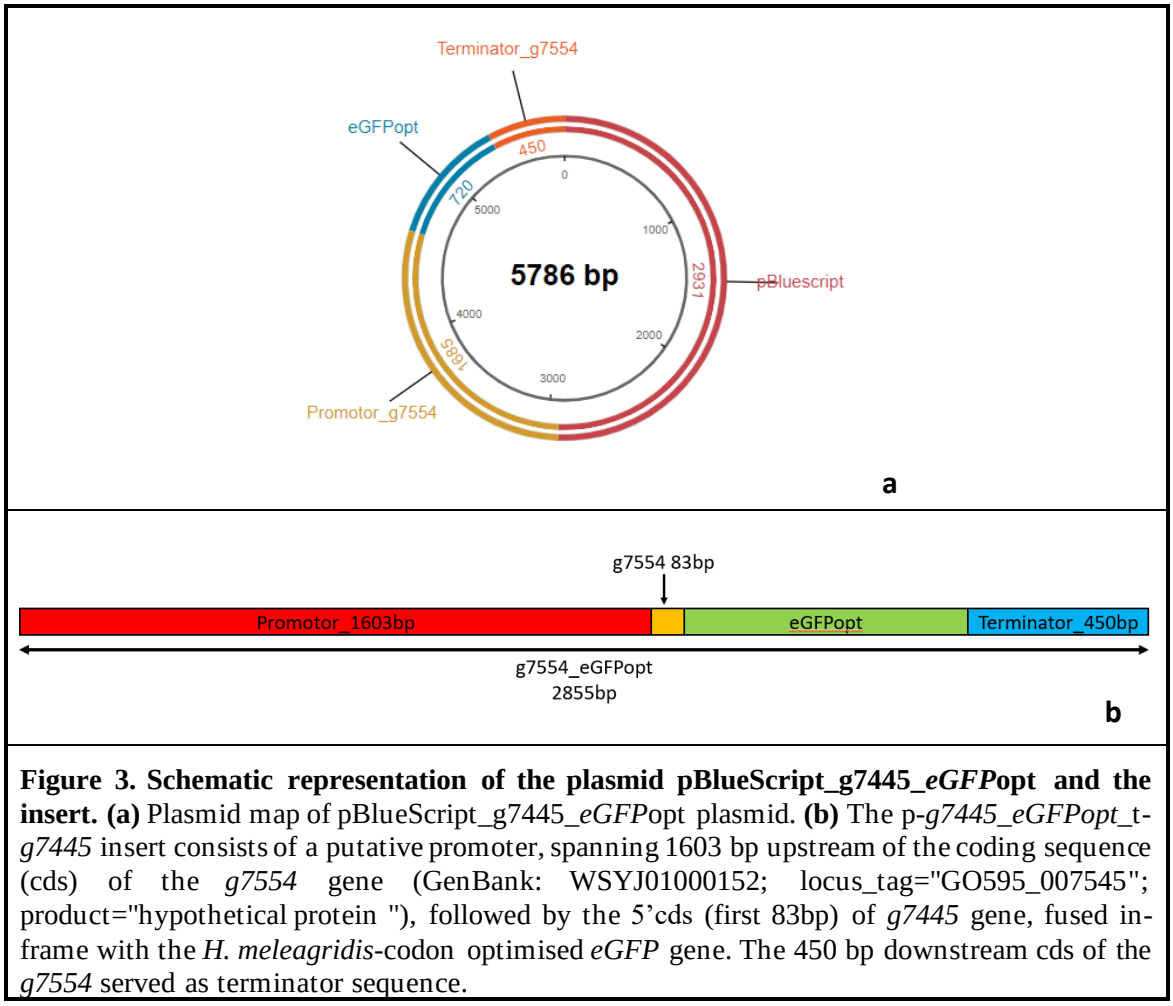



**Figure 3. Schematic representation of the plasmid pBlueScript_g7445_*eGFP*opt and the insert. (a)** Plasmid map of pBlueScript_g7445_*eGFP*opt plasmid. **(b)** The p-*g7445_eGFPopt*_t-*g7445* insert consists of a putative promoter, spanning 1603 bp upstream of the coding sequence (cds) of the *g7554* gene (GenBank: WSYJ01000152; locus_tag="GO595_007545"; product="hypothetical protein "), followed by the 5'cds (first 83bp) of *g7445* gene, fused in-frame with the *H. meleagridis*-codon optimised *eGFP* gene. The 450 bp downstream cds of the *g7554* served as terminator sequence.

The plasmids pBlue_*g7445_AcGFP* (Figure 4a) and pBlue_Tricho_*AcGFP* (Figure 4b) were constructed by replacing the *eGFPopt* gene from the pBlue_*g7445_eGFP*opt and pBlue_Tricho_*eGFPopt* plasmids, respectively, with the *AcGFP1* gene (Figure 4c) sourced from the commercial p*AcGFP1*-C1 plasmid. This substitution was made to eliminate the potential impact of codon optimization on the GFP protein. All other sequences remained unchanged. Complete insert sequence is given in Annex 4.


**Figure 4. Schematic representation of pBlue_*g7445_AcGFP1*_ and pBlue_Tricho_*AcGFP1*_ plasmids and the corresponding insert. (a)** Plasmid map of pBlue_*g7445_AcGFP*, **(b)** Plasmid map of pBlue_Tricho_*AcGFP* and **(c)** The *AcGFP*1+MS insert contains the *AcGFP1* gene and the multiple cloning site (MS) from the p*AcGFP*1-C1 plasmid.
Both plasmids, pBlue_*g7445_AcGFP1* and pBlue_Tricho_*AcGFP1*, are initially very similar to pBlue_*g7445_eGFPopt* and pBlue_Tricho_*eGFPopt* with the exception of the *eGFPopt* gene that was exchanged with the *AcGFP1* gene from a commercial p*AcGFP*1-C1 plasmid, which means the promoter and terminator sequences remained the same.

The promotor / terminator-GFP insert sequence of the pBlue_Tricho_*eGFPopt* (Figure 5a) plasmid was synthesised completely and cloned into the supplied pBluescript II sk (+) vector (Agilent Technologies, Inc., United States) by Eurofins Genomics, GmbH (Ebersberg, Germany). The design contained a copy of the promoter and terminator sequences of the previously published *T. vaginalis* vector (plasmid *α*-SCSB-CAT) [48] and a reporter gene *eGFP*. The promoter and terminator sequences were taken from NCBI entry GenBank: U07779; product=" succinyl CoA synthetase α subunit" (Figure 5b) and the *eGFP* gene from SnapGen [53] that was optimized for *H. meleagridis* using a published codon optimization table [19]. The complete insert sequence is given in Annex 5. The plasmids were transformed into NEB 5-α competent *E. coli* cells (New England Biolabs Inc., USA) by heat shock according to the manufacturer's instructions. England Biolabs Inc., USA).

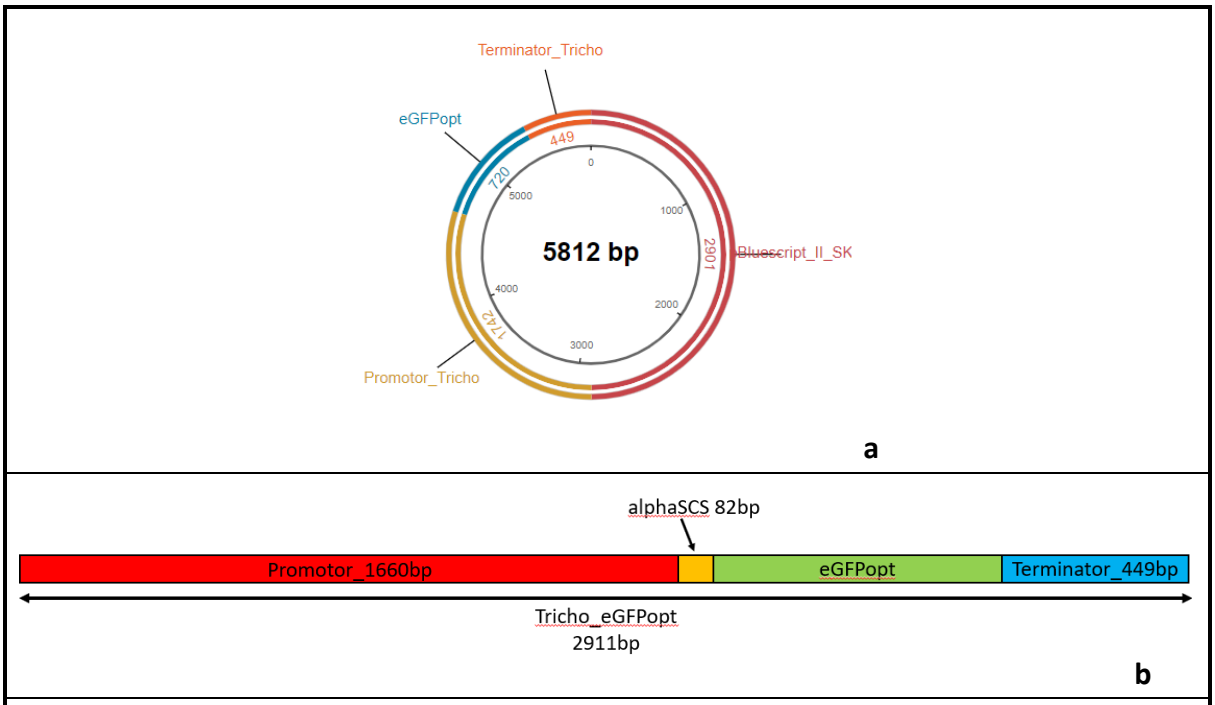



**Figure 5. Schematic representation of the pBlue_Tricho_*eGFPopt* plasmid and the corresponding insert. (a)** Plasmid map of pBlue_Tricho_*eGFPopt* **(b)** The Tricho_*eGFPopt* insert consists of the 1660 bp long region upstream cds of *T. vaginalis αSCSB* gene (GenBank: U07779; product=" succinyl CoA synthetase α subunit"), which served as putative promoter, followed by the 5'-cds region of *αSCS* containing the first 82bp cloned into frame with the *H. meleagridis*-codon optimised *eGFP* gene. The 449 bp region downstream cds of *αSCS* gene served as putative terminator.
Design of this insert was based on the information on the *α*-SCSB-CAT plasmid used in transfection experiments of *T. vaginalis* [48]. The entire insert was synthesised and cloned into a pBluescript II sk (+) vector by Eurofins Genomics, GmbH, Germany.

### 3.2.3 Plasmid construction

#### 3.2.3.1 Fragment amplification

All DNA fragments utilized in the Gibson assembly were generated and amplified through PCR, using the NEB Q5 High-Fidelity 2x Master Mix (New England BioLabs Inc., USA). The plasmid backbone fragment was generated by inversion PCR, whereby the plasmid template was not linearized prior to PCR amplification. Primers containing a 40 bp overhang essential for the Gibson assembly reaction were generated by the NEBuilder assembly tool (https://nebuilder.neb.com/#!/) (Complete primer list is given in Annex 6). The PCR reactions were carried out according to the NEB's protocol with the recommended annealing temperature of the NEBuilder assembly tool. The PCR fragments were visualized using agarose gel electrophoresis (1% agarose gel, 60min, 100V). DNA fragments of the correct size were cut

from the gel and purified using the QIAquick Gel Extraction Kit (Qiagen GmbH, Austria) according to the manufacturer's instructions.

#### 3.2.3.2 Gibson assembly

To join the purified DNA fragments, the NEBulider HiFI DNA Assembly Cloning kit (New England BioLabs Inc., USA) was used according to the manufacturer's instruction. The individual fragments were mixed at a concentration of 0.005 pmol each, with a total amount of 0.02-0.5 pmol of DNA fragments and combined with the NEBuilder HiFi DNA Assembly Master Mix. The mixture was then incubated at 50°C for a duration of one hour. The assembled plasmid was transformed into NEB 5-α chemically competent *E. coli* cells (NEBuilder Hifi DNA Assembly Cloning Kit, New England Biolabs Inc., USA) by heat shock (42°C for 30 seconds) according to the manufacturer's instruction. Following the heat shock the cells recovered in SOC medium (950µL) at 37°C for one hour. 100µL of cells were spread on selection plates (LB plates with 100µg/mL ampicillin, 100µg/mL x-Gal and 100mM IPTG).

#### 3.2.3.3 Small scale plasmid amplification and isolation

For plasmid isolation after the cloning reaction, positive clones were picked from selective plates and incubated in Luria Broth (LB) medium (10mL) with ampicillin (100µg/mL) on a shaking platform (250 rpm) at 37°C, overnight. Plasmid DNA was isolated from 7 mL cultures (3-4 $x10^9$ cells/mL) using the GeneJet Mini Plasmid Prep Kit (Invitrogen, ThermoFisher Scientific Inc., USA). The remaining culture was centrifuged (10 minutes at 4000xg) and re-suspended in 1mL freezing medium (Brain Heart Infusion broth (Oxoid, Thermo Scientific, UK) with 20% glycerol (Sigma-Aldrich, Taufkirchen Germany)) and cryopreserved for long term storage at -80°C.

#### 3.2.3.4 Plasmid verification

In order to confirm the accuracy of the plasmid constructs, a restriction endonuclease digest was conducted. The NEB (New England Biolabs Inc., USA) restriction enzymes were utilized for the reaction in accordance with the manufacturer's instructions. The complete list of the plasmid constructs and the corresponding restriction enzymes used for the verification is listed in Annex 7. Additionally, to rule-out any potential mutations that could have occurred during the cloning process, all plasmid constructs were additionally verified by sequencing (LGC Genomics GmbH, Germany). Each plasmid construct was sequenced with four primers - two binding outside and two within the insert. Detailed list of all primers used for sequencing is listed in Annex 8.

#### 3.2.3.5 Large-scale plasmid amplification and isolation

The large-scale plasmid amplification process was used to get sufficient quantity of plasmid DNA for the transfection experiments. To achieve this, a culture of *E. coli* containing each plasmid construct was grown overnight in LB medium (3x 200mL) supplemented with ampicillin (100 µg/mL). The plasmid isolation was performed using the Maxiprep kit (PureLink HiPure Plasmid Maxiprep Kit Invitrogen, ThermoFisher scientific Inc., USA). The plasmid DNA concentration for each sample was measured using the NanoDrop system (Thermo Scientific NanoDrop 2000, ThermoFisher scientific Inc., USA).

## 3.3 Transfection Methods

### 3.3.1 Transfection using *E. coli* Pulser (*E. coli* Pulser™ Transformation Apparatus, Bio-Rad Laboratories, Inc., USA)

100 µL cells ($1x10^6$ cells/mL) were mixed with 10 µL plasmid (220 ng/µL) (Table 3) at room temperature and were placed in a pre-cooled (-20°C) electroporation cuvette (0.4 cm electrode gap cuvette; Bio-Rad Laboratories, Inc., USA) carefully avoiding air bubbles. An electric pulse was applied as given in Table 4. After the pulse, the cells were immediately placed in a 12-well cell culture plate (Figure 6) with 900 µL of recovery medium and incubated for 24 or 48 hours. The recovery medium utilized for the transfection of *H. meleagridis* cells was an overnight *E. coli* DH5α culture, grown in M199 medium with 15% FBS and diluted in 1:1 ratio with M199 with 15% FBS. The recovery medium for the transfection of *T. gallinae* or LMH was their respective normal culture medium. Incubation of both *H. meleagridis* and *T. gallinae* occurred at 40°C, while LMH cells were incubated at 37°C with 5% $CO_2$.

**Table 3. List of the plasmid tested with various organisms and cell lines.**

| Plasmid | *H. meleagridis* | *T. gallinae* | LMH |
|---|---|---|---|
| pUC19_g6464_*AcGFP*opt | **Yes** | No | No |
| pBlue_g6464_*AcGFP*opt | **Yes** | **Yes** | No |
| pBlue_g7445_*eGFP*opt | **Yes** | No | **Yes** |
| pBlue_Tricho_*eGFP*opt | **Yes** | **Yes** | **Yes** |
| pBlue_g7445_*AcGFP* | **Yes** | **Yes** | No |
| pBlue_Tricho_*AcGFP*t | **Yes** | **Yes** | No |
| pMaxGFP | No | No | **Yes** |

**Table 4: List of voltage settings used for electroporation with the *E. coli* Pulser™ transformation apparatus.**

| Program setting | *H. meleagridis* | *T. gallinae* | LMH |
|---|---|---|---|
| 1100V | No | No | **Yes** |
| 1200V | **Yes** | **Yes** | **Yes** |
| 1300V | **Yes** | **Yes** | **Yes** |
| 1400V | **Yes** | **Yes** | **Yes** |
| 1500V | **Yes** | **Yes** | No |
| 1600V | **Yes** | **Yes** | No |
| 1700V | **Yes** | **Yes** | No |


**Figure 6. Example of a 12-well cell culture plate used for the experimental setup.** For the transfection, 100µL (1x10$^6$ cells/mL) *H. meleagridis* or *T. gallinae* cells were used. Following electroporation with either the *E. coli* Pulser or the Neon Transfection System, the cells were added to the well and incubated in 900µL of recovery media at 40°C for a duration of 24h or 48h. In contrast, the 100µL transfected LMH cells were incubated in 900µL recovery medium at 37°C with 5% $CO_2$ for 24h.

### 3.3.2 **Transfection using Neon Transfection System** (Thermo Fisher Scientific, USA)

100 µL of cells (1x10$^7$ cells/mL) were mixed with 20 µL of plasmid (220 ng/µL) at room temperature and then subjected to electroporation using the Neon Transfection System (Thermo Fisher Scientific, USA). Following electroporation with the specific pulse parameters (Table 5), the cells were transferred in a 12-well cell culture plate (Figure 6) with 900 µL of recovery medium. For transfection of *H. meleagridis* cells, the recovery medium consisted of an overnight culture of *E. coli* DH5α, grown in medium M199 with 15% FBS and diluted in 1:1 ratio with M199 with 15% FBS. On the other hand, the normal culture medium was used as the recovery medium for transfection of *T. gallinae* or LMH. Each cell type was tested with different program settings (Table 5) and with different plasmid combinations (Table 6).

**Table 5. List of the transfection programs used.** Program settings voltage, pulse quantity, pulse length.

| Program setting | *H. meleagridis* | *T. gallinae* | LMH |
|---|---|---|---|
| 1005V, 2Puls, 35ms | **Yes** | **Yes** | **Yes** |
| 1100V, 2Puls; 20ms | **Yes** | **Yes** | **Yes** |
| 1350V, 1Puls, 30ms | **Yes** | **Yes** | **Yes** |
| 1600V, 3Puls, 10ms | **Yes** | **Yes** | No |
| 1680V, 1Puls, 20ms | **Yes** | **Yes** | No |
| 1800V, 1Puls, 10ms | **Yes** | No | No |
| 2000V, 1Puls, 10ms | **Yes** | No | No |
| 2200V, 1Puls, 10ms | **Yes** | No | No |
| 2500V, 1Puls, 10ms | **Yes** | No | No |
| 2500V, 2Puls, 10ms | **Yes** | No | No |

**Table 6. List of the plasmids tested with various organisms and cell lines**

| Plasmid | *H. meleagridis* | *T. gallinae* | LMH |
|---|---|---|---|
| pUC19_g6464_*AcGFP*opt | **Yes** | No | No |
| pBlue_g6464_*AcGFP*opt | **Yes** | **Yes** | No |
| pBlue_g7445_*eGFP*opt | **Yes** | **Yes** | **Yes** |
| pBlue_Tricho_*eGFP*opt | **Yes** | **Yes** | **Yes** |
| pGFPuv | **Yes** | **Yes** | **Yes** |
| p*AcGFP*1-C1 | **Yes** | **Yes** | **Yes** |

### 3.3.3 Lipofection

For the lipofection, liposome-based transfection, TransFast™ Transfection Reagent (Promega, GmbH, USA) was used according to the specified test protocol for suspension cells. Briefly, the incubation process involved mixing the TransFast™ Transfection Reagent / plasmid mix with $1x10^6$ cells in 0.5mL of medium, followed by an incubation period of 1 hour at 37°C. After the incubation time, 5mL of recovery medium was added to the mix. The experimental design comprised four different concentration of plasmid that were tested. Specifically, *H. meleagridis*

were tested with the plasmids pUC19_*g6464_AcGFP*opt, pBlue_*g6464_AcGFP*opt, pBlueScript_Tricho_*eGFP*opt and *T. gallinae* with the plasmid pBlue_Tricho_*eGFP*opt.

## 3.4 GFP detection

### 3.4.1 Fluorescent microscope

The 12-well cell culture plates were inspected with an inverse fluorescence microscope (Leica Microsystems GmbH, Germany) using a high quality yellow GFP bandpass emission filter set: exciter U-HQ500/20x, beamsplitter U-Q515LP, emitter U-HQ535/30m.

### 3.4.2 Protein detection

#### 3.4.2.1 Protein isolation

For protein isolation samples were collected and pooled from *H. meleagridis* and *T. gallinae* cells after 24h and 48h of incubation. Samples of LMH cells were only collected and pooled after 24 hours incubation time. The cells were collected by centrifugation at 1800xg for 10 minutes. The pellet obtained was resuspended in 600 µl of lysis buffer (Triton x-100 buffer; 1% Triton x-100, 50mM Tris/Hcl, 150mMNaCl, 1mM EDTA, 0,5% SDS; pH 7.4) with cOmplete™ proteinase inhibitor cocktail (Merck, Austria). The cell solution was placed in a Tissuelyser (2x 2 minutes, 250 Hz) (Qiagen GmbH, Austria) and then centrifuged (10 min, 10000xg, 4° C). The supernatant containing the purified proteins was transferred to new Eppendorf tubes and stored at -80°C until further processing.

#### 3.4.2.2 SDS-PAGE

To separate the proteins, a 12% polyacrylamide SDS gel was utilized. Firstly, 20 µl of protein was mixed with 5 µl of 6x SDS loading buffer, denatured at 95° C for 5 minutes and centrifuged at maximum speed for 2-3 minutes to separate any particulates or aggregates. Next, 20µL sample were loaded per well and proteins were separated for 1.5h with 120V. Two SDS-PAGE gels were made for each sample. The first gel was stained by being covered with Coomassie Brilliant Blue (Bio-Rad Laboratories, Inc., USA) for 1 hour and de-stained with water overnight at room temperature. The second gel was used for Western blotting.

Proteins from an overnight culture of *E. coli* with pGFPuv plasmid served as a positive control in 1:10 dilution (dilution rate was determined previously through serial dilution, data not shown).

#### 3.4.2.3 Western blot

The second polyacrylamide SDS gel was electro transferred to a PVDF membrane using a Trans-Blot Turbo RTA kit (Trans-Blot Turbo RTA kit Mini PVDF Transfer Kit; Bio-Rad Laboratories, Inc., USA) and the Trans-Blot Turbo Transfer System (Bio-Rad Laboratories, Inc., USA) using the Mixed MW program (2,5A; 25V; 7min).

The membrane was blocked overnight in a 3% skim milk/ 0.05% Tween/ 1xPBS solution. The membrane was incubated for one hour with the primary antibody against GFP (1: 5000 ratio; Living Colors® A.v. Monoclonal Antibody (JL-8), Takara Bio Inc., USA) in a 0,1% skim milk/1xPBS solution. The membrane was washed four times with a washing buffer (0.05% Tween /1xPBS solution). The detection was performed with a horseradish peroxidase-conjugate secondary antibody (1:3000 ratio; Goat Anti-Mouse HRP Conjugate, Bio-Rad Laboratories, Inc., USA). The membrane was incubated for one hour with the secondary antibody in a 0,1% skim milk/ 1xPBS solution. Subsequently the membrane was washed four times with the washing buffer. The blot was developed by using Clarity Western ECL Substrate (Bio-Rad Laboratories, Inc., USA) according to the manufacturer's instruction. The proteins on the membrane were visualized with the Gel and Blot imaging system (Bio-Rad Laboratories, Inc., USA) with an exposure time of up to 300 seconds.

### 3.4.3 GFP gene detection with PCR

In addition to Western blotting, detection of a plasmid transfected into *H. meleagridis* cells was performed by PCR for the GFP gene. For that purpose, *H. meleagridis* cells were electroporated with the Neon Transfection System (6 samples of each setting 1800V, 2000V, 2200V) and counted after 24h incubation time. The cells were washed twice with 6mL of M199 medium (Gibco, Thermo Fisher Scientific, USA) and centrifuged for 5 minutes at 200xg. The resulting sample was divided into two parts. The first part was subjected to an additional wash with 3mL medium and centrifuged for 10 minutes at 200xg. The supernatant was discarded, and the pellet was used for the DNA extraction.

The other part was treated with DNase 1 (Bio-Rad Laboratories, Inc., USA). The cell suspension was centrifuged (5 min at 200xg), and the supernatant discarded. The pellet was re-suspended in 200µL M199 medium with 10µL DNase 1 + 4µL 1M $MgCl_2$. The DNase 1 suspension was incubated for 15 min at room temperature before inactivating with 20µL 0,5M EDTA. The cells were then centrifuged (10 min at 1800xg) and the pellet was used for DNA extraction.

In addition, four control samples were prepared (1) none electroporated cells with the DNase 1 treatment (2) none electroporated, washed, but without DNase 1 treatment (3) electroporated cells without washing and no DNase 1 treatment, (4) 200µL medium + 20µL plasmid (220ng) + DNase 1 treatment.

The DNA from all cell pellets were isolated using the DNeasy Blood & Tissue Kit (Qiagen GmbH, Austria). Two specific primers for the optimized *eGFP* (*eGFP*opt_Fw: TACCAGTTCCATGGCCAACA; *eGFP*opt_Rw: TCCTAATGTAATTCCTGCTGCTGT) were designed for the PCR. The *eGFP* specific fragments were PCR amplified with the NEB Q5 High-Fidelity 2x Master Mix (New England BioLabs Inc., USA) using 3µL DNA (50ng/µL). The PCR was carried out according to the NEB protocol with the recommended annealing temperature of 59,6°C (https://tmcalculator.neb.com/#!/main) (Thermal program: 98°C 1 min; 35 cycles of 98°C 10s + 59,6°C 15s + 72°C 1:40 min; 72°C 5min)

# 4 Results

## 4.1 Plasmids

### 4.1.1 Plasmid construction and isolation

Initial transfection experiments utilizing an *AcGFP* gene were unsuccessful. Therefore, altogether six plasmids were designed, differing in the plasmid backbone, promoter/terminator and GFP sequence. Plasmid pBlue_Tricho_*AcGFP* was synthesised (Eurofins, Germany), while the other five plasmids were assembled using the Gibson assembly method (Table 7). Each plasmid was extracted from a 600 mL overnight *E. coli* culture, resulting in a total yield of 1-2.5 mg DNA per sample. These yields were within the estimated range for the MaxiPrep kit.

**Table 7. Overview of the different plasmids:** Backbone plasmid, promoter/ terminator sequence and GFP genes

| Plasmid name | Backbone plasmid | Promoter/Terminator | GFP gene |
|---|---|---|---|
| pUC19_*g6464_AcGFP*opt | pUC19 | *g6464* (40S ribosomal protein S4, X isoform) (*H. meleagridis)* | *AcGFP*opt |
| pBlue_*g6464_AcGFP*opt | pBluescript II sk (+) | *g6464* (40S ribosomal protein S4, X isoform) (*H. meleagridis)* | *AcGFP*opt |
| pBlue_Tricho_*eGFP*opt | pBluescript II sk (+) | succinyl CoA synthetase αα subunit (*T. vaginalis*) | *eGFP*opt |
| pBlue_*g7554_eGFP*opt | pBluescript II sk (+) | *g7554* (*H. meleagridis)* | *eGFP*opt |
| pBlue_Tricho_*AcGFP* | pBluescript II sk (+) | succinyl CoA synthetase αα subunit (*T. vaginalis*) | *AcGFP* |
| pBlue_*g7554_AcGFP* | pBluescript II sk (+) | *g7554*(*H. meleagridis)* | *AcGFP* |

### 4.1.2 Restriction Digest

In order to confirm the accuracy of the plasmid constructs, incubation with restriction endonucleases were performed after each assembly reaction. The results revealed successful cloning for all plasmids; the restriction digest showed the predicted sizes on the gels and matched therefore the designed construct (Figure 7) (additional results are shown in Annex 9).

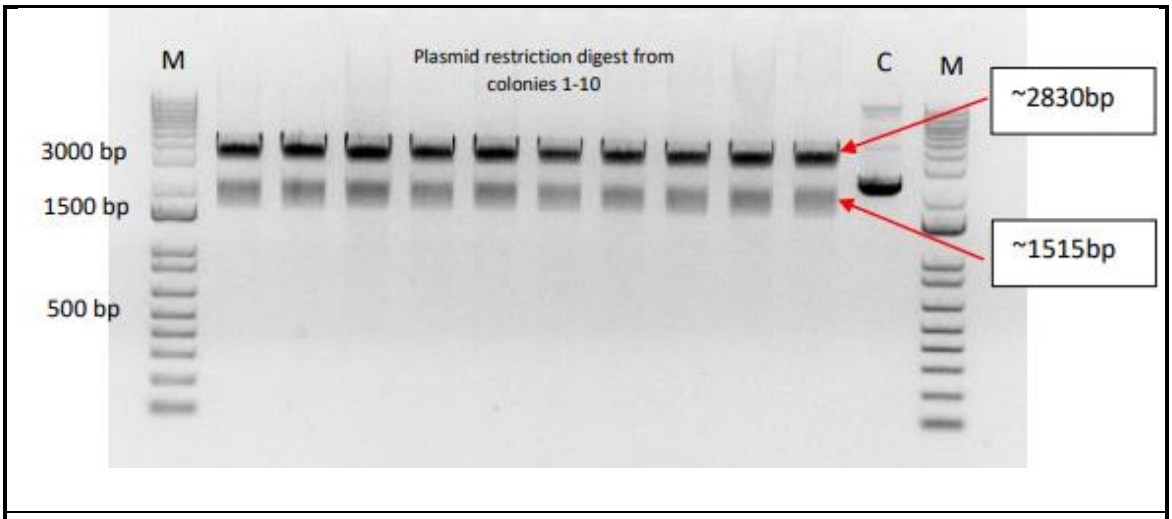



**Figure 7. A result example of restriction endonuclease digestion.** Reactions with EcoRV-HF and PstI-HF of 10 selected pUC19_*g6464_AcGFP*opt clones were separated on an 1% agarose gel. Expected bands for a correct construct were 1515bp and 2830bp. C= control undigested plasmid PUC19_HisPromS4_*AcGFP*_HisTermS4 M= 1kb Plus DNA Ladder (Invitrogen, ThermoFisher Scientific Inc., USA)

### 4.1.3 Sequencing

Sequencing results showed that the sequences of the constructed plasmid matched the designed sequence and that the *AcGFP* gene was in the correct reading frame foreseeing a correct protein translation (Figure 8) (All sequencing results are shown in Annex 10).

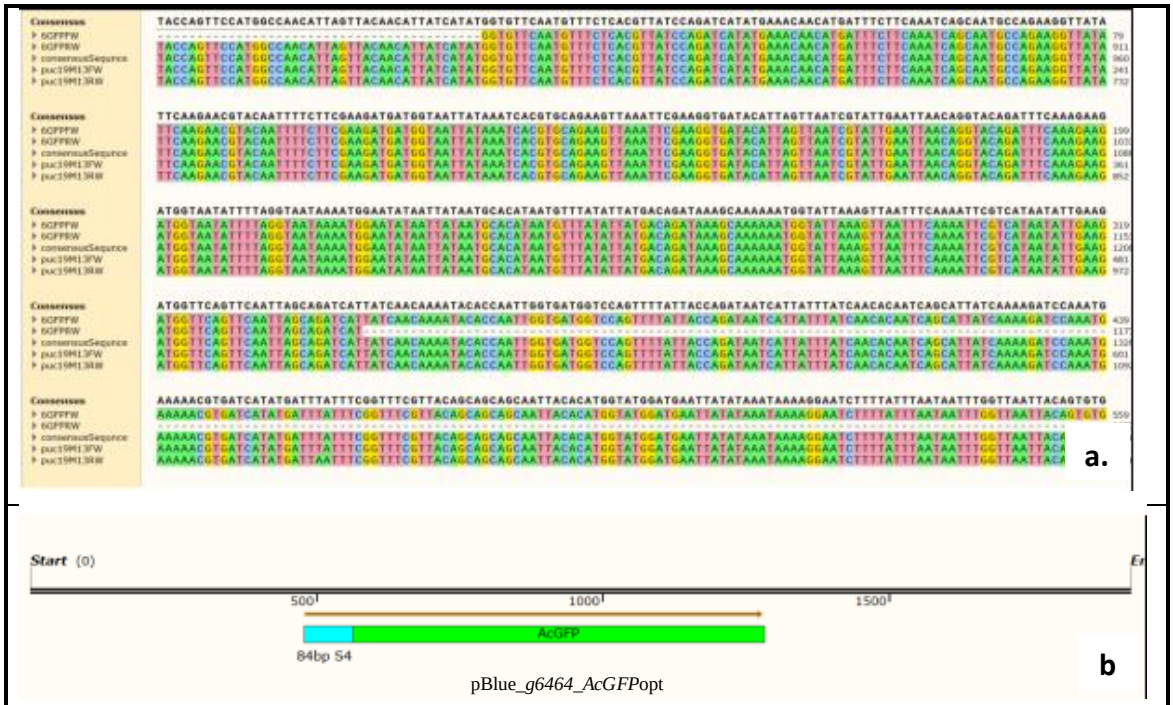



**Figure 8. Sequencing result example. (a)** A section of the sequence alignment of pUC19_*g6464_AcGFP*opt plasmid as representation for the correct sequence of all plasmids. The consensus sequence of the alignment is shown in the first row, followed by the individual sequencing results. The first two sequences have their starting point in the GFP gene (Primer GFP_FW and GFP_RW). While the third row is the designed sequence of the insert as template. The last two sequences have their starting point outside of the insert region (primer M13-21F and M13-29R). A total of ~1900bp of the construct was sequenced **(b)** Cartoon of the open reading frame of pBlue_*g6464_AcGFP*opt plasmid as an example for all plasmids. The sequence showed that the 84bp long *g6464* gene overhang and the *AcGFP* gene were in frame and a complete protein translation of the gene is possible.

## 4.2 Transfection

### 4.2.1 *E. coli* Pulser cell survival rate

Initial electroporation experiments showed that almost all *H. meleagridis* cells survived using the recommended voltage previously published for *T. vaginalis* transfection [48]. It was hypothesized that a higher voltage would result in a higher transfection rate; therefore, the cell survival rate was determined to obtain the highest possible electrical current where cells still survive (Table 8). Cell survival rate dropped significantly after 1700V. Electroporation parameters of 1900V or above produced only arcing in the cuvette. Therefore, further experiments were carried out with the electrical current range from 1200V to 1700V.

**Table 8**. **Survival rate of *H. meleagridis* with the *E. coli* Pulser.** As the voltage increased, the survival rate of the cells decreased.

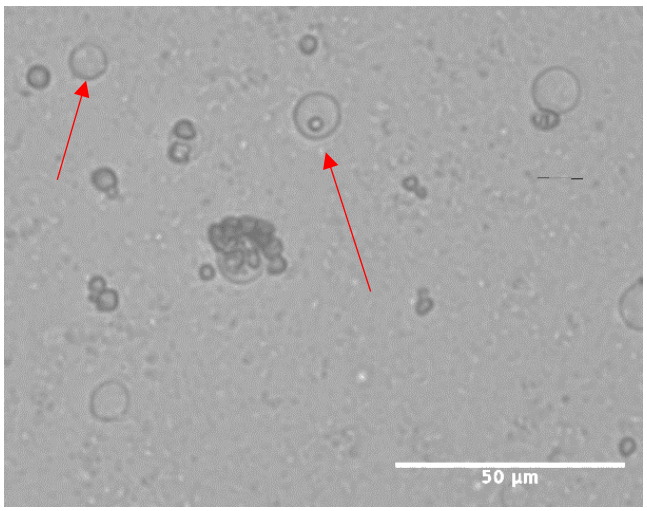
| Voltage settings | Cells numbers in 100µL | Cells numbers after electroporation (in 1mL) | Survival rate |
|---|---|---|---|
| 1200V | $1x10^5$ | $\sim0,15x10^5$ | 15% |
| 1400V | $1x10^5$ | $\sim0,12x10^5$ | 12% |
| 1600V | $1x10^5$ | $\sim0,075x10^5$ | 7,5% |
| 1700V | $1x10^5$ | $\sim0,005x10^5$ | 0,5% |
| 1800V | $1x10^5$ | $\sim0,025x10^5$ | 0,25% |
| 1900V | $1x10^5$ | 0 | 0 |

### 4.2.2 Neon transfection system cell survival rate

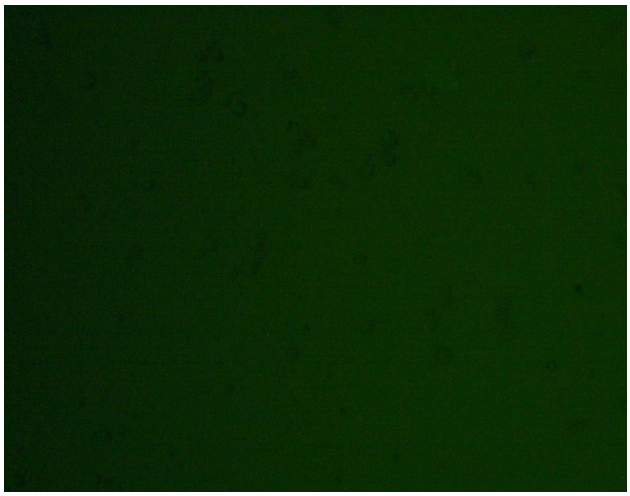
The transfection with the Neon system showed that the cells could survive higher voltage settings compared to the *E. coli* Pulser. Therefore, the transfection was tested with electrical current range from 1005V to 2500V. A survival rate of 4-20 % was achieved for *H. meleagridis* transfected with the neon transfection system. Remarkably, even at the maximum voltage (2500V) 4 % of the cells were able to survive.

## 4.3 Fluorescence microscopy

After each transfection experiment, the cells were observed under the fluorescent microscope to control cell morphology, survival and green GFP signal generation. Live *H. meleagridis* cells could be observed after each transfection experiment under the microscope (Figure 9a) after 24h and 48h, however *H. meleagridis* cells failed to produce a green GFP signal (Figure 9b).




a



b

**Figure 9. Light and fluorescence micrographs of *H. meleagridis* cells** transfected with the pBlue_*g7554_eGFP*opt plasmid using the neon transfection system after 24h incubation. **(a)** Live *H. meleagridis* cells (arrows) were observed on the phase-contrast image; however, no GFP signal was detected **(b)** on the same image with fluorescence filter.
Scale bar = 50µm

*T. gallinae* cells, as a control were also transfected with various GFP containing plasmids to investigate whether the *H. meleagridis* promotor also functions in *T. gallinae* cells. Live *T. gallinae* cells could be observed after each transfection experiment under the light microscope (Figure 10a) after 24h and 48h, however, no GFP signal was detected in the *T. gallinae* culture with the fluorescent filter (Figure 10b). The result showed that either the transfection programs or the plasmids used for *T. gallinae* were not suitable.

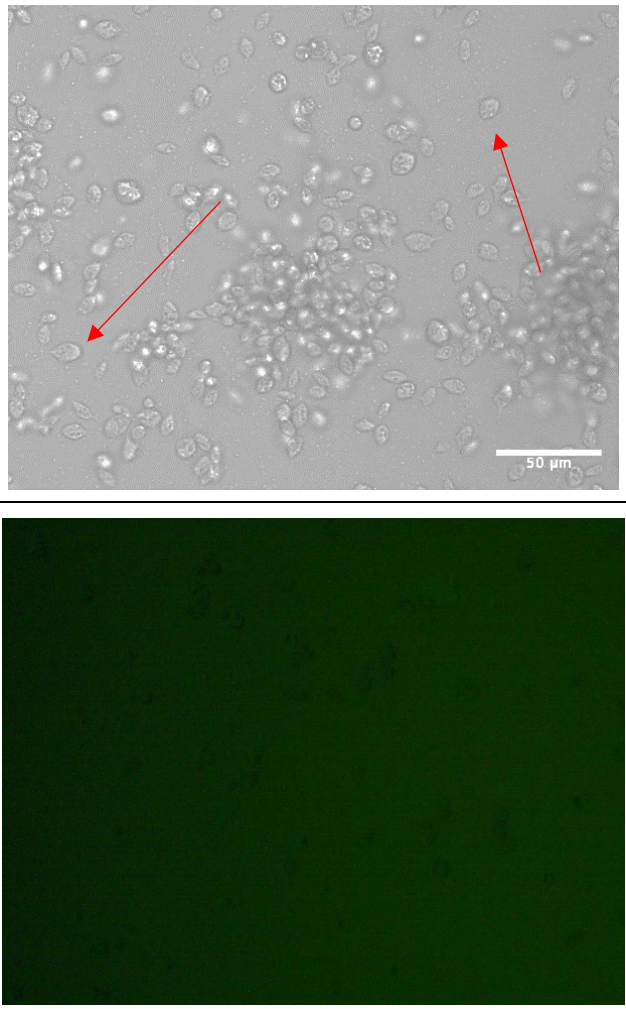
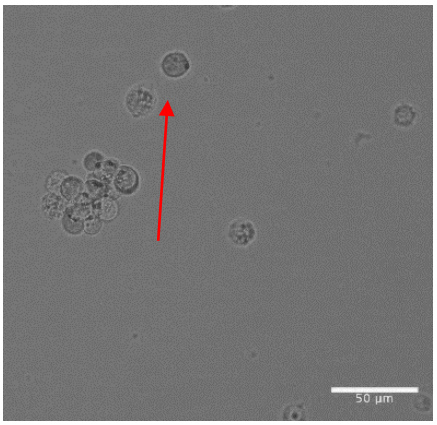
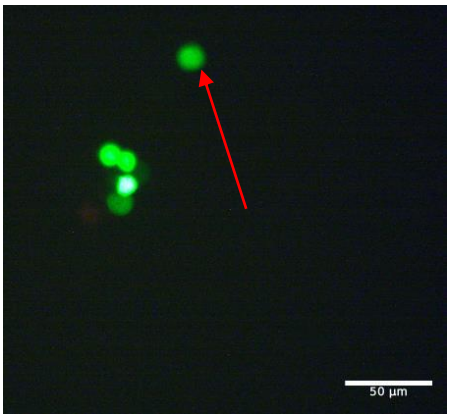



**Figure 10. Light and fluorescence micrographs of *T. gallinae* cells** transfected with the pBlue_Tricho_*eGFP*opt plasmid using the neon transfection system after 24h incubation. **(a)** Live *T. gallinae* cells (arrows) on the phase-contrast image. **(b)** Same image as (a) with fluorescent filter. No GFP signal was detected.
Scale bar = 50µm

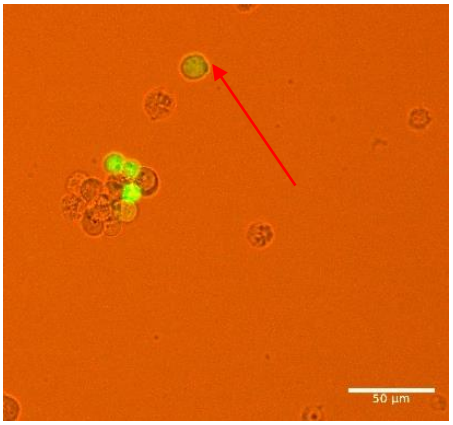
As a further control for the transfection methods, LMH cells were also tested. The p*AcGFP*1-C1 plasmid used for transfecting LMH cells produced green GFP signal after 24h incubation (Figure 11). The results could be reproduced with both the neon transfection and *E. coli* Pulser electroporator systems.




a




b




c

**Figure 11. Light and fluorescence micrographs of LMH cells** transfected with the p*AcGFP*1-C1 plasmid using the neon transfection system after 24h incubation. **(a)** Live LMH cells (arrows) on the phase-contrast image. **(b)** LMH cells emitting a GFP signal (arrows). **(c)** Overlay of phase-contrast and fluorescence light images.
Scale bar = 50µm

## 4.4 Western blot

Western blot assay was used as an alternative detection method for the expressed GFP protein. Western blot using the anti-GFP primary monoclonal antibody showed a GFP band only for the transformed *E. coli* cells with the expected size of 30kD for the GFPuv protein (Figure 12.1.-12.3.). In the other protein samples, no GFP protein was detected.

The Coomassie Blue staining of all corresponding SDS-PAGE gels showed proteins with various sizes (Figure 12) without a clear distinct band for the GFP protein at 30 kDa of size. To achieve a uniform loading without measuring the exact protein amounts, harvested cell numbers were calculated prior to the loading (Table 9). Extracted proteins from $2.2x10^3$-$9.6x10^4$ (*H. meleagridis, T. gallinea,* LMH) and $1.3x10^8$ from *E. coli* cells were loaded onto the gels. It was observed that 100% of the *E. coli* cells expressed the GFPuv protein. About 12,5% of the counted LMH cells produced GFP corresponding to about $5x10^2$ LMH cells.

Calculation:

1. Total cell number in the SDS-PAGE ≙ cell number in 20µl of the total cell count in 600µl

=> $\frac{total\ cell\ count}{600} * 20$

2. Fluorescent cells in the gel => Percentage of fluorescent cells was counted and calculated

$flourescent\ cells\ \% = \frac{counted\ \ fluorescent\ cells}{counted\ total\ \ cell\ number}$ => $total\ cell\ number\ in\ the\ SDS\_Page * percentage\ of\ fluorescent\ cells$ )

**Table 9. Overview of the calculated cell numbers used for loading onto the SDS-PAGE gel.** Total harvested live cell numbers after 48h of incubation in a 12-well cell culture plate, if not otherwise mentioned in the sample name; cell number loaded onto a single well of the SDS-PAGE gel.

| Sample name | Total cell number after harvesting | Total cell number in the SDS-PAGE gel | fluorescent cells | Total number of fluorescent cells in the gel |
|---|---|---|---|---|
| Transformed ***E. coli*** with the pGFPuvplasmid from a 10mL cell culture diluted 1:10 | ~$39x10^9$ | ~-$1.3x10^8$ | 100% | ~$1.3x10^8$ |
| Transfected **LMH** cells with the p*AcGFP*1-C1 plasmid 24h incubation | $1.3x10^5$ | $4,3x10^3$ | ~12,5% | ~$5x10^2$ |
| ***H. meleagridis*** cells transfected with various plasmid using the ***E. coli*** **Pulser** | | | | |
| pBlue_*g6464_AcGFP*opt, | $0,68x10^6$ | $2,27x10^4$ | 0% | 0 |
| pBlue_Tricho_*eGFP*opt, | $0,91x10^6$ | $3,03x10^4$ | 0% | 0 |
| pBlue_g7445_*AcGFP* | $1,06x10^6$ | $3,53x10^4$ | 0% | 0 |
| pBlue_Tricho_*AcGFP* | $0,84x10^6$ | $2,80x10^4$ | 0% | 0 |
| ***T. gallinae*** cells transfected with various plasmid using the ***E. coli*** **Pulser** | | | | |
| pBlue_*g6464_AcGFP*opt, | $0,95x10^6$ | $3,17x10^4$ | 0% | 0 |
| pBlue_Tricho_*eGFP*opt, | $1,2x10^6$ | $4,00x10^4$ | 0% | 0 |
| pBlue_g7445_*AcGFP* | $2,67x10^6$ | $8,90x10^4$ | 0% | 0 |
| pBlue_Tricho_*AcGFP* | $2,88x10^6$ | $9,60x10^4$ | 0% | 0 |
| ***H. meleagridis*** cells transfected with various plasmid using the **Neon transfection system** | | | | |
| p*AcGFP*1-C1 | $0,082x10^6$ | $0,27x10^4$ | 0% | 0 |
| pBlue_*g6464_AcGFP*opt | $0,15x10^6$ | $0,50x10^4$ | 0% | 0 |
| pBlue_g7445_*eGFP*opt-48h | $0,25x10^6$ | $0,83x10^4$ | 0% | 0 |
| pBlue_Tricho_*eGFP*opt | $0,23x10^6$ | $0,77x10^4$ | 0% | 0 |
| pBlue_g7445_*eGFP*opt-24h | $0,067x10^6$ | $0,22x10^4$ | 0% | 0 |
| ***T. gallinae*** cells transfected with various plasmid using the **Neon transfection system** | | | | |
| p*AcGFP*1-C1 | $2,36x10^6$ | $7,87x10^4$ | 0% | 0 |
| pBlue_*g6464_AcGFP*opt | $0,56x10^6$ | $1,87x10^4$ | 0% | 0 |
| pBlue_ Tricho_*eGFP*opt-48h | $0,49x10^6$ | $1,63x10^4$ | 0% | 0 |
| pBlue_ g7445_*eGFP*opt | $0,59x10^6$ | $1,97x10^4$ | 0% | 0 |
| pBlue_ Tricho_*eGFP*opt-24h | $2,45x10^6$ | $8,17x10^4$ | 0% | 0 |
| **Lipofection** of *H. meleagridis* and *T. gallinae* with various plasmids | | | | |
| Histomonas transfected with pUC19_*g6464_AcGFP*opt | $1,06x10^6$ | $3,53x10^4$ | 0% | 0 |
| Histomonas transfected with pBlue_*g6464_AcGFP*opt | $0,76x10^6$ | $2,53x10^4$ | 0% | 0 |
| Histomonas transfected with pBlue_*g7554_eGFP*opt | $0,89x10^6$ | $2,97x10^4$ | 0% | 0 |
| Trichomonas transfected with pBlue_ Tricho_*eGFP*opt | $1,34x10^6$ | $4,47x10^4$ | 0% | 0 |

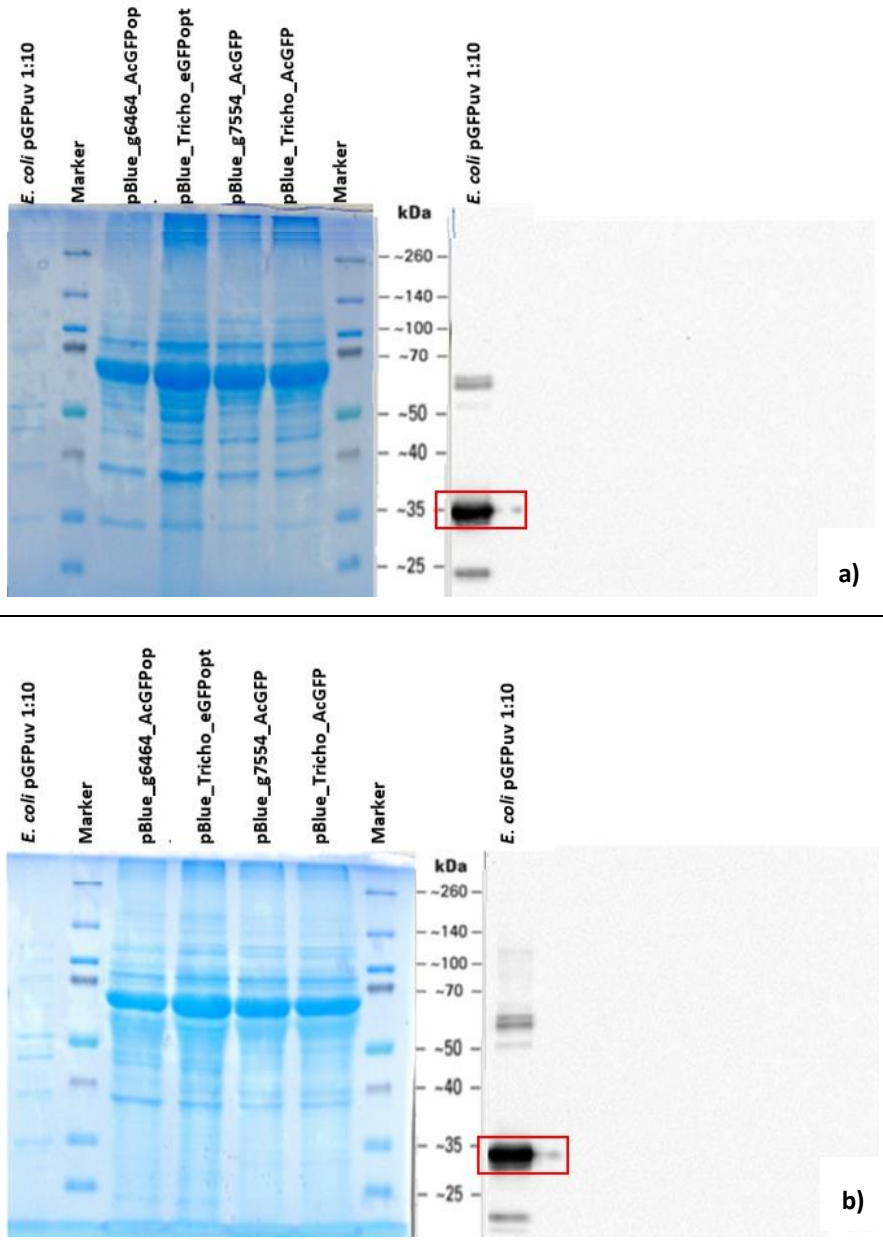
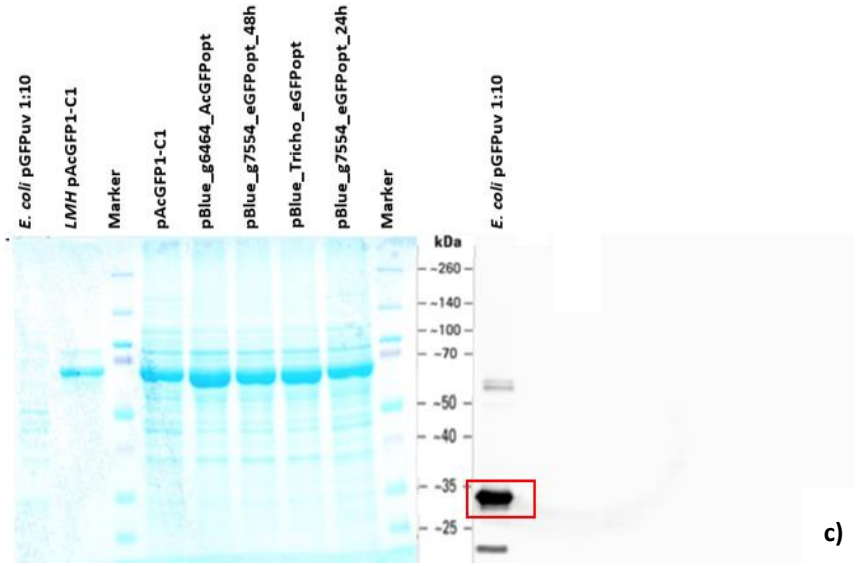



**Figure 12.1. Coomassie blue stained SDS-PAGE and the corresponding western blot result of proteins extracted from the transfected cells.** On the left-side proteins separated in a 12% SDS-PAGE and stained with Coomassie blue are presented. On the right-side, the western blot result with the anti-GFP primary antibody (Living Colors® A.v. Monoclonal Antibody (JL-8), Takara Bio Inc., USA) diluted 1:5000, is shown. The red boxed band represents the expressed GFPuv protein from the *E. coli* cells transfected with the pGFPuv plasmid. The other detected bands could be dimers or truncated GFP proteins. In no other sample GFP protein was detected**. (a)** *H. meleagridis* cells transfected with various plasmids using the *E. coli* Pulser **(b)** *T. gallinae* cells transfected using the *E. coli* Pulser with several plasmids



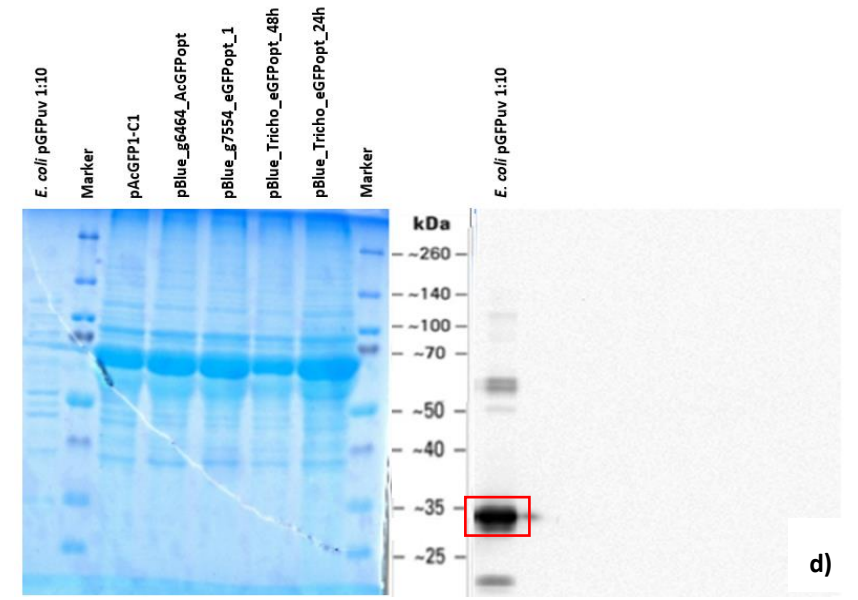




**Figure 12.2. Coomassie blue stained SDS-PAGE and the corresponding western blot of proteins extracted from the transfected cells.** On the left-side proteins separated in a 12% SDS-PAGE and stained with Coomassie blue are presented. On the right-side, the western blot result with the anti-GFP primary antibody (Living Colors® A.v. Monoclonal Antibody (JL-8), Takara Bio Inc., USA) diluted 1:5000, is shown. The red boxed band represents the expressed GFPuv protein from the *E. coli* cells transfected with the pGFPuv plasmid. The other detected bands could be dimers or truncated GFP proteins. In no other sample GFP protein was detected**. (c)** *H. meleagridis* cells transfected with various plasmids using neon transfection system **(d)** *T. gallinae* cells transfected with several plasmids using the neon transfection system

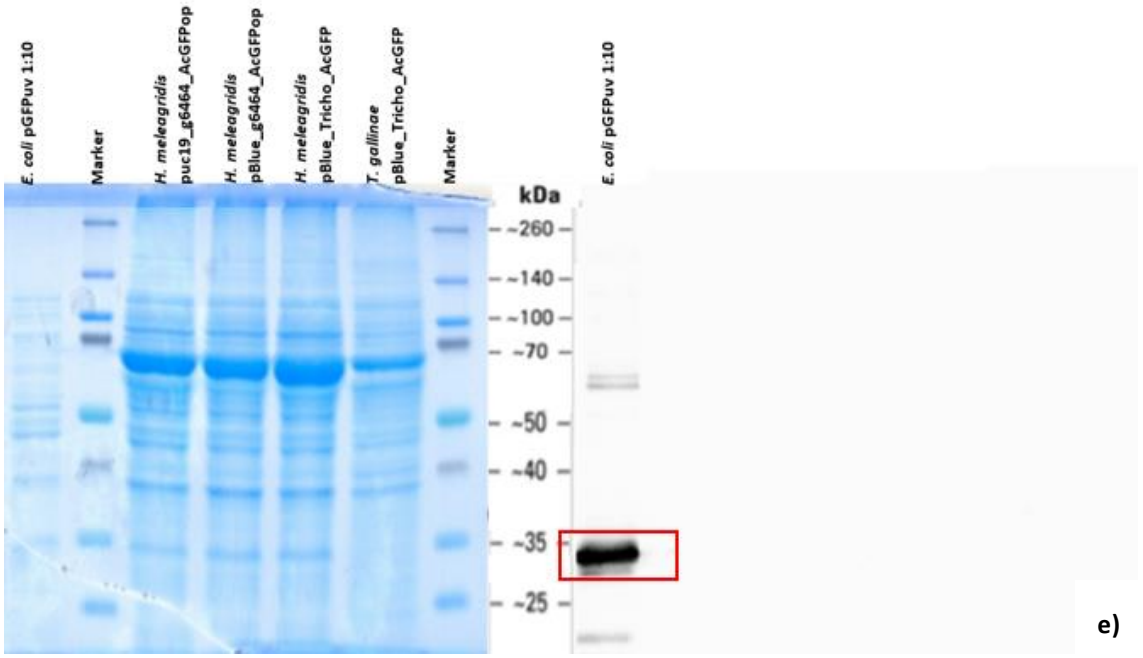



**Figure 12.3. Coomassie blue stained SDS-PAGE and the corresponding western blot result of proteins extracted from the transfected cells.** On the left-side proteins separated in a 12% SDS-PAGE and stained with Coomassie blue are presented. On the right-side, the western blot result with the anti-GFP primary antibody (Living Colors® A.v. Monoclonal Antibody (JL-8), Takara Bio Inc., USA) diluted 1:5000, is shown. The red boxed band represents the expressed GFPuv protein from the *E. coli* cells transfected with the pGFPuv plasmid. The other detected bands could be dimers or truncated GFP proteins. In no other sample GFP protein was detected**. (e)** Lipofection of *H. meleagridis* and *T. gallinae* with various plasmids.

## 4.5 GFP gene detection with PCR

A PCR detection method was used to determine whether the plasmid had entered the *H. meleagridis* cells during transfection. The total cell number of the transfected samples after 24h were 1800V=0,35x$10^7$cells; 2000V=0,30x$10^7$cells; 2200V=0,27x$10^7$ cells. The GFP specific PCR product with an expected size of 539bp was detected in all samples. (Figure 13).

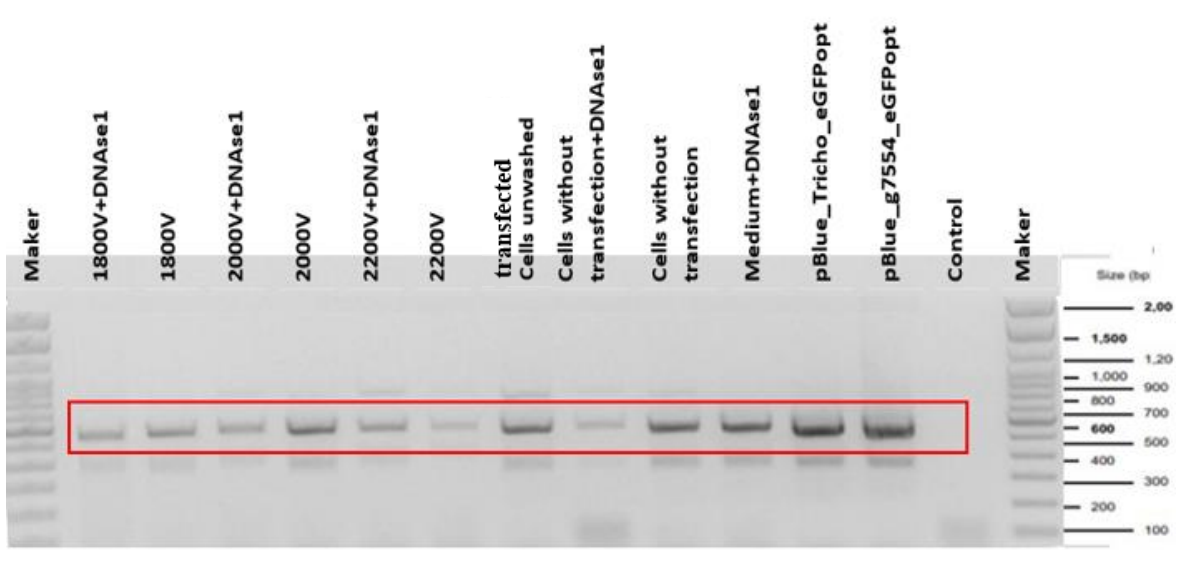



**Figure 13. Agarose gel of the GFP-PCR for the plasmid detection.** DNA extraction of various *H. meleagridis* samples with and without neon transfection treatment with various voltage setting (1800V, 2000V or 2200V) after 24h incubation time; with and without DNase1 treatment. pBlue_Tricho_*eGFP*opt plasmid or pBlue_*g7554_eGFP*opt plasmid were used for the transfection. The expected product size (539bp) is labelled with a red box. Cell without transfection represents samples where plasmid and clles were mixed without undergoing the transfection procedure. Control= PCR without a DNA template; Maker =100bp Plus DNA Ladder (Invirtogen, ThermoFisher scientific Inc, USA)

# 5 Discussion

*Histomonas meleagridis*, the causative agent of histomonosis or blackhead disease, has been known since the last century [1]. Histomonosis is of great significance in the poultry industry due to its impact on the health and welfare of infected flocks in coincidence with substantial economic losses [2]. Several aspects of the pathogen were well studied,

including life cycle [2,3], host-pathogen interaction [54] and disease control [4]. However, until today, no genetic manipulation of this parasite is reported.
The actual work, focused at establishing a protocol for incorporating a foreign DNA in the form of bacterial plasmids into *H. meleagridis* cells. As a proof of concept, transfection and the expression of a recombinant protein by *H. meleagridis* were investigated with the aim to establish a procedure that could serve as a basis for future genetic manipulation experiments such as CRISPR/Cas 9 modifications [38].

Six plasmids differing in a promoter sequence, terminator sequence, GFP gene and plasmid backbone were successfully assembled and tested in *H. meleagridis, T. gallinae* and LMH cells.

Plasmid vectors also known as "vehicles" used to introduce recombinant DNA into a host cell are an essential part of molecular cloning. The most important characteristics of such a cloning vector are: the ability to replicate itself in *E. coli* (origin of replication (ORI)), a selectable characteristics (such as antibiotic resistance), a multiple recognition site for restriction endonucleases or a multiple cloning site (MCS) [25] and a promoter along with regulatory control. Two plasmid vectors were utilised as backbone for cloning, a pUC19 plasmid and a pBlueScript II SK (+) plasmids. Both plasmids are similar, having a lac promoter, an ampicillin resistance gene and being high copy number cloning vectors [55,56]. The pUC19 plasmid was supplied with the Gibson assembly kit. However, a shift to a pBlueScript II SK (+) plasmid was made, to achieve the maximum similarity to the published *T. vaginalis* expression vector [48], which served as reference.

The green fluorescent protein (GFP) is a commonly used reporter gene. Three different GFP genes in total were tested in this work. At first, a synthesized, *H. meleagridis*-codon optimised *AcGFP*1 gene from a previous experiment (University of Veterinary Medicine Vienna, unpublished data) was used. This gene was originally derived from a jellyfish, *Aequorea coerulescens* [57]. The first transfection attempt with the codon optimised *AcGFP*1 and electroporation did not yield a fluorescence signal in the cells under the microscope. Therefore, an alternative GFP gene was explored to ensure that the hurdle is not the fluorescence signal strength of the protein. A single point mutant enhanced GFP (*eGFP*) protein with increased fluorescence and higher brightness than *AcGFP* protein was explored as an alternative. The calculated brightness values, a product of the extinction

coefficient and quantum yield [58] was 33.5 for the protein *eGFP* [59] and 27.5 for *AcGFP*1 protein [57]. Like *AcGFP* gene, the *eGFP* sequence was also codon optimised for *H. meleagridis*. However, after transfection yet again no fluorescence signal was detected in the *eGFP* transfected *H. meleagridis* cells.

To exclude the possibility that the codon optimisation had an effect on the GFP protein, an unmodified *AcGFP*1 gene from the commercially available p*AcGFP*1-C1 plasmid was tested. This commercial *AcGFP*1 protein could successfully be detected in transfected LMH cells under the fluorescence microscope, however not in *H. meleagridis* cells.

In addition to the three GFP genes, three alternative promoter/terminator sequences were examined. The entire genome of *H. meleagridis* has been published only recently and could be deciphered, making it possible to identify organism specific DNA regions [7]. Despite this progress, the specific promoter sequences for *H. meleagridis* are yet to be identified. In general, promoter sequences are regions located upstream of a gene and are essential areas for gene regulation. A typical promoter region contains a core promoter sequence as well as regulatory domains. The presence of known promoter components, such as TATA box, GC-box, CCAAT-box, BRE, and INR box, indicate the structure of a promoter [60]. Computer programs, so called promoter prediction programs, can identify a specific promoter sequence by using information about (1) the genome structure, (2) the know promoter sequences, (3) transcription factors, (4) epigenic markers, etc. However, since the regulatory mechanisms of *H. meleagridis* are poorly understood these computer programs can give numerous false positives because these programs are depending on the understanding of the organism [61,62]. In this study, no computer program was used to search for promoters in the *H. meleagridis* genome. The promoter region was manually determined to be the sequence upstream of an identified gene that spans a TATA box. It is unknown how long the promoter sequence must be or whether it must contain another promoter component beside the TATA box like a GC-box, CCAAT-box etc. In some cases, viral promoters can replace organism-specific promoter sequences in eukaryotes. In mammalian cells, the CMV or SV40 promoters are commonly used. Both are strong promoters derived from cytomegalovirus (CMV) respectively simian virus 40 (SV40) [63]. The commercial plasmid p*AcGFP*1-C1 contains a CMV promoter [64]. This plasmid was tested in all cell types; however, the GFP protein was only expressed and could be detected in LMH cells. The reason for this result is probably that the LMH cells are mammalian cells therefore the CMV

promoter was expected to work in these cells. In addition, the promoters mentioned before were tested in other protozoa, including *Entamoeba histolytica* [63] and *Toxoplasma gondii* [64] without success, indicating that *H. meleagridis* and *T. gallinae* might need organism-specific promoters as these microorganisms belong to a different phylum than LMH cells. Therefore, promoter sequences specific to *H. meleagridis* were designed and tested. The general design for the promoter/terminator sequence was based on the published *T. vaginalis* plasmid, where 1660 nucleotides upstream and 449 nucleotides downstream of the α-succinyl CoA synthetase B gene served as promoter and as terminator sequences, respectively [48]. The gene with the hypothetical promoter sequence was selected based on the *H. meleagridis* transcriptome analysis data [52]. In general, a suitable gene should ideally be expressed in every cell and during the entire cell cycle. Based on these criteria, the first promoter/terminator sequence was taken from the 40S ribosomal protein S4 encoding *g6464* gene (NCBI: GenBank: WSYJ01000142; locus_tag= "GO595_006457"; product="40S ribosomal protein S4, X isoform"). In the majority of cell types, ribosomal protein genes are amongst the most highly expressed genes. Their products are components of the ribosome and are therefore important in cell development and proliferation [65]. Therefore, it was assumed that the upstream region of the gene *g6464* might be a good candidate for a strong promoter. The second promoter/terminator pair was selected from the upstream and downstream untranslated region of the *g7554* gene. This gene encodes a hypothetical protein (GenBank: WSYJ01000152; locus_tag= "GO595_007545"; product=" hypothetical protein "). The selection was made with the goal of increasing promoter activity and consequently, protein expression, as this gene has one of the highest transcriptional read counts in the virulent *H. meleagridis* strain [52]. The transcript of the *g7554* gene has 19 times more reads than the previously used *g6464* gene (contig03106_HP=*g7554*=6874 reads; contig1818=*g6464*= 355 reads [52]). Since neither of the two specific promoter/terminator sequences for *H. meleagridis* designed yielded a positive result, another sequence was tested. The third tested promoter/terminator sequence was the original sequence from the published *T. vaginalis* plasmid [48]. The hypothesis was that since *T. vaginalis* is taxonomically related to *H. meleagridis* [11], with both genera belonging to the Parabasalid order, the same promoter/terminator sequence could function in both microorganisms. The used promoter/terminator sequence belongs to a endogenous alpha succinyl CoA synthetase (SCSB) gene that codes for a hydrogenosomal enzyme [66].

This gene is highly expressed in *T. vaginalis* and it was expected to provide a strong promoter [48,66]. A positive result could not be achieved with this promoter/terminator sequence either. A possible explanation for this is that the sequence was recreated using the database rather than the original plasmid. Unfortunately, we were unable to obtain the published plasmid despite numerous efforts.

Beside various plasmids, also different cell types were used to test the transfection methods.

The main cell type tested was *H. meleagridis* grown *in vitro* as a monoxenic culture, which means in the presence of a single bacterial strain, the *E. coli* DH5α strain [45]. In parallel, the same plasmids were also tested in *T. gallinae* that served as a positive control for the plasmids containing the *T. vaginalis* promoter. Since *T. gallinae* and *T. vaginalis* belong to the same taxonomic family, Trichomonadidae it was hypothesized that the *T. vaginalis*-based regulatory sequence would function in *T. gallinae* too. Furthermore, both taxa are taxonomically closely related to *H. meleagridis* [11]. Despite their close relationship, these microorganisms have distinct differences in genome size and gene count with unknown effect on the transfection efficiency. *T. vaginalis* has the largest genome, (160Mb in size and ~60.000 genes [67]), followed by *T. gallinae*, with a genome size of 54.7Mb genome and ˜22.000 genes [68]. *H. meleagridis* has the smallest genome of these three organisms, measuring 43Mb and encoding ~11.000 genes [7]. Another distinction between the *T. vaginalis* and the *T. gallinae* cells employed in this study was the type of *in vitro* culture. To have more similar condition to the *H. meleagridis* culture a xenic *T. gallinae* culture was used in this work which is different to *T. vaginalis* that was cultured axenically [48]. These differences in culture could be a possible reason that no GFP signal in the *T. gallinae* cells was observed. The effect of axenic and monoxenic culture on the transfection process was studied in *Dictyostelium discoideum*. It was shown that the axenic culture has a higher transfection rate, and the protocol for the monoxenic culture must be modified [69]. There is no data on how other organisms in a culture affect the transfection in protozoa, but the experiment should be repeated with an axenic *T. gallinae* culture.

Additionally, all plasmids were tested in LMH cells, a cell line obtained from an induced hepatocellular carcinoma of a male leghorn chicken [47]. The LMH cells were used here as a positive control for testing transfection methods in general. However, since LMH cells

and *H. meleagridis* represent very different organisms unfortunately no conclusions can be drawn about the functionality of the plasmids after transfection.

One of the most difficult challenges during a transfection experiment is to introduce the plasmids into the cells. In 1987, *Trypanosoma brucei* was the first protozoan parasite to be successfully transfected. Following this, similar transfection systems for other kinetoplastida like *E. histolytica*, *G. lamblia*, and apicomplexan parasites, including many Plasmodium species, could be established [70]. In most of these studies electroporation was used as the preferred transfection method [71–74]. For transfecting *T. vaginalis* an electroporator was also used [48]. For this reason, electroporation was the primary focus in the actual study. Two distinct electroporation systems were evaluated. The first electroporator tested was the *E. coli* Pulser from Bio-Rad, in which only the voltage setting could be adjusted. To establish the optimal voltage setting, it was hypothesized that the increase in voltage would lead to higher transfection efficiency but a decrease in the cell survival rate. Therefore, the survival rate of *H. meleagridis* cells was monitored, and the transfection experiments were carried out by voltage near the minimum of surviving cells. *T. gallinae* was subjected to the same conditions as *H. meleagridis*. However, no GFP signal was detected in either of the parasites, suggesting that the plasmids were not successfully introduced into the cells. One explanation could be the difference in the electroporator device used here and for *T. vaginalis*[48]. Delgadillo et al. (1997) [48] describe the electroporation process using 350V whereas in this work the cells were subjected to 1200V-1700V. This high discrepancy in applied voltage during the transfection experiments was attributed to the difference in machines used (different setting options). It was also hypothesized that the electroporator used here required a significantly higher voltage for successful transformation due to the lack of a capacitance extender. With a change in the capacitance, it is possible to use a lower voltage because a longer and constant electronical pulse can be applied [75]. Considering the *E.coli* Pulser's limited setting options and the requirement for precooled cuvette, which can be a problem because *H. meleagridis* is sensitive to low temperatures [76], a second alternative system was tested. The neon transfection system is an improved electroporator that is commonly used in CRISPR/Cas experiments with eukaryotic cells [77–79]. This device utilizes an easy-to-use pipette system that operates at room temperature. Furthermore, the voltage, pulse length, and pulse quantity can be adjusted individually [77,79]. In order to identify the best possible setting,

several programs were tested. While both electroporation systems gave a detectable fluorescent signal with LMH cells, no fluorescent signal was detected in *H. meleagridis* nor in *T. gallinae* cells. Therefore, it can be assumed that the lack of transfection success was not due to the electroporator devices.

In addition to electroporation, applying lipofection was tested as an alternative transfection method. For lipofection, the plasmids are packaged in lipid droplets that can bind and merge to the cell membrane, since they are both made of a phospholipid bilayer. This allows for the plasmid to be taken up and processed further by the cell [42,43]. Despite the lack of success with lipofection in the current study, this approach should be pursued further, as positive results have already been reported in other protozoan cells, such as *Brucella abortus*, which is considered difficult to transform due to the structure of tits cell wall and outer membrane. By optimizing the lipofection in *Brucella abortus*, a much higher transformation efficiency than electroporation could be obtained [80].

The efficiency of transfection is commonly monitored using GFP reporter genes, as their fluorescent signal can easily be detected via microscope. However, in the current study, only the LMH cells produced a detectable fluorescent signal under the microspore. As an alternative approach, a Western blot technique was employed to investigate whether GFP protein was present in the cells. Western blotting shows high sensitivity when it comes to detecting proteins by using highly specific primary antibodies and the optimised protocol [81,82]. Despite the observed fluorescent signal in LMH cells under the microscope, the Western Blot technique did not detect any AcGFP1 or eGFP protein in the cells. Only the GFPuv protein expressed in *E. coli*, serving as the positive control, could be detected via Western blot. One possible explanation for this result could be a very low number of transformed cells. However, the transfection rate could not accurately be quantified in this study. Therefore, it is possible that the number of transformed *H. meleagridis* cells lies below the detection threshold of both the fluorescent microscope and Western Blot techniques, and that these methods are of low efficacy in this case.

Since the detection of the GFP protein with Western blot was unsuccessful, a further attempt was made to detect the plasmid DNA by using PCR method. The rationale was that if the GFP protein could not be detected directly, due to potential problems with the expression of this construct, detecting plasmid DNA would confirm successful transfection of parasitic

cells. PCR requires only a small amount of DNA, which is then amplified through several cycles of amplification. Detection of the resulting PCR product on an agarose gel is dependent on the number of cycles of amplification [83,84]. To ensure that only intracellular DNA was extracted, the samples were washed and treated with DNase1 enzyme prior to DNA extraction. The DNase1 enzyme is an endonuclease [85] that will degrade any DNA including plasmids in the medium, which have not been integrated into cells. The PCR results demonstrated the presence of GFP in all samples whether the samples were transfected or not, showing no difference between DNase1 treated or untreated samples. This indicated that a single DNase1 treatment was not sufficient to remove all extracellular plasmids from the culture medium and the obtained results were unreliable parameter for detecting transfected cells.

As mentioned above, despite all the modifications of the plasmid constructs, transfection and detection methods, no GFP signal could be detected in neither *H. meleagridis* nor *T. gallinae* cells. One possibility is that the promoter sequence(s) used in the construction of the plasmids do not support the expression of the GFP. Another possibility is the rapid degradation of the plasmid by the parasite cell machinery, which would also result in the absence of the GFP expression. Such situation would require that *H. meleagridis* or *T. gallinae* have a, to date, unknown, defence mechanism such as the CRISPR/cas9 system in bacteria and archaea [86]. A third possibility could be the failure in the electroporation process presumably due to an incorrect voltage setting, in which case a plasmid never entered the parasite cells. Electroporation depends on many factors such as: (1) composition of the cell membrane, (2) cell size, (3) conformation of the DNA molecule, (4) temperature, (5) duration and voltage of the pulse, (6) post-pulse manipulation and (7) medium [44,87].

Considering all these factors, it is possible that none of the tested programs was correct. Furthermore, the treatment of the cells after the electronic pulse might have been inadequate and would need to be further optimized.

In the future, the pBlue_Tricho_eGFPopt plasmid and the pBlue_Tricho_AcGFP plasmid should be tested in *T. vaginalis* to ensure that changing of the target organism has no influence on the detection of fluorescent signals in *T. gallinae* cells. In order to reproduce the condition described for *T. vaginalis* [48], the originally published *T. vaginalis* plasmid should be obtained for testing, and the Bio-Rad gene Pulser equipped with a capacitance

extender should be tested as an electroporator. Further, the transfection efficiency should be improved. Making the cells electro-competent is one method for increasing transfection efficiency. For this, the cells must be mixed with glycerol and kept at a low temperature (4°C) throughout the process [88]. Another approach would be to increase the efficiency by adding Mg2+ or K+ to the culture media [89] or the addition of DMSO (Dimethyl Sulfoxide) as reported for mammalian cells, [90]. However, it is known that *H. meleagridis* is sensitive to low temperatures and changes in the culture medium [76]. Therefore, it is unlikely that *H. meleagridis* cells can be made electro-competent, but the addition of ions may be beneficial. Further, there is no precise information on how rice starch affects electroporation, but it is likely that the less rice starch present in the medium, the better. In general, other transfection methods could also be tested, like a different lipofection reagent, magnetofection or nucleofection. The Nucleofector technology is a more advanced electroporation method, similar to the Neon transfections system. It was created specifically for primary cells and difficult-to-transfect cell lines. Nucleofection is a technique that combines electrical parameters and cell-type solutions to introduce plasmid DNA directly into the cell nucleus. As a result, cell division is no longer required for transfection, and high transfection efficiencies can be obtained in fresh or immortalized primary cells [91]. This system was already successfully applied to *T. vaginalis* [92] and *Eimeria tenella* [93]. The basic principle of magnetofection is straightforward: create a magnetic vector, add it to the medium that surrounds cultured cells, and then apply a magnetic field to direct the vector towards the target cells [94]. However, magnetofection was tested in *Leishmania,* with little success. *Leishmania* are considered to be difficult to transfect even though these cells are prone to protein uptake [95]. This difference in uptake of proteins and nucleotides should also be investigated for *H. meleagridis*. There are many different commercially available lipofection reagents, which could be tested. Suspension cells frequently cause issues during lipofectamine transfection. To overcome the non-adherent character of the cells, the culture plate can be coated. Coating materials commonly used are collagen and poly-L-lysine but there is a study about using egg white as an simple coating reagent [96]. Since *H. meleagridis* are suspension cells, the lipofection should be tested on additional coated culture plates. Besides that, if the expressed GFP protein cannot be detected directly by florescent microscopy or western blot, a suitable proof of transfection must be found. In addition to the possibility of detecting plasmids by PCR, plasmid DNA can be visualized

under the electron microscope [97–99]. This could theoretically be a method of testing transfection, but it would most likely be time-consuming and technically complex to screen a certain number of cells.

In conclusion, there are still many possibilities that can be tested to create a protocol for the genetic manipulation of *H. meleagridis*. Nevertheless, it will probably still take some time before the genome of *H. meleagridis* can be changed with CRISPR/cas9 for the first time.

## References

1. Tyzzer EE. The Flagellate Character and Reclassification of the Parasite Producing "Blackhead" in Turkeys: *Histomonas* (Gen. nov.) *meleagridis* (Smith). J Parasitol. 1920;6:124.

2. Hess M, Liebhart D, Bilic I, Ganas P. *Histomonas meleagridis*—New insights into an old pathogen. Vet Parasitol. 2015;208:67–76.

3. McDougald LR. Blackhead Disease (Histomoniasis) in Poultry: A Critical Review. Avian Dis. 2005;49:462–76.

4. Liebhart D, Ganas P, Sulejmanovic T, Hess M. Histomonosis in poultry: previous and current strategies for prevention and therapy. Avian Pathol. 2017;46:1–18.

5. Liebhart D, Hess M. Spotlight on Histomonosis (blackhead disease): a re-emerging disease in turkeys and chickens. Avian Pathol. 2020;49:1–4.

6. Cushman S. The production of turkeys. R I Agric Exp Stn Bull. 1893;25:89–123.

7. Palmieri N, de Jesus Ramires M, Hess M, Bilic I. Complete genomes of the eukaryotic poultry parasite *Histomonas meleagridis*: linking sequence analysis with virulence / attenuation. BMC Genomics. 2021;22:753.

8. Gerbod D, Edgcomb VP, Noel C, Zenner L, Wintjens R, Delgado-Viscogliosi P, et al. Phylogenetic Position of the Trichomonad Parasite of Turkeys, *Histomonas meleagridis* (Smith) Tyzzer, Inferred from Small Subunit rRNA Sequence1. J Eukaryot Microbiol. 2001;48:498–504.

9. van der Heijden HMJF, Landman WJM, Greve S, Peek R. Genotyping of *Histomonas meleagridis* isolates based on Internal Transcribed Spacer-1 sequences. Avian Pathol. 2006;35:330–4.

10. Hauck R, Hafez HM. Partial sequence of the beta-tubulin of *Histomonas meleagridis* and the activity of benzimidazoles against H. meleagridis in vitro. Parasitol Res. 2009;104:1183–9.

11. Cepicka I, Hampl V, Kulda J. Critical Taxonomic Revision of Parabasalids with Description of one New Genus and three New Species. Protist. 2010;161:400–33.

12. Hauck R, Hafez HM. Systematic Position of *Histomonas meleagridis* Based on Four Protein Genes. J Parasitol. 2010;96:396–400.

13. Lollis L, Gerhold R, McDougald L, Beckstead R. Molecular Characterization of *Histomonas meleagridis* and Other Parabasalids in the United States Using the 5.8S, ITS-1, and ITS-2 rRNA Regions. J Parasitol. 2011;97:610–5.

14. Bilic I, Jaskulska B, Souillard R, Liebhart D, Hess M. Multi-Locus Typing of *Histomonas meleagridis* Isolates Demonstrates the Existence of Two Different Genotypes. Leung FCC, editor. PLoS ONE. 2014;9:e92438.

15. Hatfaludi T, Rezaee MS, Liebhart D, Bilic I, Hess M. Experimental reproduction of histomonosis caused by *Histomonas meleagridis* genotype 2 in turkeys can be prevented by oral vaccination of day-old birds with a monoxenic genotype 1 vaccine candidate. Vaccine. 2022;40:4986–97.

16. Mazet M, Diogon M, Alderete J, Vivares C, Delbac F. First molecular characterisation of hydrogenosomes in the protozoan parasite *Histomonas meleagridis*. Int J Parasitol. 2008;38:177–90.

17. Bilic I, Leberl M, Hess M. Identification and molecular characterization of numerous *Histomonas meleagridis* proteins using a cDNA library. Parasitology. 2009;136:379–91.

18. Lynn EC, Beckstead RB. Identification of Gene Expression Elements in *Histomonas meleagridis* Using Splinkerette PCR, a Variation of Ligated Adaptor PCR. J Parasitol. 2012;98:135–41.

19. Klodnicki ME, McDougald LR, Beckstead RB. A genomic analysis of *Histomonas meleagridis* through sequencing of a cDNA library. J Parasitol. 2013;99:264–9.

20. Monoyios A, Patzl M, Schlosser S, Hess M, Bilic I. Unravelling the differences: comparative proteomic analysis of a clonal virulent and an attenuated *Histomonas meleagridis* strain. Int J Parasitol. 2018;48:145–57.

21. Monoyios A, Hummel K, Nöbauer K, Patzl M, Schlosser S, Hess M, et al. An Alliance of Gel-Based and Gel-Free Proteomic Techniques Displays Substantial Insight Into the Proteome of a Virulent and an Attenuated *Histomonas meleagridis* Strain. Front Cell Infect Microbiol. 2018;8:407.

22. Mazumdar R, Nöbauer K, Hummel K, Hess M, Bilic I. Molecular characterization of *Histomonas meleagridis* exoproteome with emphasis on protease secretion and parasite-bacteria interaction. Parkinson J, editor. PLOS ONE. 2019;14:e0212429.

23. Ramires M de J, Hummel K, Hatfaludi T, Riedl P, Hess M, Bilic I. Comparative Surfaceome Analysis of Clonal *Histomonas meleagridis* Strains with Different Pathogenicity Reveals Strain-Dependent Profiles. Microorganisms [Internet]. 2022;10. Available from: https://www.mdpi.com/2076-2607/10/10/1884

24. Lanigan TM, Kopera HC, Saunders TL. Principles of Genetic Engineering. Genes. 2020;11:291.

25. Nora LC, Westmann CA, Martins-Santana L, Alves L de F, Monteiro LMO, Guazzaroni M-E, et al. The art of vector engineering: towards the construction of next-generation genetic tools. Microb Biotechnol. 2019;12:125–47.

26. del Solar G, Giraldo R, Ruiz-Echevarría MJ, Espinosa M, Díaz-Orejas R. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev MMBR. 1998;62:434–64.

27. Stearns T. Green Flourescent Protein: The green revolution. Curr Biol. 1995;5:262–4.

28. Stepanenko O, Verkhusha V, Kuznetsova I, Uversky V, Turoverov K. Fluorescent Proteins as Biomarkers and Biosensors: Throwing Color Lights on Molecular and Cellular Processes. Curr Protein Pept Sci. 2008;9:338–69.

29. Soboleski MR, Oaks J, Halford WP. Green fluorescent protein is a quantitative reporter of gene expression in individual eukaryotic cells. FASEB J. 2005;19:1–20.

30. Gibson DG, Young L, Chuang R-Y, Venter JC, Hutchison CA, Smith HO. Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods. 2009;6:343–5.

31. Casini A, Storch M, Baldwin GS, Ellis T. Bricks and blueprints: methods and standards for DNA assembly. Nat Rev Mol Cell Biol. 2015;16:568–76.

32. Pyhtila B. Plasmids 101: Gibson Assembly and Other Long-Homology Based Cloning Methods [Internet]. [cited 2022 Jul 28]. Available from: https://blog.addgene.org/plasmids-101-gibson-assembly

33. Gibson Assembly [Internet]. [cited 2022 Jul 29]. Available from: https://www.snapgene.com/guides/gibson-assembly

34. Wang J-W, Wang A, Li K, Wang B, Jin S, Reiser M, et al. CRISPR/Cas9 nuclease cleavage combined with Gibson assembly for seamless cloning. BioTechniques. 2015;58:161–70.

35. Sorek R, Kunin V, Hugenholtz P. CRISPR — a widespread system that provides acquired resistance against phages in bacteria and archaea. Nat Rev Microbiol. 2008;6:181–6.

36. Rath D, Amlinger L, Rath A, Lundgren M. The CRISPR-Cas immune system: Biology, mechanisms and applications. Biochimie. 2015;117:119–28.

37. Doudna JA, Charpentier E. The new frontier of genome engineering with CRISPR-Cas9. Science. 2014;346:1258096.

38. Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F. Genome engineering using the CRISPR-Cas9 system. Nat Protoc. 2013;8:2281–308.

39. Featherstone C, Jackson SP. DNA double-strand break repair. Curr Biol. 1999;9:R759–61.

40. Kanaar R, Hoeijmakers JHJ, van Gent DC. Molecular mechanisms of DNA double-strand break repair. Trends Cell Biol. 1998;8:483–9.

41. Kaulich M, Dowdy SF. Combining CRISPR/Cas9 and rAAV Templates for Efficient Gene Editing. Nucleic Acid Ther. 2015;25:287–96.

42. Fajrial AK, He QQ, Wirusanti NI, Slansky JE, Ding X. A review of emerging physical transfection methods for CRISPR/Cas9-mediated gene editing. Theranostics. 2020;10:5532–49.

43. Kim TK, Eberwine JH. Mammalian cell transfection: the present and the future. Anal Bioanal Chem. 2010;397:3173–8.

44. Gehl J. Electroporation: theory and methods, perspectives for drug delivery, gene therapy and research: Electroporation: aspects and perspectives. Acta Physiol Scand. 2003;177:437–47.

45. Ganas P, Liebhart D, Glösmann M, Hess C, Hess M. Escherichia coli strongly supports the growth of *Histomonas meleagridis*, in a monoxenic culture, without influence on its pathogenicity. Int J Parasitol. 2012;42:893–901.

46. Amin A, Bilic I, Berger E, Hess M. *Trichomonas gallinae*, in comparison to *Tetratrichomonas gallinarum*, induces distinctive cytopathogenic effects in tissue cultures. Vet Parasitol. 2012;186:196–206.

47. Kawaguchi T, Nomura K, Hirayama Y, Kitagawa T. Establishment and characterization of a chicken hepatocellular carcinoma cell line, LMH. Cancer Res. 1987;47:4460–4.

48. Delgadillo MG, Liston DR, Niazi K, Johnson PJ. Transient and selectable transformation of the parasitic protist *Trichomonas vaginalis*. Proc Natl Acad Sci U S A. 1997;94:4716–20.

49. pmaxCloning Vector | Lonza [Internet]. [cited 2022 Aug 10]. Available from: https://bioscience.lonza.com/lonza_bs/AT/en/Transfection/p/000000000000191671/pmaxCloning%E2%84%A2-Vector

50. pGFPuvVector Map. :2. [Internet] [cited 2022 Aug 10] Available from: https://www.takarabio.com/assets/documents/Vector%20Documents/pGFPuv%20Vector%20Map_121913.pdf

51. AcGFP1 fluorescent protein [Internet]. [cited 2022 Jul 28]. Available from: https://www.takarabio.com/products/gene-function/fluorescent-proteins/fluorescent-protein-plasmids/cyan-and-green-fluorescent-proteins/acgfp1-fluorescent-protein

52. Mazumdar R, Endler L, Monoyios A, Hess M, Bilic I. Establishment of a de novo Reference Transcriptome of *Histomonas meleagridis* Reveals Basic Insights About Biological Functions and Potential Pathogenic Mechanisms of the Parasite. Protist. 2017;168:663–85.

53. EGFP Sequence and Map [Internet]. SnapGene. [cited 2022 Oct 11]. Available from: https://www.snapgene.com/resources/plasmid-files/?set=fluorescent_protein_genes_and_plasmids&plasmid=EGFP

54. Hess M. Commensal or pathogen – a challenge to fulfil Koch's Postulates. Br Poult Sci. 2017;58:1–12.

55. Lin-Chao S, Chen W-T, Wong T-T. High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol Microbiol. 1992;6:3385–93.

56. How do I know if my plasmid is a high- or low copy number type? - QIAGEN [Internet]. [cited 2022 Oct 11]. Available from: https://www.qiagen.com/us/resources/faq?id=1f42840e-fbd7-4734-b0cd-e17372a9e5a4&lang=en

57. Lambert T. AcGFP1 at FPbase [Internet]. FPbase. [cited 2022 Oct 7]. Available from: https://www.fpbase.org/protein/acgfp1/

58. Glossary of Terms [Internet]. [cited 2022 Oct 11]. Available from: https://help.fpbase.org/glossary#brightness

59. Lambert T. EGFP at FPbase [Internet]. FPbase. [cited 2022 Oct 7]. Available from: https://www.fpbase.org/protein/egfp/

60. Gagniuc P, Ionescu-Tirgoviste C. Eukaryotic genomes may exhibit up to 10 generic classes of gene promoters. BMC Genomics. 2012;13:512.

61. Pedersen AG, Baldi P, Chauvin Y, Brunak S. The biology of eukaryotic promoter prediction—a review. Comput Chem. 1999;23:191–207.

62. Sudheer Menon, Shanmughavel Piramanayakam, Gopal Agarwal. COMPUTATIONAL IDENTIFICATION OF PROMOTER REGIONS IN PROKARYOTES AND EUKARYOTES. EPRA Int J Agric Rural Econ Res. 2021;21–8.

63. Rita Costa A, Elisa Rodrigues M, Henriques M, Azeredo J, Oliveira R. Guidelines to cell engineering for monoclonal antibody production. Eur J Pharm Biopharm. 2010;74:127–38.

64. PT3717-5.pdf [Internet]. [cited 2022 Aug 2]. Available from: https://www.takarabio.com/documents/Vector%20Documents/PT3717-5.pdf

65. Petibon C, Malik Ghulam M, Catala M, Abou Elela S. Regulation of ribosomal protein genes: An ordered anarchy. WIREs RNA [Internet]. 2021 [cited 2022 Oct 20];12. Available from: https://onlinelibrary.wiley.com/doi/10.1002/wrna.1632

66. Lahti CJ, Bradley PJ, Johnson PJ. Molecular characterization of the α-subunit of *Trichomonas vaginalis* hydrogenosomal succinyl CoA synthetase. Mol Biochem Parasitol. 1994;66:309–18.

67. Carlton JM, Hirt RP, Silva JC, Delcher AL, Schatz M, Zhao Q, et al. Draft Genome Sequence of the Sexually Transmitted Pathogen *Trichomonas vaginalis*. Science. 2007;315:207–12.

68. Alrefaei AF, Low R, Hall N, Jardim R, Dávila A, Gerhold R, et al. Multilocus Analysis Resolves the European Finch Epidemic Strain of *Trichomonas gallinae* and Suggests Introgression from Divergent Trichomonads. Embley M, editor. Genome Biol Evol. 2019;11:2391–402.

69. Paschke P, Knecht DA, Williams TD, Thomason PA, Insall RH, Chubb JR, et al. Genetic Engineering of *Dictyostelium discoideum* Cells Based on Selection and Growth on Bacteria. J Vis Exp. 2019;58981.

70. Meissner M, Agop-Nersesian C, Sullivan WJ. Molecular tools for analysis of gene function in parasitic microorganisms. Appl Microbiol Biotechnol. 2007;75:963–75.

71. Eid J, Sollner-Webb B. Efficient introduction of plasmid DNA into *Trypanosoma brucei* and transcription of a transfected chimeric gene. Proc Natl Acad Sci. 1987;84:7812–6.

72. Yee J, Nash TE. Transient transfection and expression of firefly luciferase in *Giardia lamblia*. Proc Natl Acad Sci. 1995;92:5615–9.

73. Wu Y, Sifri CD, Lei HH, Su XZ, Wellems TE. Transfection of *Plasmodium falciparum* within human red blood cells. Proc Natl Acad Sci. 1995;92:973–7.

74. Purdy JE, Mann BJ, Pho LT, Petri WA. Transient transfection of the enteric parasite *Entamoeba histolytica* and expression of firefly luciferase. Proc Natl Acad Sci. 1994;91:7099–103.

75. Potter H. Transfection by electroporation. Curr Protoc Mol Biol. 2003;Chapter 9:Unit 9.3.

76. Zaragatzki E, Hess M, Grabensteiner E, Abdel-Ghaffar F, Al-Rasheid KAS, Mehlhorn H. Light and transmission electron microscopic studies on the encystation of *Histomonas meleagridis*. Parasitol Res. 2010;106:977–83.

77. Neon Electroporation Transfection - AT [Internet]. [cited 2022 Aug 5]. Available from: https://www.thermofisher.com/uk/en/home/life-science/cell-culture/transfection/neon-transfection-system.html

78. Neon Transfection System Cell Line Data - DE [Internet]. [cited 2022 Oct 11]. Available from: https://www.thermofisher.com/de/de/home/life-science/cell-culture/transfection/neon-transfection-system/neon-transfection-system-cell-line-data.html

79. Brees C, Fransen M. A cost-effective approach to microporate mammalian cells with the Neon Transfection System. Anal Biochem. 2014;466:49–50.

80. Gheibi A, Khanahmad H, Kardar G, Boshtam M, Rezaie S, Kazemi B, et al. Optimization and Comparison of Different Methods and Factors for Efficient Transformation of *Brucella abortus* RB51strain. Adv Biomed Res. 2019;8:37.

81. Ghosh R, Gilda JE, Gomes AV. The necessity of and strategies for improving confidence in the accuracy of western blots. Expert Rev Proteomics. 2014;11:549–60.

82. Xiang Y, Zheng Y, Liu S, Liu G, Li Z, Dong W. Comparison of the sensitivity of Western blotting between PVDF and NC membranes. Sci Rep. 2021;11:12022.

83. PCR Troubleshooting [Internet]. Bio-Rad Lab. [cited 2022 Oct 11]. Available from: https://www.bio-rad.com/de-at/applications-technologies/pcr-troubleshooting?ID=LUSO3HC4S

84. Lorenz TC. Polymerase chain reaction: basic protocol plus troubleshooting and optimization strategies. J Vis Exp JoVE. 2012;e3998.

85. Hartmann G. Nucleic Acid Immunity. Adv Immunol [Internet]. Elsevier; 2017 [cited 2022 Dec 19]. p. 121–69. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0065277616300529

86. Gong S, Yu HH, Johnson KA, Taylor DW. DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity. Cell Rep. 2018;22:359–71.

87. Sokołowska E, Błachnio-Zabielska AU. A Critical Review of Electroporation as A Plasmid Delivery System in Mouse Skeletal Muscle. Int J Mol Sci. 2019;20:2776.

88. Gonzales MF, Brooks T, Pukatzki SU, Provenzano D. Rapid protocol for preparation of electrocompetent *Escherichia coli* and *Vibrio cholerae*. J Vis Exp JoVE. 2013;

89. Sherba JJ, Hogquist S, Lin H, Shan JW, Shreiber DI, Zahn JD. The effects of electroporation buffer composition on cell viability and electro-transfection efficiency. Sci Rep. 2020;10:3053.

90. Melkonyan H. Electroporation efficiency in mammalian cells is increased by dimethyl sulfoxide (DMSO). Nucleic Acids Res. 1996;24:4356–7.

91. Gresch O. New non-viral method for gene transfer into primary cells. Methods. 2004;33:151–63.

92. Janssen BD, Chen Y-P, Molgora BM, Wang SE, Simoes-Barbosa A, Johnson PJ. CRISPR/Cas9-mediated gene modification and gene knock out in the human-infective parasite *Trichomonas vaginalis*. Sci Rep. 2018;8:270.

93. Marugan-Hernandez V, Cockle C, Macdonald S, Pegg E, Crouch C, Blake DP, et al. Viral proteins expressed in the protozoan parasite *Eimeria tenella* are detected by the chicken immune system. Parasit Vectors. 2016;9:463.

94. Schillinger U, Brill T, Rudolph C, Huth S, Gersting S, Krötz F, et al. Advances in magnetofection—magnetically guided nucleic acid delivery. J Magn Magn Mater. 2005;293:501–8.

95. Keller A, Scheiding B, Breitling R, Licht A, Hemmerich P, Lorkowski S, et al. Transduction and transfection of difficult-to-transfect cells: Systematic attempts for the transfection of protozoa *Leishmania*. J Cell Biochem. 2019;120:14–27.

96. Basiouni S, Fuhrmann H, Schumann J. High-efficiency transfection of suspension cell lines. BioTechniques. 2012;53:1–4.

97. Burkardt HJ, Pühler A. 7 Electron Microscopy of Plasmid DNA. Methods Microbiol [Internet]. Elsevier; 1988 [cited 2022 Oct 27]. p. 155–77. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0580951708700746

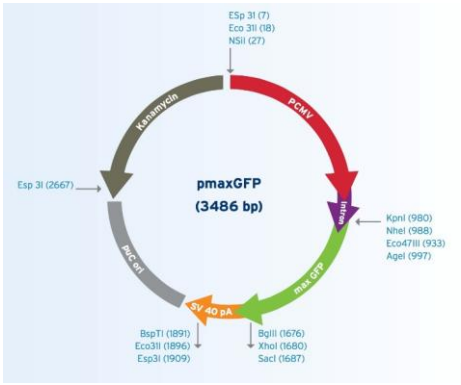
98. Hajibagheri MAN. Visualization of DNA and RNA Molecules, and Protein-DNA Complexes for Electron Microscopy. Mol Biotechnol. 2000;15:167–84.

99. Kunisada T, Yamagishi H. Rapid microscale procedure for visualizing intracellular plasmid DNA by electron microscopy. Plasmid. 1983;9:8–16.

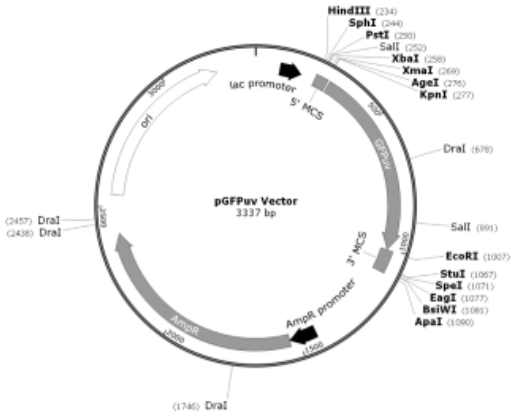
100. Does anyone know the sequence for the multiple cloning site in Lonza's pmaxGFP or at least the order of restriction sites upstream of GFP? [Internet]. [cited 2022 Aug 2]. Available from: https://www.researchgate.net/post/Does-anyone-know-the-sequence-for-the-multiple-cloning-site-in-Lonzas-pmaxGFP-or-at-least-the-order-of-restriction-sites-upstream-of-GFP

# Annex

## Annex 1. Plasmid maps of the three commercial plasmids which were used as control

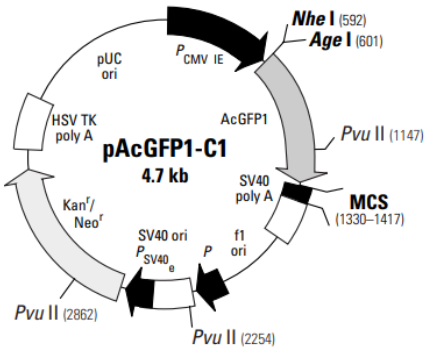



[100] **a.** (Lonza Inc, USA)




**b.** (Clontech, France)

[50]




**c.** (TAKARA BIo Inc, USA Cat. No. 632470)

[64]

**Cartoon of the commercial plasmids showing the plasmid structure in detail.** a.) Plasmid map of pMAXGFP (Lonza LTD, Switzerland The plasmid has an early cytomegalovirus (PCMV IE) promoter and a chimeric intron in front of the maxGFP gene [49,100]. b.) Plasmid map of pGFPuv (Clonetch, France) with a lac promoter and a pGFPuc gene [45,50]. c.) Plamid map of p*AcGFP*1-C1 (TAKARA Bio Inc., USA). The plasmid has a cytomegalovirus (CMV) promoter and an *AcGFP* gene [51,64].

## Annex 2. Complete *g6464_AcGFP*opt insert sequence

GTATTGCGCTCAAAATCATAATAACTCATTTCATATGAGAACTGAAATTATATAAAATTGAAATTTACC
ATATAAAAGGAAAATTTTGTAATGAAATTCATTCATCATTTCATATAATTTTATAATTTTAATCGTAACA
AATGAATTTAAAAAAACTCATATTTTAAGATTTATTATATATTCAATACTTCATATTGTTTCTCAGCTTT
ATATAGAAATGAAATATTATTTAATACAATATCCAGTGAGATATTTTAGTTTTAAAATTTAAAAAGTAA
ACTTCTTAAATTTGGTGACGGCGGTTATAAAGCGGTTTTGGTCTTCCTCTTCTCCAAATGCCTCGTTGC
CAACGTAAGCACTTGAAGCGCCTTGCTGCCCCACACCATTGGATGCTCGCCAAAACAGCTGGCAAATT
TGCTATGGTTTCAAAAGGTGCAGAATTATTCACAGGTATTGTTCCAATTTTAATTGAATTAAATGGTG
ATGTTAATGGTCATAAATTCTCAGTTTCAGGTGAAGGTGAAGGTGATGCAACATATGGTAAATTAACA
TTAAAATTCATTTGTACAACAGGTAAATTACCAGTTCCATGGCCAACATTAGTTACAACATTATCATAT
GGTGTTCAATGTTTCTCACGTTATCCAGATCATATGAAACAACATGATTTCTTCAAATCAGCAATGCCA
GAAGGTTATATTCAAGAACGTACAATTTTCTTCGAAGATGATGGTAATTATAAATCACGTGCAGAAGT
TAAATTCGAAGGTGATACATTAGTTAATCGTATTGAATTAACAGGTACAGATTTCAAAGAAGATGGTA
ATATTTTAGGTAATAAAATGGAATATAATTATAATGCACATAATGTTTATATTATGACAGATAAAGCA
AAAAATGGTATTAAAGTTAATTTCAAAATTCGTCATAATATTGAAGATGGTTCAGTTCAATTAGCAGA
TCATTATCAACAAAATACACCAATTGGTGATGGTCCAGTTTTATTACCAGATAATCATTATTTATCAAC
ACAATCAGCATTATCAAAAGATCCAAATGAAAAACGTGATCATATGATTTATTTCGGTTTCGTTACAG
CAGCAGCAATTACACATGGTATGGATGAATTATATAAATAAAAGGAATCTTTTATTTAATAATTTGGTT
AATTACAGTGTGTTAACATTTAAAAGTTGTTTGCAGTGAAGAGAATCTGTATGCGTAAGAAGACATTT
CTGTTGATTTGGTTAAACTTTTTAACTCCTGTTCAATGTATAAAATATTTTACTATCTTCATTAAGAAGC
TTACGATTTGTAATCTATATCAGGTGCATATAGGTGTTGTACGAATAATGCGTTATTCATAGAGATCAT
ATTTAATAATTTTTTCAATTTAAATTTGTTATTTTCCTATATTTTTCACTCACTTCACTTAGGAGAGCGTT
AACAATGACAAAACTTTTTTCTTTTCATTTTTCATAAAAATGATATCATTTCTGAGACTGAACTATTTTA
AAATTTCTATTTGCCAACAAACTAATTAAAAATTTATGATTTATTTAAGTTTTTGTTGTCTTCTACAATTT
TTTCATCTAAACGAACTTCAAATAAAATTTATCTTAAGAATTCTTTACTTATTAAAACTGTTCTAATATA
TTC

**Annex 3. Complete *g7554_eGFP*opt insert sequence**

AATTGGTCCATTTTTCTCACTTAATAATAACTCAAATGTATTTAAGACAATGTTTT
TATGTTCTCCATCATCAATTAAATTAAATAATATTCGAAGTGATGCTAATCGAATT
TTAATGTCATAAGTCTTAGCAAACTTTAAATATGGAAATAATACATTTTGTTTGAT
ATTACAACATTTTGTTAATGAATTTAATGCATCTAAATAATAATAACTTTCAATGT
TTTTACATGATTCATATAAAAACGTATATATCGGTAATAATGAAGTTTGCTCATC
AAATGTTAACTTCTTACGACATATAAATGGAATAAGTTGTATAACACGTTCAAGA
TTTTCATTTTCATTATATAAATAATCAGAAATTCGAACAATAATCGGAAGAAAAT
CAACGTCATTTAACGGTTTGGAATTAGCTAAAATAAGATTAATAAATTCAAAGGA
ACTTTCAGTAAAAATTAAGTTCATGTATAACTTGTAGTTCTTTTTAAGTAAAATTA
TTGTTGGATTAACATATTTGATATTTTTGCACATTTTTAAAATTAACTTATGGACT
TGTTCTTCATTGCTGATTTTGCAGAGATAATGCCATGTTTCATTGGTGATATGAA
CGTAATCAGAAAGACTAAGAGCATTGAATACACATATGAAAAACATATTCCACTT
ACTTTGTTCTGCTTCTGATAGAGTGGGAATGGAATATGAAAGAATGGGAAGGAT
AGAGGAAAGAATACCAATTGGAAATGATTTTATTGTATTATAAAAATCTTCCTTA
TGAGTAAGCTCATAAAGACTTAGCCAATGCAAAACAAATCTTTCTGCTCTACTAC
GTAGTTTTACATTTGTGTTTTTTGTAAGGGGAAAGATGTGTTCGAAGATATGAAC
AACAGTTTTTGGATCATTAAAATATGTAAGTTTCATTCGAGAAGTAATATGATAT
ACATTTTTTGTAGCTCTTTCTATGTAGTTTTTATCTGATACATCTTCTAATTTTTTA
AGGTATTCTTCGATTTCATCCATTTGATAATATTTTTAAATACTTTACATTTATTGA
ATTTTATTCTGTCTCCAAACGGTTTATTGCAAAAAAAAACCGTTAATCGCACAGA
TGCACTATTTTTGAAATTTCTTGCTTGTTTTGAAGTAAATAAATTATATCAATTCA
TATTTTAAAACATTTTTTGTAAGATCGCGAAAATAAAAAGTGTATAAGATTAATT
CATCATTTAAATAGAAATTGATCATTAATCTTTTTTTGCTAATAATTATGCAAATT
TCATGAATATCACAAAAAACAATACTAATTAGCCACAAAACCGCATTAATGGTAA
TGCTTTTTAAATGCAGAATATTTTTAAATTAATTATTTTTATCAAATATTTCCGGG
CATTTCTTTAACAACAAAAATCCTCCAAAAAGAAGGTCTCATTTTTAAACATAAT
GAAGAAACGTTGCGATTCTATGATTATCTTATAATAAAACTTCATTTTATTTCAAT
AAAAATTTACAGTTAAAATTTTGGTATTAAATGGTTTATTACAAGAAGTGATCTG
AACAATTACGTAAATTATTCAAATATCGGGAAAAATTTGTCACTTGTAATGGAAG
AAAAAAATATCTTTGGAGAAAAGCCAAATAAGGAGGAAATAAAAGTCGAAGAT
GAAGCTGAAATAAAGACTGAAAAAATGGTTTCAAAAGGAGAAGAATTATTTACA
GGAGTTGTTCCAATTTTAGTTGAATTAGATGGAGATGTTAATGGACATAAATTTT
CAGTTTCAGGAGAAGGAGAAGGAGATGCAACATATGGAAAATTAACATTAAAAT
TTATTTGTACAACAGGAAAATTACCAGTTCCATGGCCAACATTAGTTACAACATT
AACATATGGAGTTCAATGTTTTTCAAGATATCCAGATCATATGAAACAACATGAT
TTTTTTAAATCAGCAATGCCAGAAGGATATGTTCAAGAAAGAACAATTTTTTTTA
AAGATGATGGAAATTATAAAACAAGAGCAGAAGTTAAATTTGAAGGAGATACAT

TAGTTAATAGAATTGAATTAAAAGGAATTGATTTTAAAGAAGATGGAAATATTTT
AGGACATAAATTAGAATATAATTATAATTCACATAATGTTTATATTATGGCAGAT
AAACAAAAAAAATGGAATTAAAGTTAATTTTAAAATTAGACATAATATTGAAGATG
GATCAGTTCAATTAGCAGATCATTATCAACAAAATACACCAATTGGAGATGGACC
AGTTTTATTACCAGATAATCATTATTTATCAACACAATCAGCATTATCAAAAGATC
CAAATGAAAAAAGAGATCATATGGTTTTATTAGAATTTGTTACAGCAGCAGGAA
TTACATTAGGAATGGATGAATTATATAAATAAAAGTGAATTTTGGGGTGAAATTC
ACCAACATCAAGTAAAACAACCCCGAAAACCGCCTTTTCGAATTTAAAAAGTGAA
AATTTCTGAGATTTACACATTTTATACTTTAATATTTGCATTTTTATAGTTATTGCT
TAATGCTCTTAGGTTATCTAATTATGAATGAATTAATTAAACCATGAAGAAGATT
AAAAATAGGAAAAATATTGTTTACCAATTTTGTTTATCTTATTGTTAAAATCATAT
CGAACGCACACATTAAATTATGGGCTTGCCTTAGAAGAGTTCCTCTTGCATTTGG
TAGAGGAGTTCTTCTGTCTTTAAAATTGATAGGTTAATGTCGAATTTATTGGCAA
TAAGAATTTGGGCAGCTTCTTCTGTGGTAACAGGATATATGTCAACAATTGTATG
AACCATTTCTTGAAATTGAGTTGATGCAGGGAAATTTTG

## Annex 4. Complete *AcGFP*1+MS insert sequence

ATGGTGAGCAAGGGCGCCGAGCTGTTCACCGGCATCGTGCCCATCCTGATCGAG
CTGAATGGCGATGTGAATGGCCACAAGTTCAGCGTGAGCGGCGAGGGCGAGGG
CGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT
GCCTGTGCCCTGGCCCACCCTGGTGACCACCCTGAGCTACGGCGTGCAGTGCTT
CTCACGCTACCCCGATCACATGAAGCAGCACGACTTCTTCAAGAGCGCCATGCCT
GAGGGCTACATCCAGGAGCGCACCATCTTCTTCGAGGATGACGGCAACTACAAG
TCGCGCGCCGAGGTGAAGTTCGAGGGCGATACCCTGGTGAATCGCATCGAGCTG
ACCGGCACCGATTTCAAGGAGGATGGCAACATCCTGGGCAATAAGATGGAGTAC
AACTACAACGCCCACAATGTGTACATCATGACCGACAAGGCCAAGAATGGCATC
AAGGTGAACTTCAAGATCCGCCACAACATCGAGGATGGCAGCGTGCAGCTGGCC
GACCACTACCAGCAGAATACCCCCATCGGCGATGGCCCTGTGCTGCTGCCCGAT
AACCACTACCTGTCCACCCAGAGCGCCCTGTCCAAGGACCCCAACGAGAAGCGC
GATCACATGATCTACTTCGGCTTCGTGACCGCCGCCGCCATCACCCACGGCATGG
ATGAGCTGTACAAGTCCGGACTCAGATCTCGAGCTCAAGCTTCGAATTCTGCAGT
CGACGGTACCGCGGGCCCGGGATCCACCGGATCTAGATAA

## Annex 5. Complete Tricho_*eGFP*opt insert sequence

AAATGAATTCGATGAATCAGTTACAAGTGATAATTATGAAAAAATTGCTAGCGAATTAGGAGAGATTT
CATCAAAAATCATCAGCGTTAAGAAATCTCACGAAGATTACGAAGATAGTATTAAACAAATGCTGACT
AAAGCTAACAATTTGCTGCAGCAAAGCAAATCAGAGATGAGCAAAAAAGAATGTCAAAAAGTAATCG
CAGAATTGACTGCATTGATCAATACTGATAAGTCACAGCCAATTTCATTCGATGAACACAAATGGAAT
AGAAGAATGAGATCTTTAGATAATGTTGTTAAAATGAGTAACGCAGGTGTATTCTAAATATTGTTGTT
TCGATTTTAAATTTTTCCTTTCTTTCGTACATTTACTATTATATAGTTATAATTAATATTATAGTAAATCTA
TCTTGTAATCATTTGTACTCTAGATAAAGATGGTATAATAAGTATGTTTCTGGTTTAAAATACCATCTTT
ATGAATTACAAAAGCAAGTAATACTAAAATACATAATGAAGCTTTTTTTATCCGAATTGAATTGCGAAA
GAAATTTATAGAATAGTGCTATATTAAATATTATTGGACCATTATTTGTATGCAATATGAATACATGAA
TAAATACATGAATTGAAAATTATTGGACGATCATTAAAAATTTTACCCATCGAGCTTGAGTGAGAGAA
CTTTTCCATTCAATTTCCAAAACTAAAATGAAATATTAGGTAGTGAAATAAAAATATATGATATTTAGC
ATTTGGATATTTTTAACCGAATAACCGATGAACGAAGAAAATTTGAAATAAAATTGACGATCATTGAT
TTTTTTGTTAATTAAATATTATCTAATATAACCATCAAATTATAATATAATAAATTGTGATGGATAAATT
AAAATTTGAATACTACGACATAAAATAAATAACAAGTAATAATACCACTTTAAATAACAAATTTTCACA
AAGCTGCTAAAGAATTAGAATAGAGATAATAATAACGGTGAAGTTTATTTTGGGGCCAATTACTATAG
GACTATTTTAAACGCTATACATGCTTTTATATAAACGTACTTATGCAATAGAATTTCGTGGAAATGAAA
TTATTTAATTGCCTAAGATCGGAAATTTGCGAGCAAATACTTTATTAAATGCTGTCTTCTAAATATAGA
ACAATTTTTATCAATTAGCTTCGCAAAATATAAAAAATGATCATTAATTTTGGTAATAAGCTATAGTAA
TGATATATAATAGTTTAGCATGATAAAAAATTAATTCGAATACTTGAACCGAATTTGTATTCAAAATAA
AAAATTTTATCTCAAAAAATGGAAAAAATTTGAGGAACCGAAAATTCTTATCTATGCAAAAGAATTTTA
AATATTAAATTAGGTCATCATGTTTTTTCTCAAAAATTCTTTCATAGAAAATAAAATGTCAATGATATGA
CAATATATCATTAAAACACTAATAATTTGATAAAGAATAATTTCCAACTTAACAAAAAATGAAAGGAG
AAAAATATGTGCTGAAAAAAACAGGAGAAACGATTGCCCCTTAAAAATTGAGAAGATATAGCGGTTA
TAAAATTGACATCGGCCACTTACGCTTCAATTAAGGGGTAAAATTGTAAGCCTAAAATGGCTTTGTTCA
CTTCACATTAATGCTCTCTTCATCATTCGAGCGCAATCTTCATCAACCACTCCTCTTCATTGACAAGGAC
ACAAGAGTTGTTATCCAAGGTAGCATGGTTTCAAAAGGAGAAGAATTATTTACAGGAGTTGTTCCAAT
TTTAGTTGAATTAGATGGAGATGTTAATGGACATAAATTTTCAGTTTCAGGAGAAGGAGAAGGAGAT
GCAACATATGGAAAATTAACATTAAAATTTATTTGTACAACAGGAAAATTACCAGTTCCATGGCCAAC
ATTAGTTACAACATTAACATATGGAGTTCAATGTTTTTCAAGATATCCAGATCATATGAAACAACATGA
TTTTTTTAAATCAGCAATGCCAGAAGGATATGTTCAAGAAAGAACAATTTTTTTTAAAGATGATGGAA
ATTATAAAACAAGAGCAGAAGTTAAATTTGAAGGAGATACATTAGTTAATAGAATTGAATTAAAAGG
AATTGATTTTAAAGAAGATGGAAATATTTTAGGACATAAATTAGAATATAATTATAATTCACATAATGT
TTATATTATGGCAGATAAACAAAAAAATGGAATTAAAGTTAATTTTAAAATTAGACATAATATTGAAG
ATGGATCAGTTCAATTAGCAGATCATTATCAACAAAATACACCAATTGGAGATGGACCAGTTTTATTAC
CAGATAATCATTATTTATCAACACAATCAGCATTATCAAAAGATCCAAATGAAAAAAGAGATCATATG
GTTTTATTAGAATTTGTTACAGCAGCAGGAATTACATTAGGAATGGATGAATTATATAAATAAATCTAC
TGCTTACTTTAAAAACTCATACTTCTAAATAGGAAAGCTTCTCTCTAAGAAAAATGCATTCTTTTTTTTA

AATACCGAATTTTATATTATAATAATTAAACAGGAACAATAGGACATTTAAATATAATAGGAACAGGA
GAAACAGATTGGATTCTCCTCTCAGAGTGTCTGCGCGTCCTCGCACGCGGCTTACCTAACATCAAATTT
TCCATCTTGAGTCTTTCAGCATCATTTATGGTAATCCTCACGAGTTCGGTGATGGATACATAATGGAAT
AAACCATTATCCTTACCATCAACAGCGTTTTTATGATCGCCATCCGTTTTAATAGAATTCTGGCAATCAT
AAATAATGTCGCTCATATCGTCAATCTTAACGATTGATCCAATAGCTATATTATTCTGTGCGCTATCTCC
AACTTCTGGGTTCGGATCGAAATATGA

**Annex 6. Primer List.** All primers used for the plasmid's construction. Sequence segment in capital letters represents the regions that bound to the template DNA; sequence in lowercase letters corresponds to the overhang to the neighbouring sequences. Corresponding annealing temperature for each primer pair are listed in the last column.

| Name | Sequence | Temperature [°C] |
|---|---|---|
| pUC19_*g6464_AcGFP*opt | | |
| pUC19_fwd | aaaactgttctaatatattcCCGAGCTCGAATTCACTG | 64 |
| pUC19_rev | tatgattttgagcgcaatacTACCCGGGGATCCTCTAG | |
| Promoter_*g6464*_puc_fwd | ctctagaggatccccgggtaGTATTGCGCTCAAAATCATAATAACTC | 62,8 |
| Promoter_*g6464*_puc_rev | tctgcacctttgaaaccatAGCAAATTTGCCAGCTGTTTTG | |
| pex_GFP_puc_fwd | aaacagctggcaaatttgctATGGTTTCAAAAGGTGCAG | 56,6 |
| pex_GFP_puc_rev | tattaaataaaagattccttTTATTTATATAATTCATCCATACCATG | |
| Terminator_*g6464*_puc_fwd | tggatgaattatataaataaAAGGAATCTTTTATTTAATAATTTGGTTAATTAC | 56,6 |
| Terminator_*g6464*_puc_rev | gccagtgaattcgagctcggGAATATATTAGAACAGTTTTAATAAGTAAAG | |
| pBlue_*g6464_AcGFP*opt | | |
| pBluescript_II_SK (+) _fwd | aaaactgttctaatatattcGGGGATCCACTAGTTCTAGAG | 64 |
| pBluescript_II_SK (+) _rev | tatgattttgagcgcaatacCGATACCGTCGACCTCGAG | |
| Promoter_*g6464*_pB_fwd | cctcgaggtcgacggtatcgGTATTGCGCTCAAAATCATAATAACTC | 56,4 |
| Promoter_*g6464*_pB_rev | tctgcacctttgaaaccatAGCAAATTTGCCAGCTGTTTTG | |
| pex_GFP_pB_fwd | aaacagctggcaaatttgctATGGTTTCAAAAGGTGCAG | 57,2 |
| pex_GFP_pB_rev | tattaaataaaagattccttTTATTTATATAATTCATCCATACCATG | |
| Terminator_*g6464*_pB_fwd | tggatgaattatataaataaAAGGAATCTTTTATTTAATAATTTGGTTAATTAC | 60,5 |
| Terminator_*g6464*_pB_rev | tctagaactagtggatccccGAATATATTAGAACAGTTTTAATAAGTAAAG | |
| pBlue_*g7554_eGFP*opt | | |
| pBluescript_fwd | gttgatgcagggaaattttgGGGGATCCACTAGTTCTAGAG | 64,2 |
| pBluescript_rev | gtgagaaaaatggaccaattCGATACCGTCGACCTCGAG | |
| Promoter_*g7554*_fwd | cctcgaggtcgacggtatcgAATTGGTCCATTTTTCTCAC | 56,2 |
| Promoter_*g7554*_rev | tcttctccttttgaaaccatTTTTTCAGTCTTTATTTCAGC | |
| *eGFP*_fwd | ctgaaataaagactgaaaaaATGGTTTCAAAAGGAGAAG | 58,1 |
| *eGFP*_rev | tttcaccccaaaattcacttTTATTTATATAATTCATCCATTCCTAATG | |
| Terminator_*g7554*_fwd | tggatgaattatataaataaAAGTGAATTTTGGGGTGAAATTC | 59 |
| Terminator_*g7554*_rev | tctagaactagtggatccccCAAAATTTCCCTGCATCAAC | |
| pBlue_*g7554_AcGFP* | | |
| pBlue_*g7554_eGFP*opt_fwd | gatccaccggatctagataaAAGTGAATTTTGGGGTGAAATTC | 59,6 |
| pBlue_*g7554_eGFP*opt_rev | tcggcgcccttgctcaccatTTTTTCAGTCTTTATTTCAGCTTC | |
| *AcGFP*1-C1+MS_fwd | ctgaaataaagactgaaaaaATGGTGAGCAAGGGCGCC | 68,5 |

| | | |
|---|---|---|
| *AcGFP*1-C1+MS_rev | tttcaccccaaaattcacttTTATCTAGATCCGGTGGATCCCG | |
| pBlue_Tricho_*AcGFP* | | |
| pBlue_Tricho_*eGFP*_fwd | gatccaccggatctagataaATCTACTGCTTACTTTAAAAACTC | 57,8 |
| pBlue_Tricho_*eGFP*_rev | tcggcgcccttgctcaccatGCTACCTTGGATAACAAC | |
| *AcGFP*1-C1+MS_fwd | gagttgttatccaaggtagcATGGTGAGCAAGGGCGCC | 68,5 |
| *AcGFP*1-C1+MS_rev | ttttaaagtaagcagtagatTTATCTAGATCCGGTGGATCCCG | |

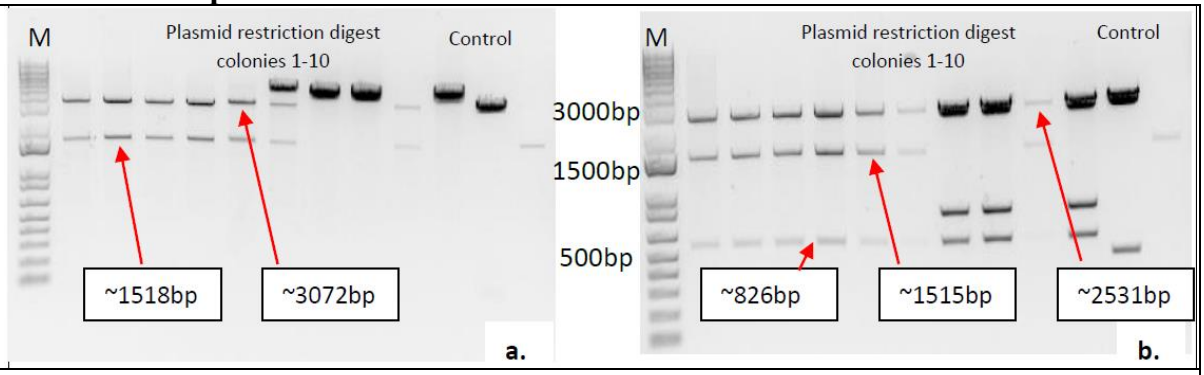
**Annex 7. Restriction enzyme list.** Enzymes were used for the diagnostic restriction digest to confirm the plasmid structure. All enzymes were purchased from NEB (New England Biolabs Inc., USA)

| **Plasmid name** | EcoRV/PstI | PvuII | KpnI/EcoRV | PstI |
|---|---|---|---|---|
| pUC19_*g6464_AcGFP*opt | **Yes** | No | No | No |
| pBlue_*g6464_AcGFP*opt | No | **Yes** | **Yes** | No |
| pBlue_*g7554_eGFP*opt | No | No | **Yes** | No |
| pBlue_Tricho_*AcGFP* | No | No | No | **Yes** |
| pBlue_*g7554_AcGFP* | No | No | No | **Yes** |

**Annex 8: List of the sequencing primer.** Primer sequence and binding position in the plasmid sequence. All primers used for the plasmid's verification. (The sequencing was carried out by LGC Genomics GmbH, Germany)

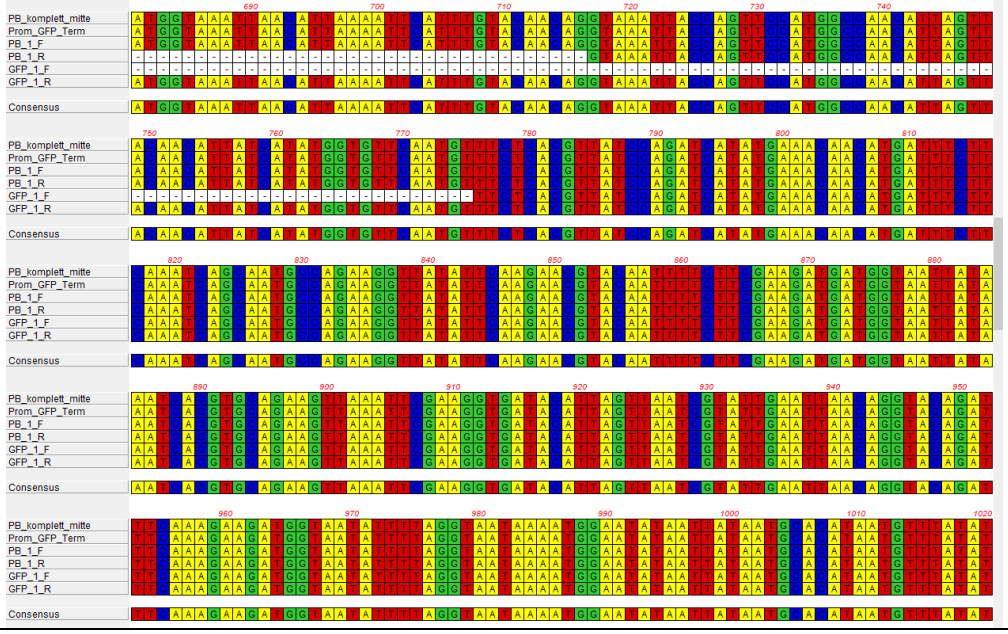
| | Primer sequence | Binding position |
|---|---|---|
| **pUC19_*g6464_AcGFP*opt** | | |
| M13-21F | TGTAAAACGACGGCCAGT | Outside of the insert |
| M13-29R | CAGGAAACAGCTATGACC | Outside of the insert |
| GFP_RW | AACTGGACCATCACCAATTGG | Within the GFP |
| GFP_FW | CCAGTTCCATGGCCAACATT | Within the GFP |
| **pBlue_*g6464_AcGFP*opt** | | |
| PB_stand_rw | GGAAACAGCTATGACCATG | Outside of the insert |
| PB_stand_fw | GTAAAACGACGGCCAG | Outside of the insert |
| GFP_RW | AACTGGACCATCACCAATTGG | Within the GFP |
| GFP_FW | CCAGTTCCATGGCCAACATT | Within the GFP |
| **pBlue_*g7554_eGFP*opt** | | |
| M13-21F | TGTAAAACGACGGCCAGT | Outside of the insert |
| M13-29R | CAGGAAACAGCTATGACC | Outside of the insert |
| *eGFP*_R | TCCTGCTGCTGTAACAAATTC | Within the GFP |
| *eGFP*_F | GAAGGAGAAGGAGATGCAAC | Within the GFP |
| Prom_F | CTGCTTCTGATAGAGTGGGA | Within the Promoter |
| Prom_R | CAAAAATAGTGCATCTGTGCG | Within the Promoter |
| **pBlue_Tricho_*AcGFP*** | | |
| *AcGFP*_Fw | GCACGACTTCTTCAAGAGC | Within the GFP |
| *AcGFP*_Rw | CTTGATGCCATTCTTGGCCTTG | Within the GFP |
| Tricho_Fw | TGCCTAAGATCGGAAATTTGC | Outside of the insert |
| Tricho_Rw | TCGTCATATTTCGATCCGAAC | Outside of the insert |
| **pBlue_*g7554_AcGFP*** | | |
| *AcGFP*_Fw | GCACGACTTCTTCAAGAGC | Within the GFP |
| *AcGFP*_Rw | CTTGATGCCATTCTTGGCCTTG | Within the GFP |
| *g7554*_Fw | TCGCACAGATGCACTATTTTTG | Outside of the insert |
| *g7554*_Rw | GCCCATAATTTAATGTGTGCG | Outside of the insert |

**Annex 9. Result of the diagnostic restriction digest of the pBlue_*g6464_AcGFP*opt plasmid as another example.**

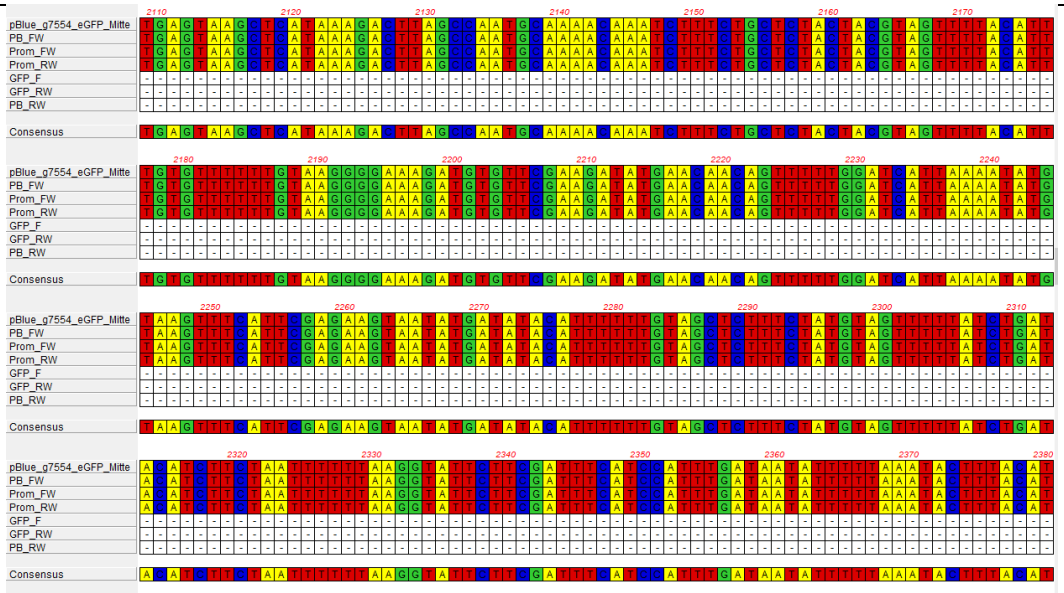



**(a)** Result of pBlue_*g6464_AcGFP*opt restriction digest with EcoRV-HF and KpnI-HF Sample 1-5 & 9 upper band=3072bp, lower band= 1518bp; **(b)** Result of pBluescriptIISK(+)_HisPromS4_*AcGFP*_HisTermS4 digest with PvuII-HF Sample 1-5 & 9 upper band=2513bp, middle band=1515, lower band= 862bp; control= pBluescriptIISK(+) digested and undigested; M= 1kb Plus DNA Ladder (Invitrogen, ThermoFisher Scientific Inc., USA)

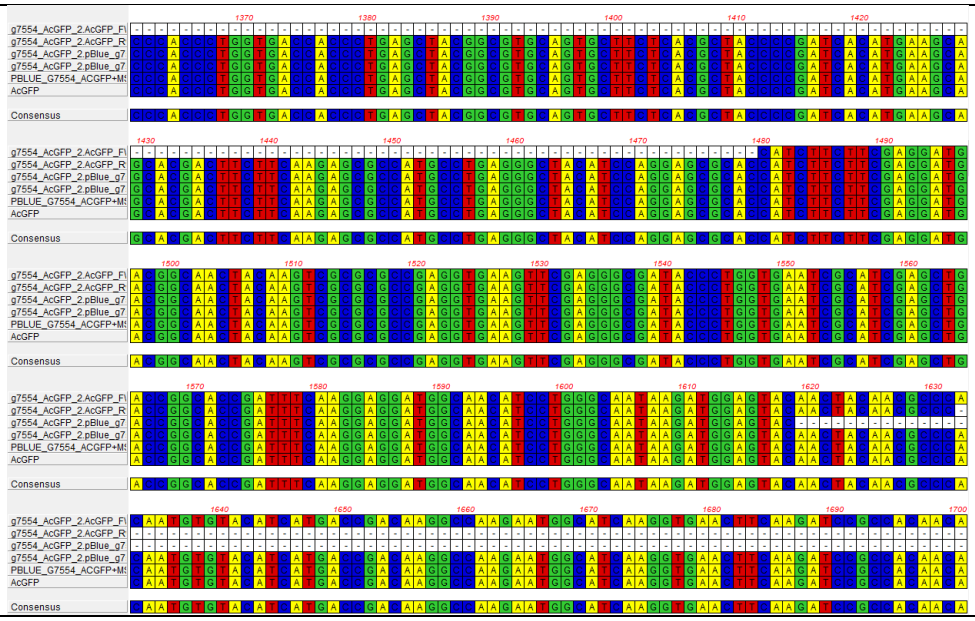
**Annex 10. Results of the plasmid sequencing.** The graphs show a section of the sequence alignments of each insert as the representation of the correct plasmid sequence.



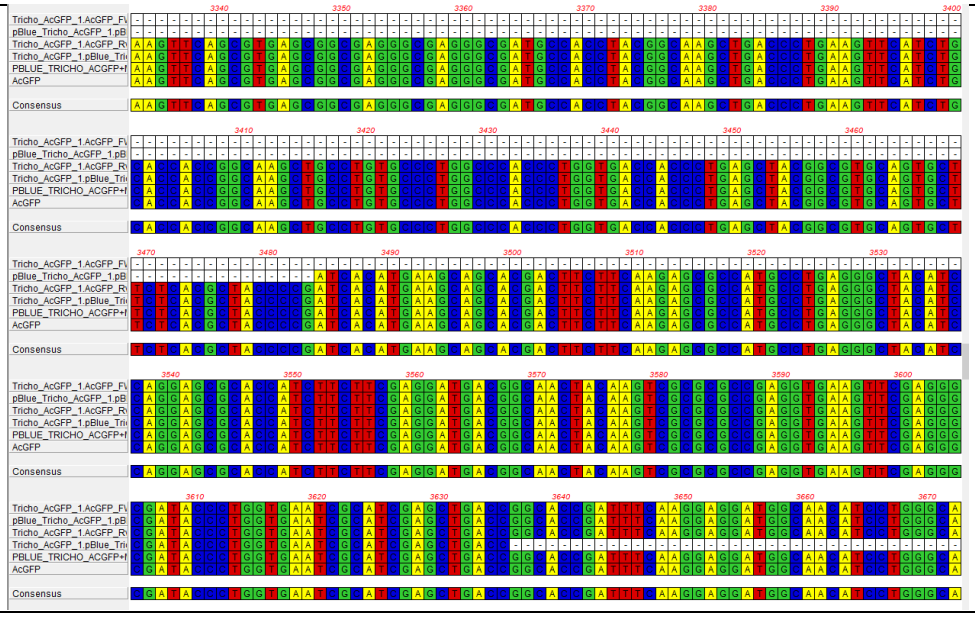

A section of the sequence alignment of pBlue_*g6464_AcGFP*opt as representation of the correct sequence of the plasmid. In the first two rows is the designed sequence of the whole plasmid and the insert as templet. The next two sequences have their starting point outside of the insert. The third and fourth sequences have their starting point in the GFP. The last row is the consensus sequence of the alignment after this the single sequences are shown. In the third row. In total ~1900bp of the construct were sequenced.




A section of the sequence alignment of pBlue_*g7554_eGFP*opt as representation of the correct sequence of the plasmid. In the first row is the designed sequence of the insert as templet. The next sequence has their starting point outside of the insert. The third and fourth sequences have their starting point within the promoter. The next two sequences have their starting point within the GFP. The seventh sequence has the starting point outside of the insert. The last row is the consensus sequence of the alignment after this the single sequences are shown. In the third row.
In total ~2500bp of the construct were sequenced.



A section of the sequence alignment of pBlue_*g7554_AcGFP* as representation of the correct sequence of the plasmid. The first two sequences have their starting point in the GFP. The third and fourth sequences have their starting point outside of the insert. The next two sequences have their starting point within the GFP. The next two rows are the designed sequence of the whole plasmid and the insert as templet. The last row is the consensus sequence of the alignment after this the single sequences are shown. In the third row.
In total ~1600bp of the construct were sequenced.



A section of the sequence alignment of pBlue_Tricho_*AcGFP* as representation of the correct sequence of the plasmid. The first two sequences have their starting point in the GFP. The third and fourth sequences have their starting point outside of the insert. The next two sequences have their starting point within the GFP. The next two rows are the designed sequence of the whole plasmid and the insert as templet. The last row is the consensus sequence of the alignment after this the single sequences are shown. In the third row.
In total ~1600bp of the construct were sequenced.