# Validation of drug targets in malaria parasites predicted by metabolic modelling using CRISPR-Cas9

Supannee Taweechai

Submitted in accordance with the requirements for the degree of Master of Science by Research

The University of Leeds

School of Biology

September 2018

I confirm that the work submitted is my own and that appropriate credit has been given where reference has been made to the work of others.

## Acknowledgements

I would like to express my deepest and sincere gratitude to my supervisor, Prof. Glenn McConkey for his valuable supervision, creative guidance, especially his kindness throughout the year of my study. I also wish to give particular thanks to Dr. Francis Isidore Garcia Totañes who taught me how to culture such a complicated protozoan parasite, helped me solving technical problems and gave me practical advice through my project. In addition, Thanks to all members in GMc group and everyone in Room. 8.11 for their kind friendship and assistance.

I would like to thank the Royal Thai Government for the generous support of a full scholarship to study in this program. I would also like to thank all the supervisors at BIOTEC for giving me the great opportunity to study here. Thanks to all friends in MPMB lab for morale boosting

Lastly, and most importantly, I thank my family for their unflagging love and unwavering support. I have no words to fully express my gratitude.

## Abstract

In recent years, despite the significant decline in number of malaria infections, malaria continues to be a threatening disease to humans with the emergence of artemisinin-resistant parasites. During the symptomatic, infection stage, malaria treatments rely on antimalarial chemotherapy to control the progress of infection. Hence, new drugs that can control infection and prevent transmission are urgently needed. In the early phase of drug discovery, metabolic modeling has been conducted to model growth at a genomic level and predict drug targets against *Plasmodium falciparum*, leading to the uncovering of 18 new predicted drug targets. However, those targets need to be verified by genetic or chemical target validation. Herein, two candidate drug targets; UMP-CMP kinase and dicarboxylate/ tricarboxylate carrier (DTC) protein were examined for necessity using gene knockouts by clustered regularly interspaced short palindromic repeats and Cas9 endonuclease-mediated genome editing (CRISPR-Cas9). The gene disruption of both target loci along with the non-essential KAHRP gene was validated through genotyping analysis and confirmed by DNA sequencing. Subsequently, phenotypic analysis by microscopy implicated the essentiality of DTC and UMP-CMP kinase as, there were no visible viable parasites by thin blood smearing from transgenic parasites, while KAHRP2 transgenic parasites were observed on day 20 post-transfection. However, quantitative RT-qPCR data is important to validate the visual observations; comparing with the KAHRP2 transgenic cultures. Demonstrating essentiality of UMP-CMP kinase and DTC is the first step in targeting these proteins for future drug development. Furthermore, setting up the gene disruption system here could be applied to other predicted malaria drug targets in the future.

## Table of contents

## Lists of figures

## List of table

## Abbreviations

| | |
|---|---|
| CMP | Cytidine monophosphate |
| CTP | Cytidine triphosphate |
| DC | Dicarboxylate carrier |
| dCMP | Deoxycytidine monophosphate |
| dCTP | Deoxycytidine triphosphate |
| DTC | Dicarboxylate/tricarboxylate carrier |
| dTTP | Deoxythymidine triphosphate |
| GOI | Gene of interest |
| *hdhfr* | Human dihydrofolate treductase |
| HR | Homology region |
| Hs | *Homo sapiens* |
| KAHRP | Knob-associated histidine rich protein |
| mito OxoC | Mitochondrial 2-oxodicarboxylate carrier isoform 1 |
| sgRNA | Single guide RNA |
| TCT iso_A | Tricarboxylate transport protein isoform a |
| TCT iso_B | Tricarboxylate transport protein isoform b. |
| UMP | Uridine monophosphate |
| UTP | Uridine triphosphate |

# Chapter 1 Introduction

## 1.1 General information

Malaria is a major tropical disease caused by a parasite in the genus *Plasmodium*. There are four well-established malaria parasites that infect humans. The two species that mainly infect humans, *Plasmodium falciparum* and *Plasmodium vivax*, are found in most endemic areas, while *Plasmodium ovale* and *Plasmodium malariae* are less commonly found in human cases and only have a sporadic and limited distribution (Snounou *et al.* 1993, Al-Maktari *et al.* 2003, Naing *et al.* 2014). Recently, the simian malaria parasites, *Plasmodium cynomolgi* and *Plasmodium knowlesi*, have been found in some cases of naturally acquired infection in humans; however, their transmission from human to human by an *Anopheles* mosquito has not been recorded (White 2008, Ta *et al.* 2014, Ahmed and Cox-Singh 2015).

Malaria is found in tropical and sub-tropical regions. According to the World Malaria Report 2017, there were an estimated 216 million malaria infections globally (in 91 countries) causing 445,000 deaths in 2016. Infection by the virulent malaria parasite, *P. falciparum*, is predominant in sub-Saharan Africa and is responsible for 99% of all estimated malaria cases. Recently, although the number of malaria infections has significantly declined, malaria remains a threatening disease, particularly in children under 5 years of age residing in Africa. In addition, an emergence of artemisinin resistance has been reported along the Thailand-Myanmar border, leading to great concerns about malaria elimination programmes. Moreover, the effect of global warming may accelerate malaria spread; a temperature higher than 21°C is suitable for *P. falciparum* mating and maturation in the vector, placing an increasing proportion of the world population at risk of malaria infection (World Health Organization 2017)

## 1.2 Malaria life cycle

Malaria is caused by a protozoa parasite of *Plasmodium spp*. that has a complex multistage life cycle, with sexual reproduction in the *Anopheles* mosquito and an asexual reproductive stage within the red blood cells of the

human host (Figure 1.1). The sporozoite stage in the salivary glands of the mosquito is transmitted to a human host while the mosquito takes a blood meal. Once the sporozoite enters the host, within an hour the sporozoite invades the hepatocytes where then they replicate, divide, and transform into schizonts, which contain thousands of merozoites. Next, the infected hepatocytes rupture and the merozoites are released into the bloodstream, beginning the asexual reproductive stage. After the merozoites invade the erythrocytes, they envelop themselves within a layer that becomes a parasitophorous vacuole membrane when the invasion is complete. Over the subsequent 48 hours, *P. falciparum* parasites develop within the vacuole from trophozoites (ring-form) into schizonts that contain 12-16 daughter merozoites. The merozoites released from the erythrocytes invade fresh red blood cells to continue the cycle of parasite multiplication. In this stage, some parasites differentiate into gametocytes, which are the sexual form of the parasites. The transcription factor AP2-G has been shown to regulate the gametocytogenesis transition (Kafsack *et al.* 2014). When a mosquito ingests the sexual form of parasites during a blood meal, fusion of the male and female gametes occurs in the stomach of the mosquito to form a zygote. Then the zygote elongates into an ookinete (a motile zygote). After that, the ookinete migrates across the midgut epithelium and develops into an oocyst. The oocyst undergoes a series of mitoses to form sporozoites. Once the oocyst has matured and ruptured, the sporozoites are released and they migrate to the mosquito salivary glands, where the parasites can infect a new host to complete its life cycle.

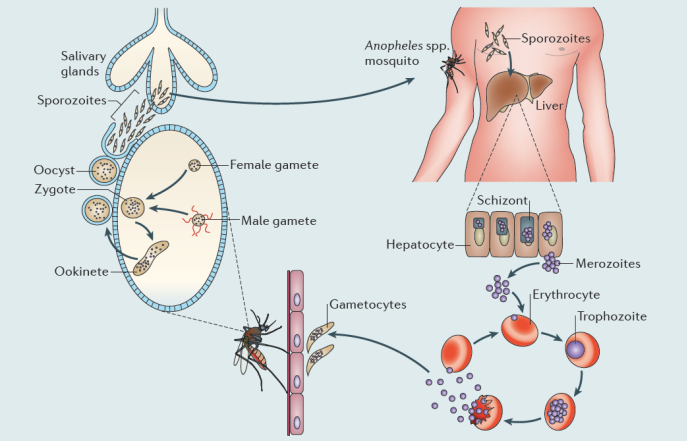



**Figure 1.1 The life cycle of *Plasmodium spp*.**
The malaria life cycle can be divided into 2 stages; the asexual stage in humans and the sexual stage in mosquitoes. The figure was taken from (de Koning-Ward *et al.* 2015).

## 1.3 Malaria treatment and drug resistance

Throughout history, malaria fevers have been treated with herbal remedies, namely cinchona bark and qinghao. Then, quinine was isolated from cinchona bark in 1820 and used for malaria treatment as a pure chemical compound thereafter. However, the global quinine supply was cut off during World War II, inducing an effort to develop new synthetic antimalarial drugs. Eventually, a powerful antimalarial, chloroquine, was successfully synthesized in the 1940s and widely used throughout the world (Meshnick and Dobson 2001). Chloroquine was used as the first line treatment for two decades before the emergence of chloroquine-resistant parasites was reported from South America and South-East Asia in the early 1960s. Then, the *P. falciparum* resistance to chloroquine spread to Africa during the 1980s (D'alessandro and Buttiens 2001, Trape 2001). The mechanism of parasite resistance to chloroquine has been previously discussed. Mutations in the chloroquine resistance transporter (*Pf*CRT) and multidrug resistance gene (*pfmdr*1), coding for the membrane transporter (*Pf*MDR1), can contribute to chloroquine resistance in *P. falciparum*. Mutations occurring in both targets causes the concentration of chloroquine in the digestive vacuole to decrease considerably, allowing the parasites to avoid chloroquine drug pressure (Anderson *et al.* 2005, Cojean *et al.* 2006). Despite the widespread presence of chloroquine resistance, most countries in Africa are still using this drug as the first-line treatment since chloroquine is the cheapest antimalarial drug that is available on the market. Some areas (Malawi, Kenya, Botswana, and South Africa) use sulphadoxine-pyrimethamine (SP) instead of chloroquine (D'alessandro and Buttiens 2001). However, SP resistant parasites have also been identified. These drugs have a prolonged half-life, leading to selection of resistant parasite strains that have developed a mutation in the active site of the drug target (Triglia *et al.* 1997, Nzila *et al.* 2000, Mita *et al.* 2009).

Another effective antimalarial drug, artemisinin, is the active ingredient that was extracted from a herb (sweet wormwood) and has been used to cure malaria infections since 1979. This compound was successfully synthesized 4 years later and has been used as a first line treatment since then. Artemisinin is considered to be a fast-acting drug that rapidly reduces parasitaemia up to 10000-fold every 48 hours (White 2008). Due to the occurrence of SP-resistant parasites, an alternative drug or new treatment regimen was urgently needed. Hence, artemisinin-based combination therapies (ACT) were implemented for

malaria treatment in all endemic countries. For combination therapy, artemisinin or artemisinin derivatives are used as a partner drug with other effective antimalarial drugs with an unrelated mode of action, such as lumefantrine, amodiaquine, mefloquine, piperaquine, or sulfadoxine-pyrimethamine as first-line treatment of uncomplicated *P. falciparum* malaria (Guidelines for Treatment of Malaria Third Edition, World Health Organization, 2015) (World Health Organization 2015)

However, cases of uncomplicated *P. falciparum* infection with decreasing artemisinin susceptibility were reported in the Greater Mekong subregion in 2009. The resistance of *P. falciparum* to artemisinin derivatives presents by delaying the rate of clearance of parasites after monotherapy (Dondorp *et al.* 2009, Noedl *et al.* 2009). The mechanism of artemisinin resistance is still unclear. *Pf*kelch13 has been proposed as an excellent marker for artemisinin resistance in *P. falciparum* after genome sequencing of parasites isolated from a patient showed mutations in the K13 propeller protein (Ariey *et al.* 2014, Miotto *et al.* 2015). Moreover, enhancement of the cell stress response by the accumulation of ubiquitinated proteins and increases in proteasome activity have also been found in artemisinin-resistant parasites (Mbengue *et al.* 2015).

The emergence of artemisinin resistance and simultaneous resistance to the partner drugs (summarised in Table 1.1) is a serious obstacle to artemisinin-based combination therapies and the Global Malaria Eradication Programme (Phyo and Nosten 2018). Thus, there is an urgent need for new drugs that are safe to use and have novel modes of action against these resistant parasites. Because malaria treatment relies on antimalarial chemotherapy to control the parasites (Schlitzer 2007) and there are no effective malaria vaccines currently available (Cowman *et al.* 2016), target identification and validation are important steps to take to initiate early drug discovery.

**Table 1.1 Summary of the emergence of antimalarial resistant drugs**

| Antimalarial drug | Year of first deployment | Place of first deployment | Year of resistance emerged | Place of emergence of resistance |
|---|---|---|---|---|
| Quinine | 1630 | South America | 1910* | Brazil |
| Chloroquine | 1945 | Global malaria eradication campaign | 1957 | Colombia, Cambodia-Thailand border |
| Amodiaquine | 1948 | Americas | 1961 | Colombia |
| Atovaquone | 1996 | Thailand | 1996 | Thailand |
| Proguanil | 1948 | Various African countries | 1949 | Aden Protectorate, Yemen |
| Sulfa + antifols[b] | 1967 | Thailand | 1967 | Thailand |
| Meloquine | 1967 | Vietnam | 1982 | Thailand |
| Piperaquine | 1978 | China | 1985 | China |
| Artemisinin | 1979 | China | 2008 | Cambodia |
| Meloquine-artesunate | 1994 | Thailand | 2002[a] | Cambodia |
| Artemether-lumefantrine | 1994 | China | 2006[a] | Cambodia |
| Dihydroartemisinin-piperaquine | 2001 | Cambodia | 2013[a] | Cambodia |

*There is no high-grade resistance to quinine.

[a]Therapeutic efficacy <90% (cut-off threshold of WHO to switch the ACT policy).

[b]Sulfa + antifols: Sulfadoxine + antifolates.

The table was adapted from (Phyo and Nosten 2018).

## 1.4 Target identification

Approximately 5300 genes (that are highly conserved and probably have essential functions) are found in *Plasmodium* genomes (Janse *et al.* 2011); however, based on the information from *P. falciparum* and *P. berghei*, only around 500 genes have been investigated by gene disruption methods (de Koning-Ward *et al.* 2015). Therefore, the remaining genes need to be further studied in regards to their function. To narrow down the list of possible therapeutic targets, an *in silico* system has been implemented for prediction of antimalarial drug target candidates. Using a highly curated metabolic network model of *P. falciparum*, iFT342 was developed by Francis Isidore Garcia Totañes as part of his Ph.D. project at the University of Leeds in 2017 (Totanes

2017), identifying 18 novel drug targets. The constraint-based genome-scale model used flux analysis to solve for increasing biomass as a measure of growth and an *in silico* simulation of single gene knockout analysis to identify essential genes, resulting in a higher percentage of true positive predictions compared with other malaria models.

### 1.5 Target validation

Once a target has been identified, it then needs to be analysed for crucial biological functions. Validation techniques in *Plasmodium* parasites require a genetic manipulation approach (e.g., gene knockout or introduction of a transgene) to directly assess the function of the predicted gene target. The conventional gene knockout method is the most common technique in use for investigation of gene function in *Plasmodium* parasites. This method involves the introduction of foreign DNA into the malaria parasite using DNA transfection. The success of the conventional method depends upon recombination of homologous sequences between the engineered plasmid and the target genomic site via spontaneous single/double crossover of the homology regions. The appropriate targeting sequence and a good selection approach are necessary for success when performing crossing homologous recombination. Using circular plasmids for transfection usually results in the episomal plasmid being maintained rather than the integration event, introducing the requirement of another drug cycle or negative selection, which is time-consuming. Moreover, the stable transfectants require regular genetic monitoring because applying on/off drug pressure could lead to a naturally drug-resistant mutant. Also, homologous integration is a rare event and hence there is a very low efficiency of generating knockouts (Crabb and Cowman 1996, Waterkeyn *et al.* 1999, Crabb *et al.* 2004). Given these limitations, the conventional gene knockout method is an inefficient and time-consuming technique. Therefore, there is a need for a robust and efficient technique to study gene function in malaria parasites.

In the past decade, genetic tools for use with *P. falciparum* have developed considerably; for instance, the fusion of the gene of interest (GOI) to a destabilization domain (DD) or an *Escherichia coli* DHFR destabilizing domain (DDD) have been developed to control the expression of essential genes since the parasites often die after an indispensable gene is disrupted (Armstrong and Goldberg 2007, de Koning-Ward *et al.* 2015). Conditional knockout using

the dimerizable Cre (DiCre) system, has also been used for controlling the deletion of essential genes (Collins *et al.* 2013). However, these approaches do not improve the frequency of integration and are not suitable for large-scale gene deletion.

More recently, a new system that utilizes programmable nucleases to make double-stranded breaks in the genome has been developed for improving the integration frequency along with site-specific modifications. It is well-known that a chromosome break is a catastrophic event in the cell, leading to cell death unless it can be repaired. By providing a donor DNA for homologous directed recombination, the cell can repair the broken-double strand through the insertion of the provided sequence, resulting in the genome being modified in any way desired. There are two techniques in current use that use a nuclease to induce targeted chromosome breaks in *Plasmodium* parasites: zinc finger nucleases (ZFNs) and the clustered regularly-interspaced short palindromic repeat (CRISPR) associated Cas9 protein (CRISPR-Cas9) system (Straimer *et al.* 2012, Ghorbal *et al.* 2014).

ZFNs have a DNA binding domain (ZFN-L and ZFN-R) for DNA recognition and contain a catalytic domain of Fok1 for cleavage. The donor plasmid encoding a pair of ZFNs (ZFN-L and ZFN-R) is co-expressed using a single promoter of viral 2A ribosome skip peptide, allowing the individual polypeptides to be expressed. When the donor plasmid is transfected into malaria parasites, the pair of ZFNs is expressed and forms a dimer of ZFN-L and ZFN-R, resulting in the assembly of an artificial nuclease enzyme, which induces a double-stranded DNA break in the genome at the specified target site. Subsequently, the broken DNA is repaired by homology-directed recombination using homologous regions provided from the donor plasmid due to the absence of a non-homologous end joining (NHEJ) system in *Plasmodium* spp. The use of this system has been successfully demonstrated in *P. falciparum* and the integrant parasites are detected after only 2 weeks; nevertheless, there are some disadvantages to this system. For example, it is costly and difficult to design ZFNs because of the high AT content of *P. falciparum* (Straimer *et al.* 2012, Lee *et al.* 2014, Lee and Fidock 2014).

The CRISPR-Cas9 system is a prokaryotic adaptive immune mechanism defence against viruses or other invading DNA. CRISPR associated with *cas* genes is involved in acquired immunity against the bacteriophages of *Streptococcus thermophiles* and was first described by Barrangou et al (2007).

They found a series of repeat sequences interspersed with short fragments of viral nucleic acid integrated into the bacteria cells that conferred resistance to phages (Barrangou *et al.* 2007). In 2008, Brouns et al. identified the mature CRISPR RNAs (crRNAs) associated with Cas protein from the host that was used to interfere with virus proliferation (Brouns *et al.* 2008). Then, Jinek et al. (2012) showed that a dual RNA structure composed of the mature crRNA base forms a duplex with the trans-activating tracrRNA and is crucial to direct the *S. pyogenes* type II Cas9 protein (spCas9) to cleave double-stranded DNA at the specific target DNA sequences *in vitro* (Jinek *et al.* 2012).

Based on a programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity of *S. pyogenes*, this system has been applied for gene editing in malaria parasites, first demonstrated by Ghorbal et al (2014). In this method, the parasites are transfected with two plasmids, one being the pUF1-Cas9 plasmid that provides a Cas9 endonuclease for making the double-strand break (DSB) in a site specific to the target DNA guided by the sgRNA, and the second being a plasmid (pL7-GOI) that contains the sgRNA and the donor DNA template for homology recombination. Regarding the pUF1-Cas9 plasmid, a *cas* gene from *S. pyogenes* was engineered to express the endonuclease Cas9 bearing nuclear localization signals (NLS) on both sides of nuclease under the control of plasmodial regulatory element while the sgRNA in the CRISPR plasmid was transcribed through the U6 small nuclear (sn) RNA regulatory elements. Once the double-stranded DNA is broken by Cas9, the cell repairs itself via its endogenous repair mechanism using the insertion of the homologous sequences of interest provided on the transfected DNA vector (the donor DNA). Using this system, they also showed the integrated parasites were achieved within 8 days of transfection (Ghorbal *et al.* 2014).

In the CRISPR plasmid constructed by this group, the original plasmid pL6 eGFP can be modified to induce gene disruption in any targeted loci in *P. falciparum* by adapting the homology sequences and by introducing a new targeting sgRNA into the plasmid. The newly engineered plasmid is then co-transfected with the unmodified Cas9 plasmid. Following a demonstration of the well-established CRISPR-Cas9 system in *P. falciparum* by Ghorbal et al (2014), this simple method is robust and has a high efficiency and precise integration into the target locus and could be a promising approach for large-scale gene disruption. Thus, two potential drug targets (UMP-CMP kinase and

dicarboxylate/ tricarboxylate carrier protein) have been chosen for establishing the CRISPR-Cas9 system in our laboratory as a model, with plans to test other predicted malaria drug targets in the future.

### 1.6 UMP-CMP kinase as a drug target for antimalarial chemotherapy

The gene for UMP-CMP kinase (PF3D7_0111500, EC 2.7.4.14) contains 1116 bp, encoding 371 amino acids, and is a bifunctional enzyme. It plays an important role in nucleic acid synthesis by converting UMP, CMP, and dCMP into their diphosphate form. In malaria parasites, UMP is the first pyrimidine nucleotide that acts as the precursor for all pyrimidine nucleotides biosynthesis. Then, these diphosphate forms are converted into UTP, CTP, dTTP and dCTP, which are essential for DNA and RNA synthesis (Krungkrai and Krungkrai, 2016). In humans, this enzyme is pivotal in some anticancer and antiviral therapies by converting nucleoside analogues into their triphosphorylated active form through a phosphorylation reaction (Liou *et al.* 2002, Segura-Peña *et al.* 2004). A comparison of UMP-CMP kinase from various species (human, slime mould, yeast, pig, and mouse) showed that this enzyme is highly conserved, suggesting a crucial function of this kinase across species (Segura-Peña *et al.* 2004). For *P. falciparum*, it shares 44.79% identity with humans (Figure 1.2) and the active site of this enzyme is strongly conserved, corresponding to other species. Since UMP-CMP kinase is essential in anticancer treatment (leukaemia, lymphoma, and solid tumours) and the malaria parasite is also a highly proliferating organism, this enzyme could possibly be a target for antimalarial agents.

In addition, the data from gene expression in *P. falciparum* (strain 3D7) using Illumina RNA-seq (López-Barragán *et al.* 2011) revealed that high expression of the UMP-CMP kinase gene is found in the late trophozoite and the schizont stages (as shown in Figure 1.3), stages characterized by having an exponential replication of parasites in red blood cells. Hence, cutting off production of nucleotides used for DNA and RNA synthesis could be a way to control the number of parasites. Moreover, high expression of this gene is also found in the sexual stages: gametocyte II, gametocyte v, and ookinete, suggesting it could also be a target for blocking or perturbing parasite transmission. UMP-CMP kinase is predicted to be a novel drug target with a potentially lethal effect by using the *in silico* simulation of single gene knockout analysis from metabolic modelling (iFT342). UMP-CMP kinase enzyme has

not been characterized in *Plasmodium spp*. until recently. We conclude that UMP-CMP kinase could be a good drug target for elimination and control of malaria.

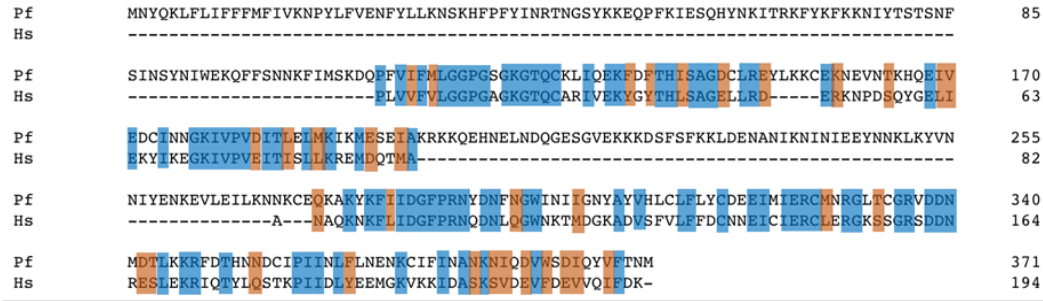



**Figure 1.2 Homology and structure-based alignment of UMP-CMP kinase**

The structure- based sequence alignment compared between UMP-CMP kinase from *P. falciparum* (PF3D7_0111500) and human enzyme (1tev.1.A, short- form open conformation), was generated using the protein structure homology modelling server Swiss Model. Blue highlight depicts identical amino acids, orange highlight represents similar residues.


**Figure 1.3 Gene expression of UMP-CMP kinase**

Transcript levels of fragments per kilobase of exon model per million mapped reads (FPKM) were obtained from the Illumina-based sequencing of *P. falciparum* 3D7 mRNA from two gametocyte stages (II and V), ookinete, and four time points of erythrocytic stages representing ring, early trophozoite, late trophozoite, and schizont. The expression level defines as numbers of reads per kilobase per million mapped reads. The percentile graph shows the ranking of expression of the PF3D7_0111500 gene compared to all others in this experiment. The figure was taken from PlasmoDB (Plasmodium Genomic Resource).

## 1.7 Dicarboxylate/tricarboxylate carrier (DTC) protein as a drug target for antimalarial chemotherapy

Dicarboxylate/tricarboxylate carrier (PF3D7_0823900) is an integral membrane protein that is 318 amino acids long (957 bp). It belongs to the mitochondrial carrier protein family, which is responsible for the transportation of metabolites and cofactors across the inner membrane of mitochondria. This protein from *P. falciparum* has been expressed using a cell-free system and characterized by Nozawa et al (2011). They showed *Pf*DTC is able to exchange oxoglutarate–malate, oxoglutarate–oxaloacetate, or oxoglutarate–oxoglutarate in the artificial membrane (Nozawa *et al.* 2011). However, data from the *in vivo* experiment using $^{13}C$ isotope-labelled mutant parasites showed that *Pf*DTC does not import oxaloacetate into the mitochondria and instead is involved in mitochondrial $\alpha$-ketoglutarate-malate transportation of tricarboxylic acid metabolism (Ke *et al.* 2015). A comparison of *Pf*DTC with other mitochondrial membrane proteins from humans shows 32.62%, 21.38%, 22.59%, and 22.80% identity with dicarboxylate carrier protein, mitochondrial 2-oxodicarboxylate carrier isoform 1, tricarboxylate transport protein isoform a, and tricarboxylate transport protein isoform b, respectively (Figure 1.4). Hence, it is a good prospect for the development of parasite-specific inhibitors especially since there is high expression of the *pfdtc* gene in the late trophozoite stage (Figure 1.5). Blocking reproduction of parasites in the red blood cell stages is necessary for treatment of symptomatic disease. The highest level of expression of the *pfdtc* gene occurs in gametocyte II (López-Barragán *et al.* 2011), which could be a good target for malaria elimination. *Pf*DTC is predicted to have a growth limiting effect in metabolic modelling (iFT342) but gene disruption of *pfdtc* has not been previously performed. This gene is an attractive target for drug discovery.

```
Pf_DTC          -MDRDIAKYDLESSSSVDVNKKGNNNKSVFEKIKPFAVGGASGMFATFCIQPLDMVKVRI     59
Hs_DC           --------------------------MAAEARVSRWYFGGLASCGAACCTHPLDLLKVHL     34
Hs_mito OxoC    ----------------MSAKP---EVSLVREASRQIVAGGSAGLVEICLMHPLDVVKTRF     41
Hs_TCT iso_A    ------MPAPRAPRALAAAAPASG-KAKLTHPGKAILAGGLAGGIEICITFPTEYVKTQL     53
Hs_TCT iso_B    MFPAALARRPRRPKSGTGEGPERQRPGGSLRSGFPVPAGGLAGGIEICITFPTEYVKTQL     60
                                                 ** :.        * : :*.::

Pf_DTC          QLNAEGK--NVLRNPFIVAKDIIKNEGFLSLYKGLDAGLTRQVIYTTGRLGLFRTFSDMV     117
Hs_DC           QTQQEVK--LRMT---GMALRVVRTDGILALYSGLSASLCRQMTYSLTRFAIYETVRDRV     89
Hs_mito OxoC    QIQRCATDPNSYKSLVDSFRMIFQMEGLFGFYKGILPPILAETPKRAVKFFTFEQYKKLL     101
Hs_TCT iso_A    QLDER-SHPPRYRGIGDCVRQTVRSHGVLGLYRGLSSLLYGSIPKAAVRFGMFEFLSNHM     112
Hs_TCT iso_B    QLDER-SHPPRYRGIGDCVRQTVRSHGVLGLYRGLSSLLYGSIPKAAVRFGMFEFLSNHM     119
                * :   .               .: .*.:.:* *:   :  .      ::  :.   . :

Pf_DTC          K-KEGEPLPFYKKCFCALAAGGLGAFI-GNPADLSLIRLQADNT--LPKELKRNYTGVFN     173
Hs_DC           AKGSQGPLPFHEKVLLGSVSGLAGGFV-GTPADLVNVRMQNDVK--LPQGQRRNYAHALD     146
Hs_mito OxoC    G-YVS-LSPALTFAIAGLGSGLTEAIV-VNPFEVVKVGLQANRNTFAEQPS---TVGYA-     154
Hs_TCT iso_A    R-DAQGRLDSTRGLLCGLGAGVAEAVVVVCPMETIKVKFIHDQT--SPNPK---YRGFFH     166
Hs_TCT iso_B    R-DAQGRLDSTRGLLCGLGAGVAEAVVVVCPMETIKVKFIHDQT--SPNPK---YRGFFH     173
                              : .  :*   ..:   * :   : :  : .    :

Pf_DTC          ALYRISK--EEGLFALWKGSVPTIARAMSLNLGMLSTYDQSKEFL--Q--KYLGVGMKTN     227
Hs_DC           GLYRVAR--EEGLRRLFSGATMASSRGALVTVGQLSCYDQAKQRVLST--GYLSDNIFTH     202
Hs_mito OxoC    --RQIIKKEGWGLQGLNKGLTATLGRHGVFNMVYFGFYYNVKNMIPVN--KDPILEFWRK     210
Hs_TCT iso_A    GVREIVR--EQGLKGTYQGLTATVLKQGSNQAIRFFVMTSLRNWYRGDNPNKPMN-PLIT     223
Hs_TCT iso_B    GVREIVR--EQGLKGTYQGLTATVLKQGSNQAIRFFVMTSLRNWYRGDNPNKPMN-PLIT     230
                   .: :    **    .* . :  :        :    . ::

Pf_DTC          LVASVISGFFAVTLSLPFDFVKTCMQKMKADPVTKKMPYKNMLDCSIQLYKKGGISIFYS     287
Hs_DC           FVASFIAGGCATFLCQPLDVLKTRL-------MNSKGEYQGVFHCAVETAKLGPLA-FYK     254
Hs_mito OxoC    FGIGLLSGTIASVINIPFDVAKSRIQGPQP--VPGEIKYRTCFKTMATVYQEEGILALYK     268
Hs_TCT iso_A    GVFGAIAGAASVFGNTPLDVIKTRMQGLEA------HKYRNTWDCGLQILKKEGLKAFYK     277
Hs_TCT iso_B    GVFGAIAGAASVFGNTPLDVIKTRMQGLEA------HKYRNTWDCGLQILKKEGLKAFYK     284
                   . ::*  :     *:*. *: :             *:   .      :   :  :*.

Pf_DTC          SYATYYVRIAPHAMITLITVDYLNNLLKK-IS--  318
Hs_DC           GLVPAGIRLIPHTVLTFVFLEQLRKNFGIKVPS-  287
Hs_mito OxoC    GLLPKIMRLGPGGAVMLLVYEYTYSWLQENW---  299
Hs_TCT iso_A    GTVPRLGRVCLDVAIVFVIYDEVVKLLNKVWKTD  311
Hs_TCT iso_B    GTVPRLGRVCLDVAIVFVIYDEVVKLLNKVWKTD  318
                .      *:     : ::  :   . :
```

**Figure 1.4 Multiple amino acid sequence alignment of DTC**

Amino acid sequence alignment of DTC from *P. falciparum* compared with the mitochondrial membrane proteins from humans was generated using Clustal Omaga protein alignment. The accession number of the proteins: XP_001349285.1; Pf_DTC, CAB59892.1; Hs_DC, NP_085134.1; Hs_mito OxoC, NP_005975.1; Hs_TCT iso_A, NP_001243463.1; Hs_TCT iso_B. Abbreviations used in the figure: Pf; *P. falciparum*, Hs; *Homo sapiens*, DTC; Dicarboxylate/tricarboxylate carrier, DC; dicarboxylate carrier protein, mito OxoC; mitochondrial 2-oxodicarboxylate carrier isoform 1, TCT iso_A; tricarboxylate transport protein isoform a, TCT iso_B; tricarboxylate transport protein isoform b. (*) amino acid positions with identical residue, (:) amino acid positions with conserved substitution amino acid, (.) amino acid positions with semi-conserved substitution amino acid, and (non-marked) amino acid with non-conserved substitution amino acid.

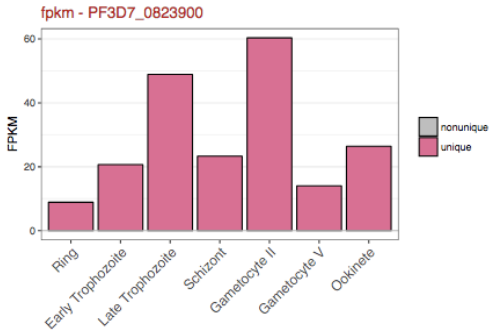



**Figure 1.5 Gene expression of DTC**
Transcript levels of fragments per kilobase of exon model per million mapped reads (FPKM) were obtained from the Illumina-based sequencing of *P. falciparum* 3D7 mRNA from two gametocyte stages (II and V), ookinete, and four time points of erythrocytic stages representing ring, early trophozoite, late trophozoite, and schizont. The expression level defines as numbers of reads per kilobase per million mapped reads. The percentile graph shows the ranking of expression of the PF3D7_0823900 gene compared to all others in this experiment. The figure was taken from PlasmoDB (Plasmodium Genomic Resource).

## 1.8 Research objectives

The main objective of this study is to utilise the CRISPR-Cas9 system to genetically validate two candidate drug targets in malaria parasites that were predicted by our metabolic modelling. This research covers the specific objectives:

1. Establish the CRISPR-Cas9 knockout system in our laboratory using the non-essential *P. falciparum* KAHRP gene as a positive control for the CRISPR-Cas9 system.
2. Knock out UMP-CMP kinase and dicarboxylate/tricarboxylate carrier using the CRISPR-Cas9 system with negative controls for transfection.
3. Analyse the target-gene disruptions using genotyping and confirm the integration by Sanger sequencing.
4. Evaluate gene essentiality using quantitative RT-qPCR by comparing it with the non-essential gene (KAHRP).

# Chapter 2 Methodology

## 2.1 Identification of guide RNA sequences

The sequences of the genes of interest (UMP-CMP kinase and DTC) and the entire genomic sequence of *P. falciparum* 3D7 were downloaded from PlasmoDB (Plasmodium Genomic Resource) to identify guide RNA sequences. The guide RNA sequence was designed according to the following characteristics: the single guide RNA (sgRNA) sequence must be 20 bases long, match the target-DNA site, and the sgRNA must be followed by a protospacer adjacent motif (PAM) sequence (5'-NGG-3'). The sgRNA sequence must start with a G (the RNA polymerase III-dependent U6 promoter requires a G at the 5' end of the RNA sequence to initiate transcription), and base 17 must either be an A or a T (Cas9 endonuclease will cleave approximately three bases upstream of the PAM). The sgRNA plus PAM sequence must only occur once in the whole genome (both in the sense and the antisense strands) to avoid off-target activity (Naito *et al.* 2014). The sgRNA sequence must have a GC content that is greater than or equal to 40%. According to the criteria above, two sequences of sgRNA, GAGTAAAGATCAACCCTTTGTGG and GGTGGATTAGGAGCCTTTATTGG were identified that were specific to UMP-CMP kinase and DTC, respectively.

## 2.2 Designing primers for amplifying the homology regions

The size of the amplified homology regions (HR) was restricted to 250–500 bases in length and one of the homology regions must be close to the cutting site while the other homology region can be relatively far away. Based on the methodology of the In-Fusion cloning system (Takara Bio USA, Inc), the primers were designed with an adaptor sequence, which consists of 20 nucleotides of homology with the target vector. For amplification of HR1 of UMP-CMP kinase, a pair of oligonucleotide primers was designed as: forward primer: 5'-CTTTCCGCGGGGAGGACTAG(TGATCA)**ATGAATTACCAAAAG**

**TTATT**-3' and reverse primer: 5'-TTTTTTTACAAAATGCTTAA**AAATTGTTT**

**TTCCCAAATA**-3', which generated a PCR product of 322 bp. The underlined sequences are the homology nucleotides (adaptor sequence) between the PCR product and the vector, which is necessary for In-fusion base cloning.

The restriction site of *Bcl*I (represented in the bracket) was added to allow for further screening of positive clones and the bold sequences are the gene-specific sequences. The primers for amplifying HR2 of UMP-CMP kinase were: forward primer: 5'-TATTTATTAAATCTAGAATT(TGATCA)**GGATGGATTAAT**

**ATAATAGG**-3' and reverse primer: 5'-ATTATTTTTACCGTTCCATG**TTACAT**

**ATTTGTAAAAACATAC**-3', which produced a 453 bp PCR product. Regarding DTC, two pairs of primers to amplify HR1 and HR2 were created using the same strategy: forward primer: 5'-CTTTCCGCGGGGAGGACTAG(TGATCA)**ATGGACAGAGATATAGCTAA**-3' and reverse primer: 5'-TTTTTTTACAAAATGCTTAA**GTTCTCCTTCTTTTT**

**TTACC**-3', which amplified a 413 bp PCR product of HR1 while forward primer: 5'-TATTTATTAAATCTAGAATT(TGATCA)**CCTGCAGATTTATCTTTAA**-3' and reverse primer: 5'-ATTATTTTTACCGTTCCATG**TTGAATAGAACAATCT**

**AACAT**-3', were designed for amplification of HR2 of DTC with a product size of 436 bp.

## 2.3 Cloning of CRISPR plasmids and insertion of guide RNA sequences

In this study, plasmid pL6 eGFP (Figure 2.1) was obtained from Prof. Jose-Juan Lopez-Rubio (Biology of Host-Parasite Interactions Unit, Institut Pasteur, Paris, France). It was used as the template plasmid for generating CRISPR plasmids. Cloning of HR fragments and guide RNA sequences of GOI into the plasmid pL-6 eGFP to generate pL7 GOI plasmids was performed by replacing the GFP regions (GFP 5' and GFP 3') with the homology regions (HR1 and HR2) of GOI. The guide RNA sequences were then replaced at the *BtgZ*I adapter region (as illustrated in Figure 2.2). The multiple cloning steps were conducted by using the In-Fusion HD Cloning Kit (Takara Bio USA, Inc).

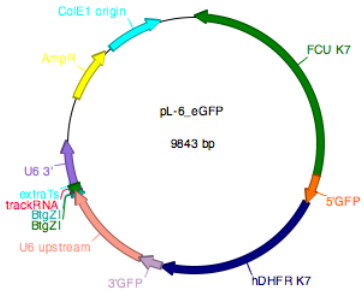



**Figure 2.1 An illustration of the pL6 eGFP plasmid**
The original plasmid pL6 eGFP used as the template plasmid for generating the CRISPR plasmids.

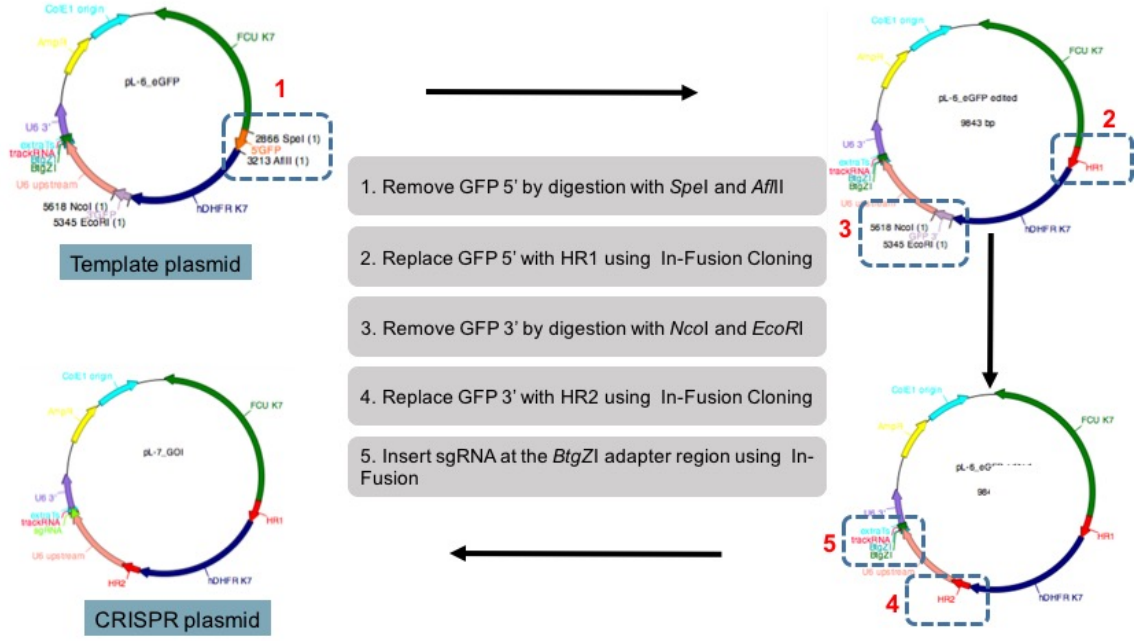



**Figure 2.2 The overall procedure for construction of the CRISPR plasmids**
The plasmid pL7 GOI was generated from the plasmid pL6 eGFP by replacing the GFP regions (GFP 5' and GFP 3') with the homology regions (HR1 and HR2) of GOI, then the targeting-oligonucleotides of sgRNA were added at the *BtgZ*I adapter region. All cloning steps were performed by using the In-Fusion HD Cloning Kit (Takara Bio USA, Inc). Then, the presence of the insertion and the correctness of its sequence were confirmed by Sanger sequencing.

### 2.3.1 Genomic DNA extraction

Genomic DNA was used as a temple for amplifying the homology region of the GOIs. To extract genomic DNA, the unsynchronized parasite *P. falciparum* strain 3D7 at 4% parasitaemia was cultured in a T75 flask and then pelleted by centrifugation at 3,000 rpm for 10 minutes. The cell pellet was then resuspended by adding a 5x pellet volume of sterile PBS buffer, gently mixing it, and adding a saponin solution at a final concentration of 0.15% to lyse the human red blood cells, then incubated at room temperature for 2 minutes with gentle mixing. To stop the reaction, PBS buffer was added to obtain a final volume of 30 ml. After transferring the solution into an Oakridge tube and centrifugation at 10,000 rpm for 10 minutes, the supernatant was discarded. Then 700 ul of lysis buffer (10mM Tris-HCl, 20mM EDTA, and 1% SDS) was added to the parasite pellet and mixed gently. Subsequently, 10 µl of RNase A (10 mg/ml) was added and incubated at 37°C for 1 hour. After that, 10 µl of Proteinase K (18.2 mg/ml) was added and incubated overnight at 50°C to digest the proteins and non-nucleic acid cellular components. Then, 700 µl of a mixture of phenol: chloroform: isoamyl alcohol (25:24:1) was added and gently mixed by inverting to separate the protein contaminants and cellular debris into the organic phase, followed by centrifugation at 13,000 rpm for 10 minutes. The isolated DNA in the aqueous phase was transferred into a clean 1.5 ml microcentrifuge tube. To increase the purity of genomic DNA, this step was repeated again. Before adding ethanol to precipitate the DNA, 700 µl of hydrated diethyl ether was added into the aqueous solution to clean up the organic compounds from the isolated DNA. This was followed by centrifugation at 13,000 rpm for 10 minutes. The top layer of ether was discarded and then repeated this step again to increase the 260/230 absorbance ratio. After centrifugation, the aqueous phase in the bottom portion was transferred into a sterile 1.5 ml microcentrifuge tube and a 1/10 volume of 3M sodium acetate pH 5.2 was added. Then, 2 volumes of absolute ethanol were added and incubated at −20°C for 30 minutes to precipitate the DNA. After centrifugation again at 13,000 rpm for 10 minutes, the supernatant was removed and the pellet was washed with 70% ethanol and centrifuged it again. The next step was removing the supernatant and allowing the DNA pellet to dry. Then the DNA was dissolved in nuclease-free water and the DNA concentration was measured by using a Nanodrop spectrophotometer, then stored at −20°C.

### 2.3.2 Amplification of homology regions

To optimise the PCR conditions for amplification of the homology regions (HR), the reaction was performed in 10 µl mixture reaction, contained 10 ng of gDNA template (*P. falciparum* strain 3D7), 500 nM of the forward primer, 500 nM of the reverse primer, 1× of Q5® High-Fidelity 2X Master Mix and the final volume was adjusted with sterile water. The PCR reaction was performed for 30 cycles; the first cycle consisted of the initial denaturation at 98°C for 30 seconds; the subsequent 10 cycles consisted of 98°C for 10 seconds, 52°C for 30 seconds, and 62°C for 30 seconds; followed by 20 cycles consisting of 98°C for 10 seconds, 62°C for 1 minutes and 62°C for 2 minutes. The optimised PCR condition here was applied for all HR amplifications except HR2 of UMP-CMP kinase. In order to obtain the expected band from HR2 of UMP-CMP kinase, the extension temperature was optimised from 62°C to 65°C while the other temperatures remained the same. The final large-scale PCR reaction was performed in a volume of 100 µl and followed the procedures as described above, then the amplified HR fragment was purified by using a gel extraction kit (QIAquick, catalogue no. 28704) and stored at −20°C.

### 2.3.3 Cloning the homology region into the pL6 eGFP plasmid

The original plasmid pL6 eGFP was digested with *Spe*I and *Afl*II to remove the GFP 5', where the purified fragment of HR1 was introduced into the plasmid. For the backbone plasmid preparation, a total of 5 µg of plasmid was used for the digestion reaction by separating it into 5 reactions. Each 50 µl reaction contained 1 µg of plasmid, 5 µl of the 10× CutSmart buffer, 0.5 µl each of the two restriction enzymes, and the rest of the volume was adjusted using sterile water. The mixture solution was then incubated at 37° for at least 2 hours. The digested reaction was further cleaned up by using a PCR purification kit (QIAquick, catalogue no. 28104). In this study, cloning of HR1 into the plasmid was performed by using the In-Fusion cloning kit. The 10 µl In-Fusion reaction contained 100 ng of digested plasmid, the amount of insert fragment that gave a 1:10 molar ratio of vector and insert, 2 µl of In-Fusion HD enzyme premix, and sterile water. An online NEB ligation calculator (https://nebiocalculator.neb.com/#!/ligation) used to calculate the required amount of the insert fragment that gave a 1:10 molar ratio of vector and insert. Then the In-Fusion cloning reaction was incubated at 50° for 15 minutes.

Subsequently, the heat shock method was used to transform the recombinant plasmid from the In-Fusion reaction into *Escherichia coli* XL10 gold ultracompetent cells. The 2.5 µl of In-Fusion reaction was incubated with 50 µl of competent cells on ice for 30 minutes then applied a heat-pulse in a 42°C water bath for 30 seconds, followed by a 2 minute- incubation on ice. Then, 250 µl of LB broth was added into the transformation reaction and incubated at 37°C with shaking at 200 rpm for 1 hour for plasmid amplification. After that, the transformants were plated on LB agar supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C. After that, plasmid extraction from a single colony was verified by Sanger sequencing before cloning of the HR2 fragment into the plasmid.

The pL6 plasmid carrying the HR1 fragment was then digested with *EcoR*I and *Nco*I to remove the original GFP 3' and replace it with the HR2 of GOI. The digestion reaction was performed as described above. For the cloning step, a ratio of 1:10 of digested plasmid and HR2 fragment was used in the In-Fusion reaction, then the transformation protocol followed all of the steps as described for the HR1 cloning. Colonies were selected from transformants of the In-Fusion reaction, then the insertion of HR2 and its correctness were confirmed by Sanger sequencing.

### 2.3.4 Insertion of guide RNA sequences into the pL6 plasmid

The annealed oligonucleotides of sgRNA were introduced into the pL6 plasmid carrying HR1 and HR2 of GOI by replacing them at the *BtgZ*I adapter region (Figure 2.2) through the use of the In-Fusion cloning kit. According to the requirements of In-Fusion based cloning, the oligonucleotides were designed with adaptor sequences (20 nucleotides of homology with the target vector). The oligonucleotides of sgRNA that were specific to UMP-CMP kinase were designed: Oligo 1: CATATTAAGTATATAATATT**GAGTAAAGATCAACCCTT**

**TGTGG**GTTTTAGAGCTAGAAATAGC and Oligo 2: GCTATTTCTAGCTCTA

AAAC**CCACAAAGGGTTGATCTTTACTC**AATATTATATACTTAATATG. The underlined sequences are adaptor sequences and the bold sequences represent the sgRNA that was identified for UMP-CMP kinase. The oligonucleotides of the guide RNA that were specific to DTC were: Oligo 1: CATATTAAGTATATAATATT**GGTGGATTAGGAGCCTTTATTGG**GTTTTAG

AGCTAGAAATAGC and Oligo 2: GCTATTTCTAGCTCTAAAAC**CCAATAAA**

**GGCTCCTAATCCACC**AATATTATATACTTAATAT, created using the same strategy as for UMP-CMP kinase.

To insert the sgRNA into the plasmid, the annealing of oligonucleotides was performed by following this procedure: the annealing reaction contained 10 µl of the plus strand oligonucleotides and 10 µl of the minus strand oligonucleotides (at a 100 µM concentration) and 2.2 µl of NEB buffer 2 in a sterile PCR tube. Consequently, the annealing program consisted of step 1, heating at 95°C for 10 minutes; followed by step 2, repeated 95°C, 1 segment, then in the step 3 the reaction was gradually cooled with a reduction of 0.6°C/second for 16-times. After that, in step 4, it was held at 85°C for 1 minute, repeated 85°C, 1 segment in step 5, followed by step 6 with a reduction of 0.6°C/second for 16-times. Then in step 7, 75°C for 1 minute, repeated 75°C, 1 segment in step 8 and gradually cooled with a reduction of 0.6°C/second for 16-times in step 9, then repeated the same cycle until reaching 25°C in step 22, ended with 25°C for 1 minute. After that, the reaction was held at 4°C. The annealed oligonucleotides (45 µM) were placed on ice and then were diluted to 500 nM in cold 5 mM Tris buffer pH 8.0. The oligonucleotides duplex was stored in a −20°C freezer.

Cloning of the oligonucleotides duplex of sgRNA into the pL6 plasmid (carrying HR1 and HR2 of GOI), the backbone plasmid was prepared in parallel by digestion with *BtgZ*I. The digestion reaction was performed in 20 µl reaction (in a total of 5 reaction tubes), contained 1 µg of plasmid, 2 µl of the 10× CutSmart buffer, 0.5 µl of restriction enzyme *BtgZ*I, and the final volume was adjusted to 20 µl with sterile water. The reaction was incubated at 60°C for 1 hour, then 0.5 µl of additional restriction enzyme was added and incubated for another hour at the same temperature to ensure that the plasmid was completely digested. Subsequently, the digested reaction was cleaned up by using a PCR purification kit. Then, the In-Fusion reaction was performed in a 10 µl reaction containing 100 ng of digested plasmid, 1.5 µl of annealed oligonucleotides (concentration 500 nM), 2 µl of In-Fusion HD enzyme premix and the final volume was adjusted with sterile water, then the mixture was incubated at 50° for 15 minutes. Consequently, the In-Fusion reaction was transformed into *E. coli* XL10 gold ultracompetent cells by following the procedure as described in the section of cloning the homology region into the

pL6 eGFP plasmid. Plasmid extractions were performed and confirmed the insertion of sgRNA by Sanger sequencing.

### 2.3.5 Plasmid extraction by the alkaline extraction method

The plasmids after extraction by the alkaline extraction method from the transformants were screened for the insertion of HR1 and HR2 by digestion with the restriction enzyme *Bcl*I. To isolate the plasmid, 5 ml of bacterial culture was centrifuged at 3,000 rpm for 10 minutes, then 100 µl of Solution I (25 mM Tris-HCl pH 8.0, 10 mM EDTA, 50 mM glucose and 150 µg/ml RNase A) was added to the cell pellet and mixed by vortexing. After that, 150 µl of Solution II (0.2 M NaOH and 1% SDS) was added and mixed by inverting 6–8 times, then the sample was left at room temperature for 2 minutes followed by adding 150 µl of 3 M sodium acetate pH 5.2 and inverting 6–8 times. Then, it was centrifuged at 13,000 rpm for 10 minutes. The supernatant was transferred into a sterile 1.5 ml microcentrifuge tube and 1 volume of a mixture of phenol: chloroform: isoamyl alcohol (25:24:1) was added into the DNA sample, mixed by inverting, then centrifuged at 13,000 rpm for 10 minutes. The aqueous phase containing the isolated DNA was then transferred into a clean 1.5 ml microcentrifuge tube and the DNA was precipitated by adding 1/10 volume of 3M sodium acetate pH 5.2. Then, 2 volumes of absolute ethanol were added and it was incubated at −20°C for 30 minutes, followed by centrifugation at 13,000 rpm for 10 minutes, removal of the supernatant, and washing the DNA pellet with 70% ethanol. After that, the supernatant was gently removed and the DNA was dried. The DNA pellet was then dissolved in sterile water and the DNA concentration was measured by using a Nanodrop spectrophotometer. It was then stored at −20°C.

## 2.4 Parasite culture and transfection

### 2.4.1 Maintenance parasite culture

Asexual blood–stage *P. falciparum* parasites (strain 3D7) were cultured in RPMI1640 growth medium (with L-glutamine, HEPES and phenol red, Life Technologies) supplemented with 5 g/L Albumax II (Gibco), 2 g/L sodium bicarbonate (Sigma), 0.1 g/L hypoxanthine (Sigma) and 0.1% (v/v) gentamicin (10 mg/ml, Gibco) at 5% haematocrit. Human red blood cells ($O^+$ blood was

obtained from the National Blood Service of the National Health Service Blood and Transplant Unit (NHSBT) Seacroft, Leeds) was washed with RPMI medium by centrifugation at 3,000 rpm for 10 minutes, then the supernatant was discarded and repeated this step three times to remove the serum, dead cells, and white blood cells prior to use. The packed red blood cells were then mixed with an equal volume of RPMI medium to achieve a 50% haematocrit.

The parasite cultures were grown in 25 $cm^2$ polystyrene non-vented tissue culture flasks (SARSTEDT) and individually gassed with a 1% oxygen, 3% carbon dioxide, and 96% nitrogen gas mixture for 10 seconds prior to sealing the flasks. The culture flasks were then incubated horizontally in a 37°C incubator. The medium was changed daily by tilting the culture flask for 30–45 minutes to allow the red blood cells to settle on the bottom corner of the flask and then carefully removing the medium by a transfer pipette. After that, pre-warmed medium was added into the flask and the culture was then gassed and sealed before placing it back into the 37°C incubator.

The total number of parasites was monitored by blood smearing. While the culture flask was tilted, a sample of blood was collected by transfer pipette and smeared on a glass slide. The blood smear was then fixed with 100% methanol for 30 seconds and stained with 10% Giemsa staining solution (VWR Chemicals), diluted in Sorensen's buffer (10 mM $NaH_2PO_4$, 28 mM $Na_2HPO_4$, pH 7.2) for 10 minutes. The slide was rinsed with tap water and air dried before reading by Olympus BH-2 light microscopy with oil-immersion (1000× magnification).

The parasite counting was performed by using a Miller graticule, a device that assists in counting that is installed in the microscope eyepiece. The Miller graticule (Figure 2.3) consists of two squares, an outer square (in blue) and an inner square (in red). The inner square represents a tenth of the whole area of the graticule field (the outer square). Fields of distributed red blood cells were used for counting where the graticule was placed. The inner square was counted, then used to estimate the total number of red blood cells in the graticule field by multiplying by 10. The total number of infected red blood cells was also counted in both the inner and outer square for the parasitaemia calculation. The parasitaemia was calculated according to Equation 1 . It should be noted that in the case of multiple infections, the infected red blood cells were only counted as one and the stages of parasites were also noted. For daily parasitaemia monitoring, at least 1,000 red blood cells were counted.

$$Parasitaemia = \frac{total\ number\ of\ infected\ red\ blood\ cells}{estimated\ total\ number\ of\ red\ blood\ cells} \times 100\% \quad \textbf{Equation 1}$$

The parasite subculturing was performed when the parasitaemia exceeded 4% by diluting with 5% haematocrit in fresh complete medium to reduce the parasitaemia to 0.5 to 1%. The formula that was used for subculturing (a final volume of 6 ml per flask) is illustrated below:

$$V_i = 6\ ml \times (P_f/P_o)$$

$$V_{medium} = 0.9 \times (6\ ml - V_i)$$

$$V_{RBC} = 0.1 \times (6\ ml - V_i)$$

Where:

$P_o$ = initial parasitaemia

$P_f$ = final parasitaemia

$V_i$ = volume of sample from initial culture

$V_{medium}$ = volume of fresh medium

$V_{RBC}$ = volume of washed RBCs (50% haematocrit)

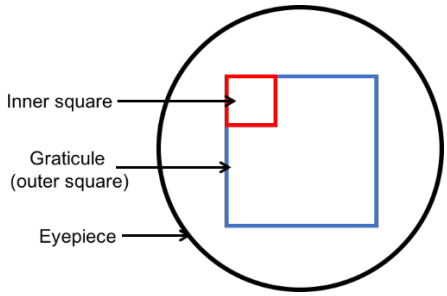



**Figure 2.3 An illustration of a Miller graticule**
The inner square (in red) represents a tenth of the whole area of the outer square (in blue). By counting red blood cells in the small square and then multiplying that number by 10, we can make an estimation of the total number of red blood cells in the whole graticule field.

### 2.4.2 Preparation of parasites for transfection

The early ring stage with a parasitaemia between 8–10% was used for transfection because the ring parasites tolerate electroporation better as compared with the other stages. A synchronised culture was prepared as described below by sorbitol treatment to get rid of any parasites at other stages than the ring stage. Prior to transfection, the parasites were cultured as described above until the parasitaemia reached 5–6% and the majority of the parasites were in the ring stage. The parasite culture (total volume 6 ml) was then transferred into a 50 ml Falcon tube and centrifuged at 3,000 rpm for 5 minutes. The medium was discarded and replaced with a 0.5× culture volume of filter-sterilised 5% sorbitol (3 ml of 5% sorbitol was added). In this step, the parasite pellet was gently mixed with a transfer pipette to ensure that the infected red cells were completely resuspended. The parasites were incubated in sorbitol for 10 minutes at room temperature and then a 2.5× culture volume (15 mL) of pre-warmed RPMI complete medium was added to dilute the concentration of sorbitol, followed by centrifugation at 3,000 rpm for 5 minutes. The supernatant was removed and replaced with 6 ml of pre-warmed complete medium, then the culture was transferred to a new culture flask, gassed and sealed before placing it back in an incubator. After synchronisation, the parasites were fed twice a day and the 24 hour synchronized parasites were assessed if the synchronisation was successful by blood smearing. The parasites were expected to be about 90–95% mid trophozoites. The early ring parasites had formed at 42 hours post-synchronisation. To obtain a highly synchronous ring stage parasites, a second synchronisation can be performed 40–42 hours later.

### 2.4.3 Plasmid preparation

Five plasmids, pL7-UMP-CMP kinase, pL7-DTC, pL7-KAHRP, pL7-KAHRP2, and pUF1-Cas9 were used in this study. The two plasmids pL7-KAHRP and pUF1-Cas9 were a gift from Prof. Jose-Juan Lopez-Rubio while the another KAHRP plasmid called pL7-KAHRP2 was obtained from Dr. Francis Isidore Garcia Totañes. The pL7-KAHRP2 plasmid was constructed in order to reduce the size of plasmid (1.1 kb smaller than the pL7-KAHRP). All plasmids were transformed into *E. coli* XL10 gold ultracompetent cells and kept in glycerol stock at −70°C. For plasmid preparation, the glycerol stocks were scraped and

cultured in 100 ml of LB broth supplemented with 100 µg/ml ampicillin with shaking at 200 rpm at 37°C for 16 hours. The bacterial cells were transferred into a 50 ml Falcon tube and harvested by centrifugation at 3,000 rpm, 4°C for 30 minutes. The plasmids were then isolated using a ZymoPURE™ Plasmid Midiprep Kit (catalogue number: D4200), following the manufacturer's protocol. According to the transfection protocol, 60 µg of circular plasmid or 10 µg of linearized plasmid are required for transfection of *P. falciparum* (Ghorbal *et al.* 2014). The circular plasmids were re-precipitated by adding 1/10 volume of 3M sodium acetate pH 5.2. Then, 0.7 volumes of isopropanol were added and it was incubated at −20°C for 30 minutes to precipitate the DNA. After centrifugation at 13,000 rpm for 10 minutes, followed by removing the supernatant gently, the pellet was washed with 70% ethanol and centrifuged again. The ethanol was removed without perturbing the pellet, then the DNA was dried. This step was conducted in a sterile cabinet to avoid any plasmid contamination. The DNA was dissolved in sterile water and stored at −20°C. The DNA concentration was measured by using a Nanodrop spectrophotometer.

For the linearization, the plasmids were digested with the restriction enzyme *HinC*II (2 cutting sites in the yFCU cassette). The 150 µl reaction included 3 µg of plasmid, 15 µl of 10× NEB buffer 3.1, 1.5 µl of restriction enzyme *HinC*II (10 U/µl) and adjusted to the final volume with sterile water. The reaction was then incubated at 37°C for at least 2 hours. A total of 7 linearized reactions from each plasmid was performed in order to obtain sufficient DNA for the transfection experiment. After digestion, the linearized DNA was purified by adding 1 volume of the phenol-chloroform mixture solution and subsequently centrifuging at 13,000 rpm for 10 minutes. The top layer was transferred into a sterile 1.5 ml microcentrifuge tube and the DNA was then re-precipitated as described above for circular plasmid precipitation. The linearization was verified on agarose gel electrophoresis. The DNA was dissolved in sterile water at an appropriate concentration and stored at −20°C.

### 2.4.4 Transfection

For the co-transfection of the pUF1-Cas9 plasmid and the pL7-GOI linear plasmids, the transfection procedure followed the previously described protocol. The two plasmids were simultaneously transfected into the

synchronized early ring stage parasites (around 8% parasitaemia). First, the synchronized culture was packed by centrifugation at 3,000 rpm for 5 minutes, then suspended in a 7.5 packed cell volume of ice-cold cytomix buffer. After that, 0.8 ml of the suspended infected cells were transferred into the 4 mm electroporation cuvette and kept on ice. Then 60 µg of pUF1-Cas9 circular plasmid and 10 µg of the linear plasmid were added into the cuvette (the maximum volume added that contained the two plasmids was 30 µl), mixed gently and electroporated using the conditions: 0.62 kV, 25 µF and 200 Ω resistance in a BioRad Gene Pulser II. After incubation of the electroporated cells on ice for 2 minutes, the cells were then transferred into a culture flask containing 6 ml of complete medium and 5% freshly washed red blood cells. The transfected parasite flask was gassed before culturing in the 37°C incubator. Drug selection was applied with 1.5 nM WR99210 and 300 µM DSM265 24 hours after transfection and maintained for 5 days. The media cultures were changed daily and the parasites were continuously monitored until the appearance of transgenic parasites was detected by Giemsa-stained thin blood smear.

In the co-transfection for the circular plasmids, the protocol obtained from the Pasteur Institute was used in this experiment. The overall protocol for introducing CRISPR-Cas9 plasmids into the parasites and verifying the presence of transgenic parasites is summarised in Figure 2.4. The transfection should be completed within 30–45 minutes. Therefore, the procedure was followed step by step as shown below:

1) Plasmids

Added 60 µg of pUF-Cas9 and 60 µg of each CRISPR (pL7) plasmid, and 270 µL of cytomix buffer (120 mM KCl, 0.15 mM $CaCl_2$, 2 mM EGTA, 5 mM $MgCl_2$, 25 mM HEPES, 10 mM $K_2HPO_4$, 10 mM $KH_2PO_4$, pH 7.6) into a sterile 1.5 ml microcentrifuge tube and incubated the DNA solution at 37°C. The mixture was conducted in the safety cabinet.

2) Culture flasks

Prepared a 6 ml pre-warmed complete RPMI medium mixed with 100 µl of fresh RBC in a 25 $cm^2$ tissue culture flask and placed it in the 37°C incubator.

3) Parasites

The parasites in the early ring stage with parasitaemia between 8–10% were pelleted using centrifugation at 3,000 rpm for 5 minutes and the supernatant was gently removed by a transfer pipette. Then the pellet was kept in the cabinet prior to transfection.

4) Electroporation

The DNA/cytomix buffer in the sterile 1.5 ml microcentrifuge tube was removed from the incubator and added with 100 µl of infected RBC pellet, mixed well, and transferred this mixture into the 2 mm electroporation cuvette (Geneflow, catalogue number: E6-0060). This step was done carefully to avoid any bubbles. The total volume in a 2 mm cuvette was 400 µl. The parasites were then electroporated using the conditions 0.31 kV, 25 µF and 200 Ω resistance using a BioRad Gene Pulser II. After the pulse, the time constant was recorded in milliseconds. Then, the cells were transferred into the pre-prepared culture flask that contained complete medium and fresh RBCs. The cuvette was rinsed with medium to collect any remaining cells, then the parasite flask was gassed and placed back in the 37°C incubator. Three hours after the transfection, the medium was changed to remove the cytomix and replaced it with the pre-warmed fresh medium. Drug selection with 1.5 nM WR99210 and 300 µM DSM265 was applied 6 hours post-transfection. The next day, the medium was changed and adjusted the haematocrit to 5%. The culture flasks were mixed twice a day to increase the invasion rate. For the next 5–8 days, the culture flasks had their media changed daily and the parasitaemia was monitored on a daily basis until no live parasites were observed in the blood smears. The drug selection was performed for only 5 days post-transfection. Afterwards, the medium was changed every 2 days, but gassing of the cultures was performed every day. Seven days post-transfection 3 mL of parasite culture was collected for genomic DNA extraction and replaced it with 3 mL of 5% haematocrit complete RPMI medium into the culture flasks. Fresh RBCs were added to the cultures every 2 weeks and continued to monitor the haematocrit. The transgenic parasites were continuously monitored for 60 days post-transfection with a routine blood smear performed every week.

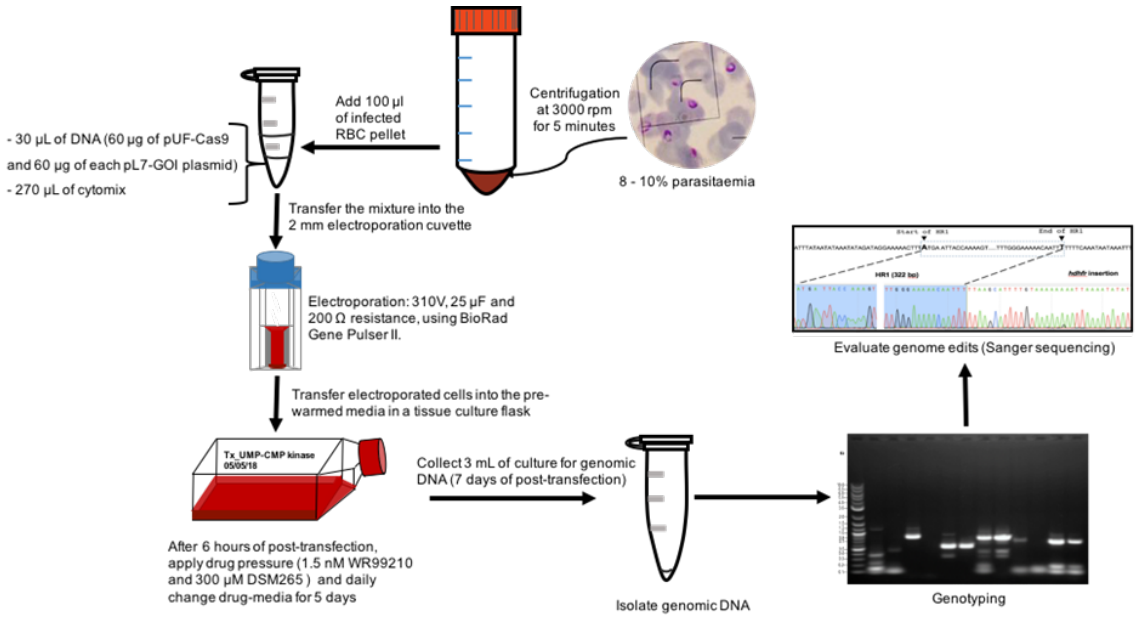



**Figure 2.4 The overall protocol of CRISPR/Cas9 system**
The synchronized ring-stage parasites were transfected with the CRISPR/Cas9 plasmids, then the transfected parasites were cultured under drug pressure for 5 days to select the parasites that contained the two plasmids. After that, genomic DNA isolation was performed (7 days of post-transfection) using the standard phenol-chloroform extraction method, then the integrated parasites were evaluated through genotyping and confirmed the gene modification by Sanger sequencing.

## 2.5 Analysis of transgenic parasites

The genomic DNAs were isolated from 3 ml of parasite cultures using the protocol for gDNA extraction as described in Section 2.3.1. These genomic DNAs were used for genotyping to evaluate the CRISPR success (whether the donor DNA (*hdhfr*) was incorporated into the target locus). Based on the CRISPR-Cas9 system, after the site specific DNA was cleaved by Cas9 endonuclease, the cleaved target locus was then repaired by homology recombination through the using of HR1 and HR2 flanking *hdhfr*, resulting in *hdhfr* to be incorporated into the genome. Therefore, the primers for the diagnostic PCR of the 5' and the 3' integrations were generated by designing the outer primers as being only genomic DNA specific without including any of the HR sequence to avoid amplifying the plasmid. The inner primers were donor DNA (*hdfhr*) specific. A diagram of the primer design is depicted in Figure 2.5. Melting temperatures of the primers around 54–59°C were chosen

(more specific binding was expected from higher melting temperature primers). Ideally, the diagnostic primers are expected to amplify only the integrated mutant-specific sequences whereas the wild-type parasites and the episomal mutant parasites are expected to produce no product.

For genotyping, the PCR reactions were performed in a 20 µl mixture reaction, containing 10–20 ng of gDNA, 500 nM of forward primer, 500 nM of reverse primer, 10 µl of Q5® High-Fidelity 2X Master Mix and the final volume was adjusted with sterile water. The PCR reaction was performed for 40 cycles; the first cycle consisted of the initial denaturation at 98°C for 30 seconds and the subsequent 40 cycles consisted of 98°C for 10 seconds, the annealing temperature was adjusted (depending on the pair of primers in the reaction) and was performed for 30 seconds, then the extension temperature was held at 62°C for 2 minutes, followed by 62°C for 2 minutes and the temperature was then held at 4°C after final step. The PCR products were monitored by agarose gel electrophoresis and visualized under a UV transilluminator.

To confirm the *hdhfr* cassette was integrated into the target locus, the obtained PCR products at the expected sizes were cloned into bacterial cells using the TOPO® TA Cloning® Kit (Invitrogen) for subsequent sequencing after gel purification of the PCR product. The TOPO® TA cloning® kit requires 3' A-overhangs that can be added by the Taq polymerase. Therefore, the Q5-PCR product in this experiment had to be purified by using a gel extraction kit (QIAquick, catalogue no.28704) to remove the Q5®-High Fidelity DNA polymerase and to get rid of any non-specific bands. For the 3' A-overhangs addition, a 20 µl reaction containing 10 µl of purified PCR product and 10 µl of Go Taq Green 2x Master Mix was incubated at 72°C for 15 minutes using a PCR machine. After that, a TOPO® cloning reaction was performed in a 6 µl reaction, contained 4 µl of PCR product (after adding the 3' A-overhangs), 1 µl of salt solution (provided from the TOPO® TA Cloning® Kit) and 1 µl of TOPO® vector, and the mixture solution was incubated for 5 minutes at room temperature. Subsequently, the TOPO® cloning reaction was transformed into *E. coli* XL10 gold ultracompetent cells by adding 2 µl of the TOPO®cloning reaction to 50 µl of competent cells, then incubated on ice for 30 minutes followed by application of a heat-pulse in a 42°C water bath for 30 seconds, followed by 2 minutes of incubation on ice. After that, 250 µl of LB broth was added into the transformation reaction and it was incubated at 37°C with shaking at 200 rpm for 1 hour. The transformants were plated on LB agar

supplemented with 100 µg/ml ampicillin and incubated overnight at 37°C. Then, plasmid extractions were performed and confirmed the insertion of the *hdhfr* cassette by Sanger sequencing. Due to the multistep of sub-cloning through the TOPO® TA Cloning® Kit, some of the obtained PCR products at the expected sizes were directly subjected to DNA sequencing after gel purification.

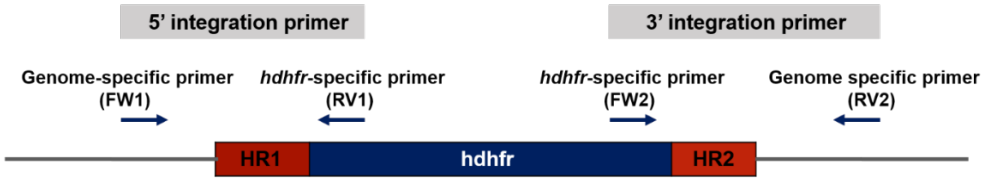



**Figure 2.5 An illustration of the primers designed for genotyping**
The outer primers are genomic DNA specific and the inner primers are *hdhfr* specific. These diagnostic primers were used to amplify only the integrated mutants.

# Chapter 3 Result

## 3.1 Cloning of the CRISPR plasmid and insertion of the guide RNA sequences targeting the UMP-CMP kinase

To investigate the gene function of the UMP-CMP kinase, the plasmid pL7-UMP-CMP kinase that contains sgRNA and the donor DNA template was generated from the plasmid pL6 eGFP, then this plasmid was used for co-transfection with the unmodified pUF1-Cas9 plasmid to knockout the targeted gene in *P. falciparum*. Construction of the pL7-UIMP-CMP kinase plasmid was started by cloning of HR1 into the GFP 5' region of the pL6 eGFP plasmid. The genomic DNA was extracted from the *P. falciparum* strain 3D7 by the standard phenol-chloroform extraction method and used as a template for the amplification of HR1. The fragment of HR1-UMP-CMP kinase was successfully amplified with a single band at the expected size of 322 bp as shown in Figure 3.1 A. Then, a large-scale PCR reaction was performed and purified by a gel extraction kit before cloning this fragment into the plasmid pL6 eGFP. After the purification, the PCR product was monitored on an agarose gel (Figure 3.1 B), compared with the reference size 336 bp of the non-SERCA-type $Ca2^{+}$-transporting P-ATPase gene.

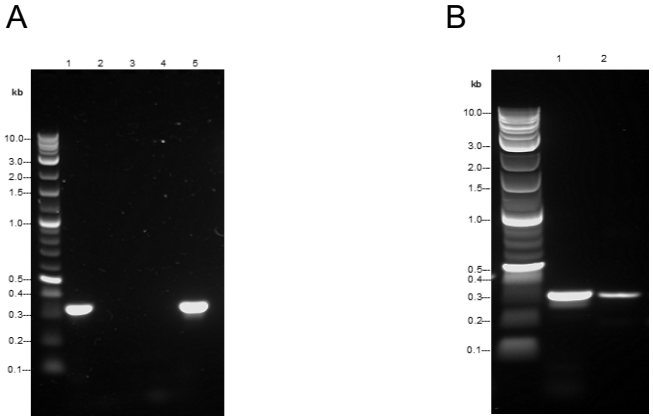



**Figure 3.1 Amplification of HR1-UMP-CMP kinase**
(A) The PCR product of HR1- UMP-CMP kinase on 2% agarose gel electrophoresis, Lane 1: PCR product of HR1- UMP-CMP kinase using gDNA as a template (expected size 322 bp), Lane 2: negative control PCR reaction without the HR1- UMP-CMP forward primer, Lane 3: negative control PCR reaction without the HR1- UMP-CMP reverse primer, Lane 4: negative control PCR reaction without the gDNA template, Lane 5: positive control PCR reaction using the primers for amplification of non-SERCA-type $Ca2^{+}$ -transporting P-ATPase gene (expected size 336 bp). (B) HR1- UMP-CMP

kinase on 2% agarose gel electrophoresis after gel purification, Lane 1: the purified fragment of HR1 (322 bp), Lane 2: the reference size (336 bp) of the non-SERCA-type $Ca2^{+}$- transporting P-ATPase gene.

For the cloning of the HR1- UMP-CMP kinase into the plasmid, the plasmid pL6 eGFP was digested with *Spe*I and *Afl*II to remove the GFP 5', the site where the purified fragment of HR1 was introduced into the plasmid. The expected bands of 9496 bp and 347 bp were obtained after digestion, indicating that the plasmid was successfully digested by the restriction enzymes (Figure 3.2). After that, a 1:10 ratio of digested plasmid and HR1-UMP-CMP kinase fragment were used for the In-Fusion cloning, then they were transformed into *E. coli* XL10 gold ultracompetent cells. Consequently, the transformants were plated on LB agar supplemented with 100 µg/ml ampicillin. Five clones of the transformants were picked up for plasmid extraction and the positive clones were then screened by digestion with *Bcl*I (since the restriction site of this enzyme was added into the forward primer for the positive clone screening). Due to the absence of this restriction site in the template plasmid, after enzyme digestion the positive clones would show a shifted band when compared with the undigested plasmid (the migration of the linearized plasmid is slower than the circular plasmid) as shown in Figure 3.3 after digestion with *Bcl*I. It was difficult to see the shifted band after the digestion among the five selected clones, however, four of them (clones HR1-U2, HR1-U3, HR1-U4 and HR1-U5) presented a slightly shifted band. Therefore, those plasmids were then subjected to DNA sequencing to confirm the presence of the HR1-UMP-CMP kinase and the correctness of its sequence. It was noted that for further experiments, screening for the positive clone by enzyme digestion will be skipped because of the indistinguishable digested bands.

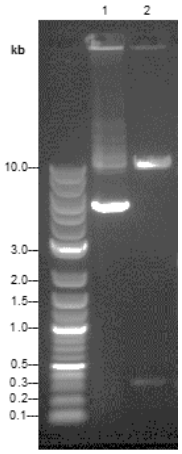



**Figure 3.2 Plasmid pL6-eGFP as a backbone for cloning of HR1-UMP-CMP kinase**
The pL6 eGFP was digested with *Spe*I and *Afl*II and analysed on 0.6% agarose gel electrophoresis, then it was used as the backbone for cloning of HR1. Lane 1: undigested plasmid pL6-eGFP, Lane 2: digested plasmid pL6-eGFP with *Spe*I and *Afl*II (expected bands after digestion are 9496 bp and 347 bp).

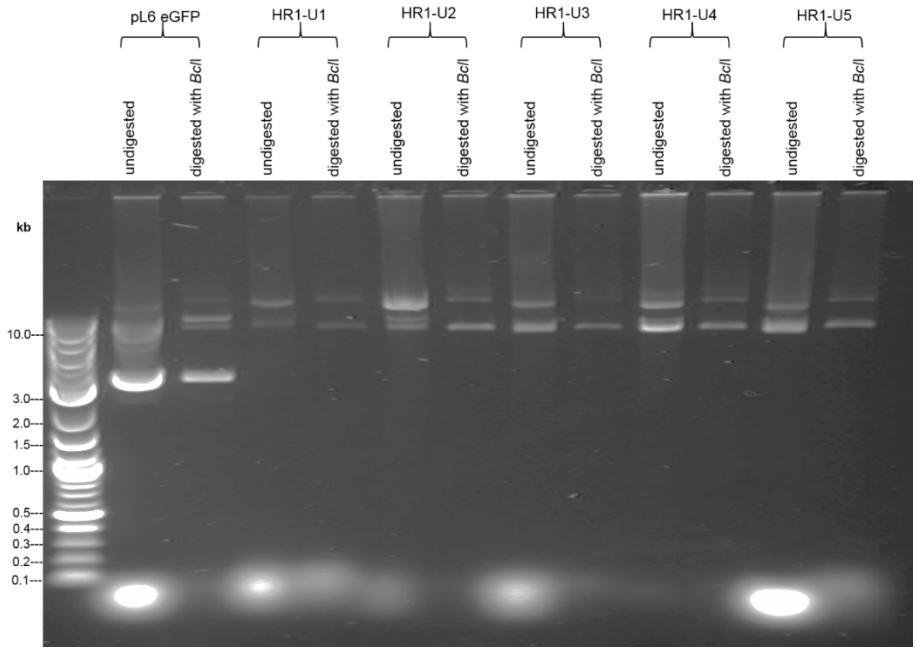



**Figure 3.3 Positive clone screening of HR1-UMP-CMP kinase**
Five plasmids were extracted from bacterial cells using the alkaline extraction method and those plasmids were then digested with *Bcl*I to check the insertion of HR1-UMP-CMP kinase. The positive clone was expected to have a shifted band when comparing digested parental plasmid with clones.

From the DNA sequencing results, the DNA sequence from the selected clones was analysed by multiple sequence alignment compared with the

sequence of UMP-CMP kinase extracted from PlasmoDB (Gene ID: PF3D7_0111500). The alignment result is illustrated in Figure 3.4, where all four selected clones had the DNA sequence of HR1-UMP-CMP kinase as expected. The correct clone HR1-U5 (named pL6 HR1-U5) with a good chromatogram result was chosen for the further cloning of the HR2 of UMP-CMP kinase.

The next step of cloning was the insertion of the HR2 fragment into the plasmid pL6 HR1-U5. The amplification of the HR2-UMP-CMP kinase followed the same procedures as used for the HR1-UMP-CMP kinase. However, the expected PCR product was obtained with non-specific bands (Figure 3.5 A). To amplify a single PCR product, the PCR conditions were optimised by changing the extension temperature from 62°C to 64°C and 65°C. As a property of Q5® High-Fidelity DNA Polymerase, the annealing temperature must be followed as calculated by the NEB Tm Calculator (https://tmcalculator.neb.com/#!/main). Hence, only the extension temperature can be adjusted. After increasing the extension temperature, a single band of the PCR product was achieved from both temperatures (Figure 3.5 B); however, the extension temperature of 64°C showed a greater yield when compared to 65°C. This condition was therefore applied in the large-scale PCR. Consequently, the obtained HR2 fragment from the large-scale reaction was purified by gel purification and monitored on agarose gel electrophoresis as depicted in Figure 3.6, lane 3. The fragment of HR2-UMP-CMP kinase was then cloned into the plasmid pL6 HR1-U5.

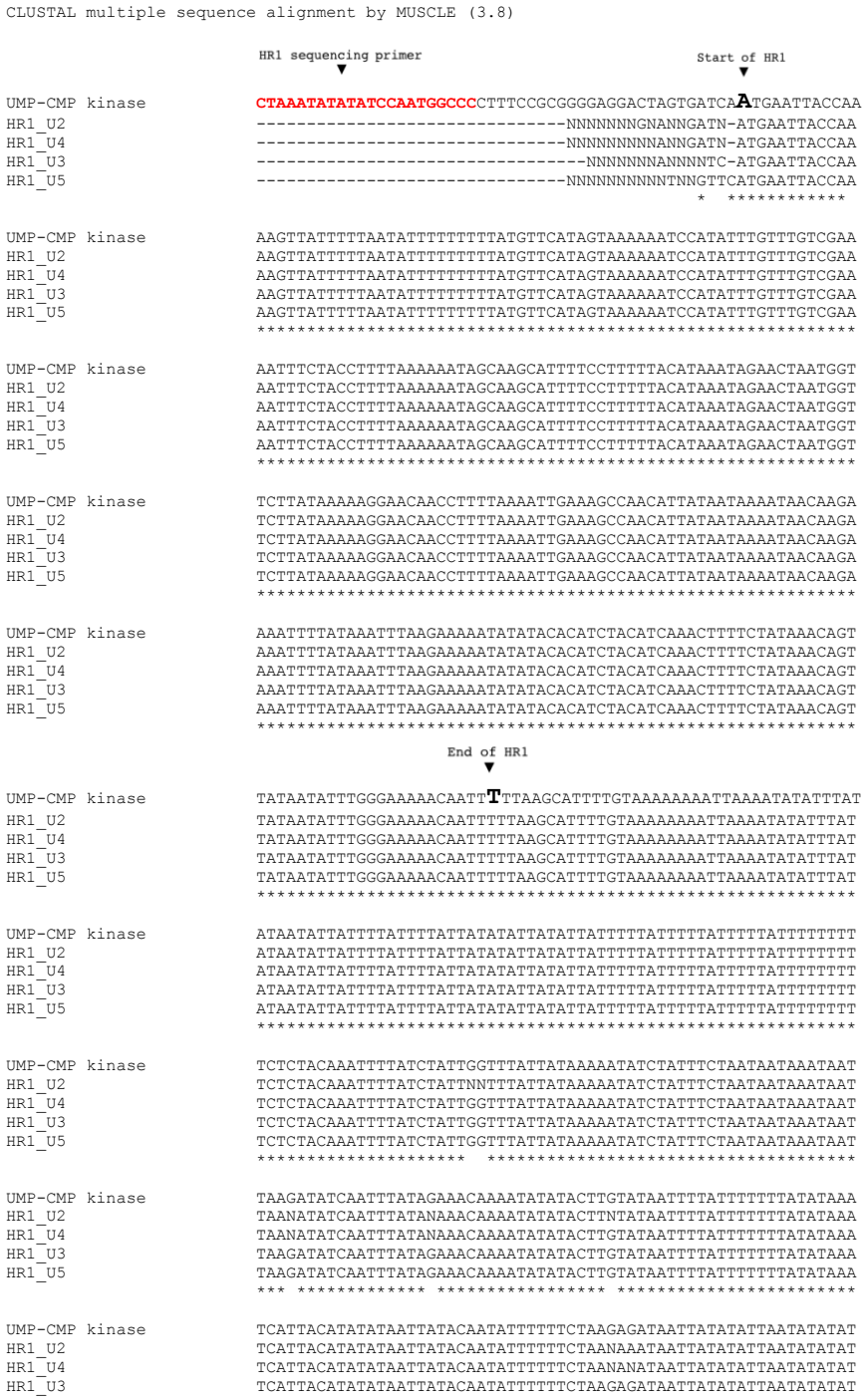



**Figure 3.4 Sequence alignment of HR1-UMP-CMP kinase**
DNA sequencing of the selected clones were compared with the UMP-CMP kinase sequence from PlasmoDB (Gene ID: PF3D7_0111500) using Clustal multiple sequence alignment. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.

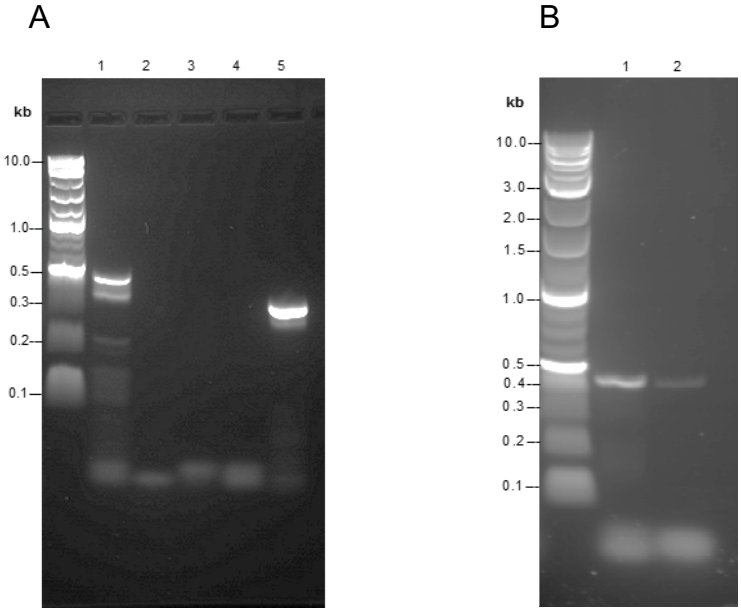



**Figure 3.5 Amplification of HR2-UMP-CMP kinase**
(A) PCR product of HR2-UMP-CMP kinase on 2% agarose gel electrophoresis, using the extension temperature at 62°C, Lane 1: PCR product of HR2- UMP-CMP kinase using gDNA as a template (expected size 453 bp), Lane 2: negative control PCR reaction without the HR2- UMP-CMP forward primer, Lane 3: negative control PCR reaction without the HR2- UMP-CMP reverse primer, Lane 4: negative control PCR reaction without gDNA, Lane 5: positive control PCR reaction using the primers for non-SERCA-type $Ca2^{+}$-transporting P-ATPase gene (expected size 336 bp). (B) PCR product of HR2 of UMP-CMP kinase on 2% agarose gel electrophoresis after adjusting the extension temperature, Lane 1: PCR product of HR2-UMP-CMP kinase using an extension temperature at 64°C, Lane 2: PCR product of HR2- UMP-CMP kinase using an extension temperature at 65°C.

For the backbone preparation, the plasmid pL6 HR1-U5 was digested with *EcoR*I and *Nco*I to remove the original GFP 3' and replace it with the fragment of the HR2-UMP-CMP kinase. The expected fragments of the backbone plasmid were obtained after digestion, but undigested plasmids were also observed (Figure 3.6, lane 2). The cloning of the HR2-UMP-CMP kinase fragment into the digested pL6 HR1-U5 plasmid was performed by using the In-Fusion cloning kit. Thirteen clones from the transformants were picked up for plasmid extraction and sent for Sanger sequencing. Four of them, clones HR2-U9, HR2-U11, HR2-U12, and HR2-U13, had the sequence of the HR2-UMP-CMP kinase, but the deletion of one adenine base was found in clone HR2-U12. The multiple sequence alignment of the sequencing clones was compared with the sequence of the UMP-CMP kinase from the database as

shown in Figure 3.7. From the sequencing data (shown in the appendix), clone HR2-U9 (named pL6 HR2-U9) had the expected sequence with a good chromatogram result and was selected for the further experiment of adding sgRNA.

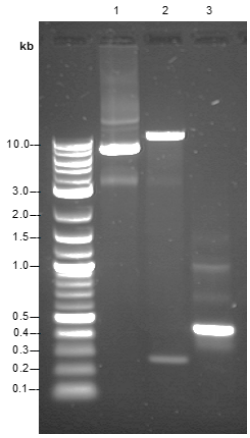



**Figure 3.6 Plasmid pL6 HR1-U5 as a backbone for cloning of HR2- UMP-CMP kinase**
The plasmid pL6 HR1-U5, carrying the fragment of HR1-UMP-CMP kinase was digested with restriction enzymes *EcoR*I and *Nco*I and then the digested plasmid was used as the backbone plasmid for cloning of the HR2-UMP-CMP kinase fragment. The DNA fragments were analysed on 0.6% agarose gel electrophoresis. Lane 1: undigested plasmid pL6 HR1-U5, Lane 2: digested plasmid pL6 HR1-U5 with *EcoR*I and *Nco*I (expected bands were 9509 bp and 273 bp), Lane 3: HR2 fragment (453 bp) of UMP CMP kinase after gel extraction.

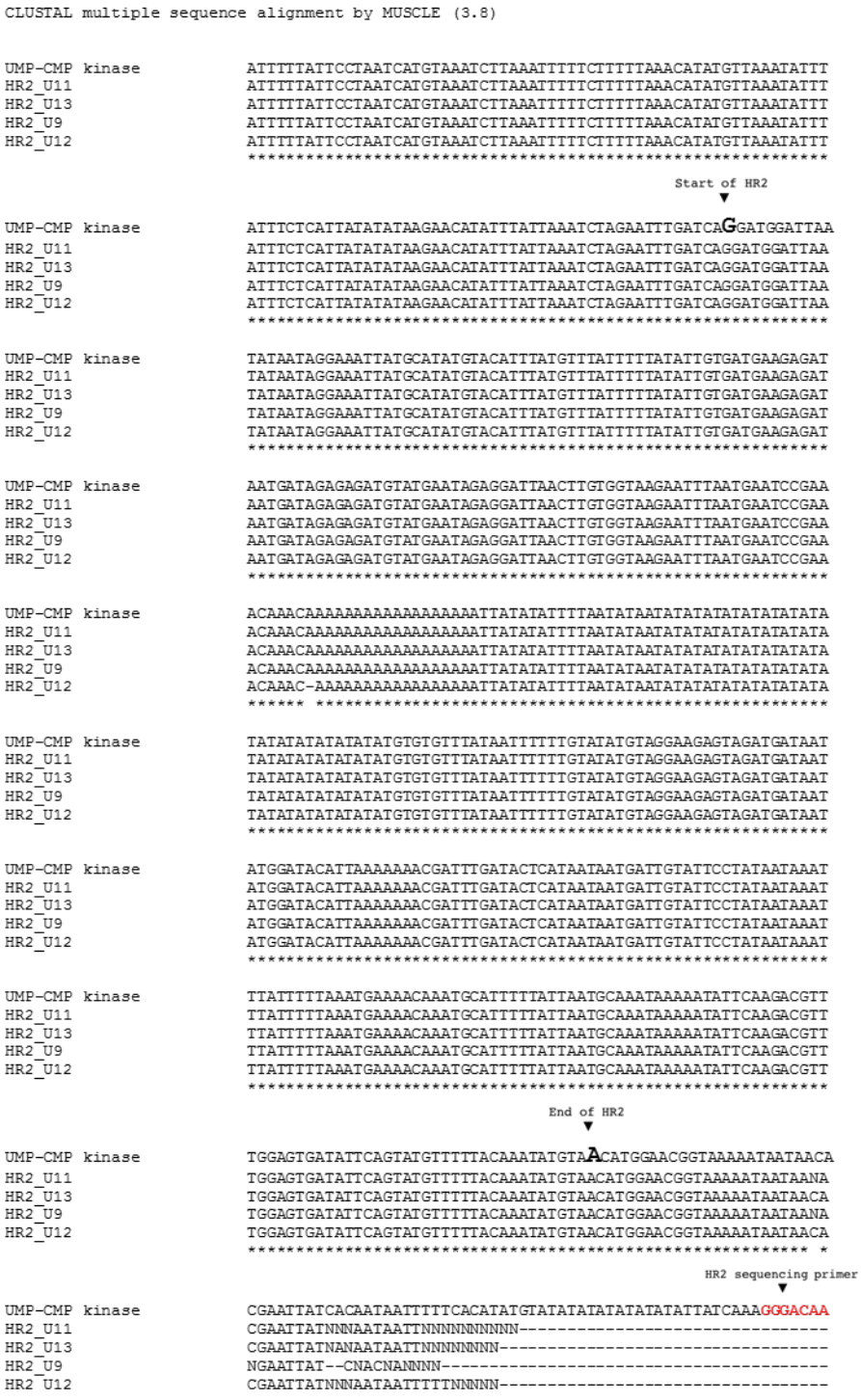



**Figure 3.7 Sequence alignment of HR2-UMP-CMP kinase**
DNA sequencing of the selected clones were compared with the UMP-CMP kinase sequence from PlasmoDB (Gene ID: PF3D7_0111500) using Clustal multiple sequence alignment. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.

In the final step of cloning the plasmid pL7–UMP-CMP kinase, the oligonucleotides duplex of sgRNA (targeting the UMP-CMP kinase) was

inserted into the pL6 HR2-U9 plasmid (the CRISPR plasmid carrying the two fragments of HR1 and HR2 UMP-CMP kinase) by replacing the *Btg*ZI adaptor. The backbone plasmid pL6 HR2-U9 after digested with the restriction enzyme is shown in Figure 3.8. The results show that the backbone plasmid was successfully digested into a single fragment, then the annealed oligonucleotides of sgRNA were inserted into the digested plasmid through the use of an In-Fusion cloning kit. After cloning and transformation of the In-Fusion reaction into bacterial cells, four clones were selected from the transformants to confirm the insertion of sgRNA. Three of them showed the presence of sgRNA with the correct sequence inserted into the plasmid at the expected site. The DNA analysis by multiple sequence alignment is demonstrated in Figure 3.9.

In summary, the pL7-UMP-CMP kinase carrying the sgRNA and the donor DNA template (*hdhfr*) flanked by two homology regions (HR1 and HR2) was successfully constructed and confirmed by Sanger sequencing. This plasmid was then co-transfected along with the plasmid pUF1-Cas9 into the malaria parasites *P. falciparum* to knockout and evaluate the biological function of UMP-CMP kinase.

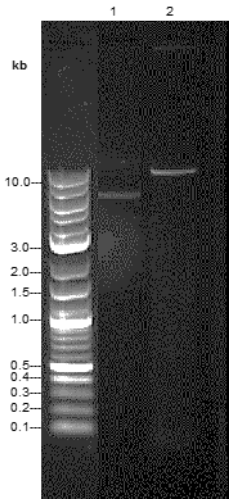



**Figure 3.8 Plasmid pL6 HR2-U9 as a backbone for adding sgRNA of UMP-CMP kinase**
The pL6 HR2-U9 plasmid, carrying the two fragments of HR1 and HR2 UMP-CMP kinase was digested with the restriction enzyme *Btg*ZI and then the digested plasmid was used as the backbone plasmid for insertion of sgRNA. The DNA fragments were analysed on 0.6% agarose gel electrophoresis. Lane 1: undigested plasmid pL6 HR2-U9, Lane 2: digested plasmid pL6 HR2-U9 with *Btg*ZI (expected band, 9926 bp).


**Figure 3.9 Sequence alignment of sgRNA-UMP-CMP kinase**
DNA sequencing of three clones were compared with the predicted sequence of sgRNA-UMP-CMP kinase using Clustal multiple sequence alignment. The predicted sequence was generated by using APE plasmid editor. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.

## 3.2 Cloning of the CRISPR plasmid and insertion of guide RNA sequences targeting the dicarboxylate/tricarboxylate carrier (DTC)

The fragment of HR1 from the DTC kinase was amplified from the genomic DNA and cloned into the template plasmid (pL6 eGFP). The fragment of HR1 was successfully amplified with a single band as an expected size of 413 bp (Figure 3.10 A). Then, a large-scale PCR reaction was performed and cleaned up using a gel purification kit. After that, the purified fragment was cloned into the digested pL6 eGFP plasmid. The preparation of the backbone plasmid

(pL6 eGFP) by digestion with the restriction enzymes *Spe*I and *Afl*II to remove GFP 5' was performed in parallel. The fragments of plasmids after digestion and the purified PCR product of HR1 are illustrated in Figure 3.11 lane 1 and lane 2, respectively. The results show that the digested fragments were obtained as the expected sizes but undigested plasmids were also found, indicating that the reaction showed incomplete digestion. However, this digested backbone plasmid was subsequently used in the In-Fusion cloning reaction for cloning the fragment of HR1 into the plasmid. After cloning and transformation, three clones were selected and subjected to DNA sequencing to confirm the presence of HR1-DTC and the correctness of its sequence. From DNA analysis by multiple sequence alignment compared with the sequence of DTC from the database (Figure 3.12), two clones (HR1-DTC3 and HR1-DTC4) that had the expected sequence of HR1-DTC (the DNA sequencing data is shown in the appendix) were identified. Then the clone HR1-DTC4 (named pL6 HR1-DTC4) was selected as the template for cloning HR2-DTC.

Cloning of the HR2-DTC fragment into the plasmid pL6 HR1-DTC4, the fragment of HR2-DTC that was amplified from genomic DNA with the expected size (436 bp) was obtained as illustrated in Figure 3.10 B. In the cloning step, the template plasmid pL6 HR1-DTC4 was also prepared in parallel by digestion with the restriction enzymes *EcoR*I and *Nco*I for removing the GFP 3' and replacing it with the HR2-DTC fragment. The expected bands of the digested plasmid and the purified fragment of HR2-DTC were verified on agarose gel electrophoresis (Figure 3.11, Lane 4 and lane 5). For the plasmid digestion, although the fragments with the expected sizes were obtained, incomplete digestion was also observed. After that, a 1:10 ratio of digested plasmid and insert fragment was applied in the In-Fusion cloning, followed by transformation into bacterial cells. A total of 10 clones from the transformants were selected for plasmid extraction, then the plasmids were sent for DNA sequencing to confirm successful cloning. From the sequencing results, only clone HR2-DTC7 (named pL6 HR2-DTC7) had the correct sequence of HR2-DTC when aligned with the DTC sequence, which was extracted from PlasmoDB (Figure 3.13). It should be noted that three of ten clones showed the sequence of the GFP 3' (the backbone plasmid sequence) while the other clones showed failure of DNA sequencing with multiple peaks. In the final step

of cloning, the plasmid pL6 HR2-DTC7 was used as the template for insertion of the sgRNA.

The backbone plasmid (pL6 HR2-DTC7) for adding the sgRNA was prepared through its digestion with the restriction enzyme *BtgZ*I. The DNA fragment after digestion is illustrated in Figure 3.11, lane 7. The results showed successful digestion with a single band when compared with the undigested plasmid. Then, the annealed oligonucleotides of sgRNA (targeting DTC) were inserted into the digested plasmid using the In-Fusion cloning kit and the obtained plasmids from the In-Fusion reaction were transformed into *E. coli* XL10 gold ultracompetent cells. Eleven clones were subjected to DNA sequencing, and two of them, clone sgDTC7 and sgDTC8, were confirmed to have the oligonucleotides sgRNA sequence inserted into the plasmid (the DNA sequencing data is shown in the appendix) whereas the other clones failed to read the DNA sequencing. DNA analysis by multiple sequence alignment was performed as depicted in Figure 3.14, and both clones had the expected sequence of sgRNA targeting DTC and the sgRNA was inserted at the corresponding site as predicted.

In summary, the plasmid pL7-DTC was successfully constructed and confirmed to be correct by Sanger sequencing. The gene function of DTC for parasite survival was then further investigated by co-transfection of the plasmid pL7-DTC and pUF1-Cas9 into the ring-stage parasites of *P. falciparum* to knock out the target gene, then evaluated its function.

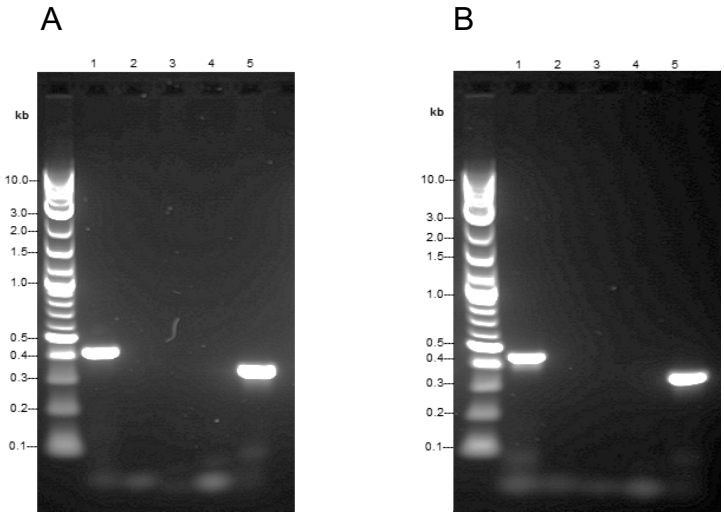



**Figure 3.10 Amplification of HR1-DTC and HR2-DTC**
(A) PCR product of HR1-DTC on 1.5% agarose gel electrophoresis, Lane 1: PCR product of HR1-DTC using gDNA as a template (expected size 413 bp), Lane 2: negative control PCR reaction without the HR1-DTC forward primer, Lane 3: negative control PCR reaction without the HR1-DTC reverse primer, Lane 4: negative control PCR reaction without gDNA, Lane 5: positive control PCR reaction using the primers for amplification of non-SERCA-type $Ca2^{+}$ -transporting P-ATPase gene (expected size 336 bp). (B) PCR product of HR2-DTC on 1.5% agarose gel electrophoresis, Lane 1: PCR product of HR2- DTC using gDNA as a template (expected size 436 bp), Lane 2: negative control PCR reaction without the HR2-DTC forward primer, Lane 3: negative control PCR reaction without the HR2-DTC reverse primer, Lane 4: negative control PCR reaction without gDNA, Lane 5: positive control PCR reaction using the primers for amplification of non-SERCA-type $Ca2^{+}$-transporting P-ATPase gene (expected size 336 bp).

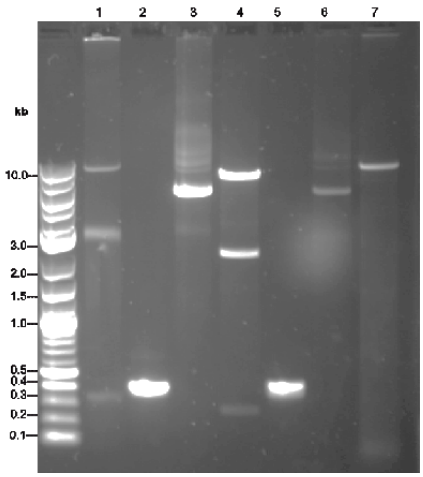



**Figure 3.11 The backbone plasmids and the insert fragments HR1 and HR2 of DTC**

The plasmid pL6-eGFP was digested with *Spe*I and *Afl*II, then it was used as the backbone for cloning of HR1 while the plasmid pL6 HR1-DTC4 was used as the template for cloning of HR2 after digested with *EcoR*I and *Nco*I. The DNA fragments were analysed on 0.6% agarose gel electrophoresis. Lane 1: digested plasmid pL6-eGFP with *Spe*I and *Afl*II (expected bands, 9496 bp and 347 bp), Lane 2: HR1 homology fragment (413 bp) of DTC after gel extraction, Lane 3: undigested plasmid pL6 HR1-DTC4, Lane 4: digested plasmid pL6 HR1-DTC4 with *EcoR*I and *Nco*I (expected bands, 7274 bp, 2326 bp and 273 bp), Lane 5: HR2 homology fragment (436 bp) of DTC after gel extraction, Lane 6: undigested plasmid pL6 HR2-DTC7, Lane 7: digested plasmid pL6 HR2-DTC7 with *BtgZ*I (expected band, 9999 bp).

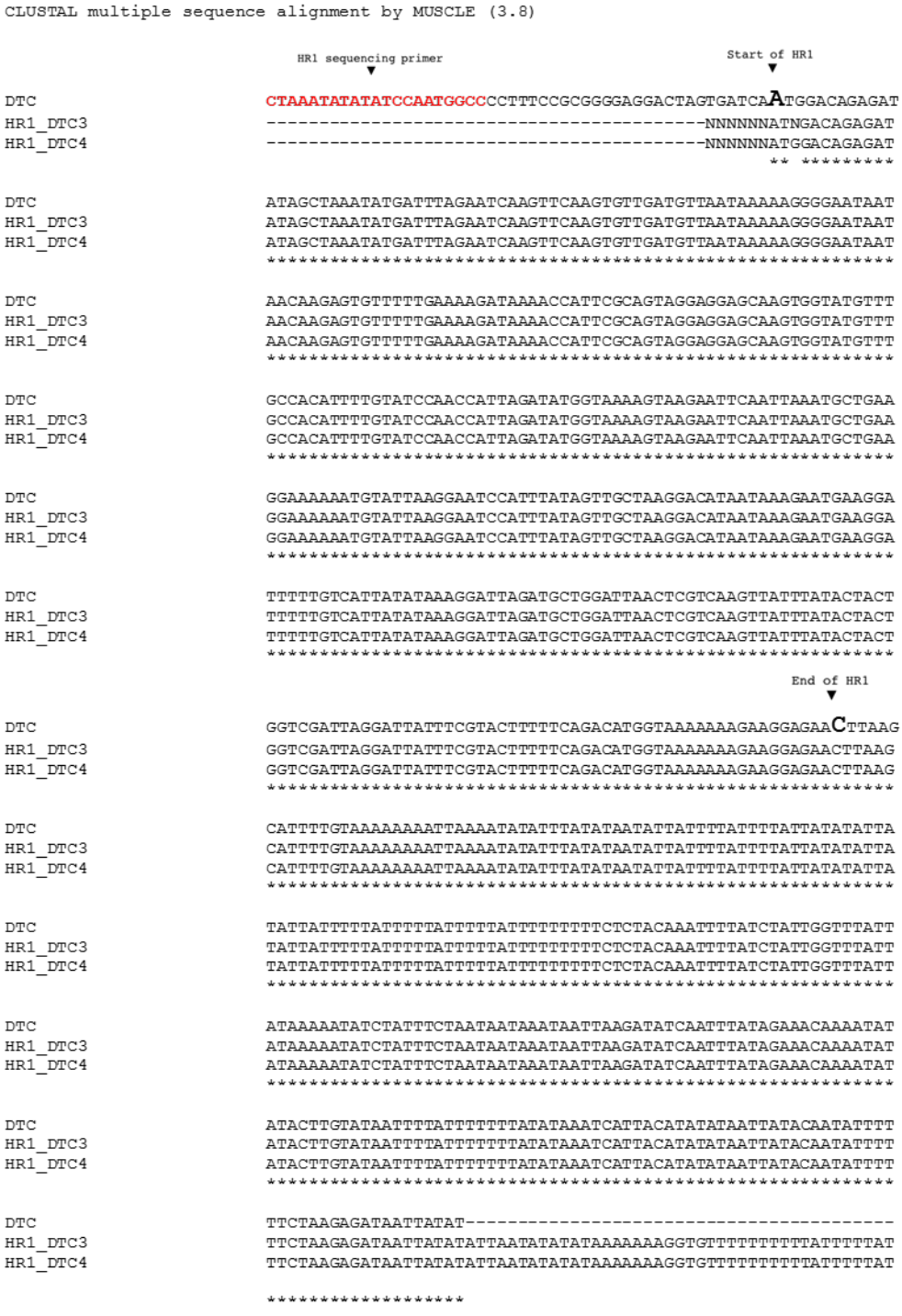



**Figure 3.12 Sequence alignment of HR1- DTC**

DNA sequencing of two clones were compared with the DTC sequence from PlasmoDB (Gene ID: PF3D7_0823900) using Clustal multiple sequence alignment. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.

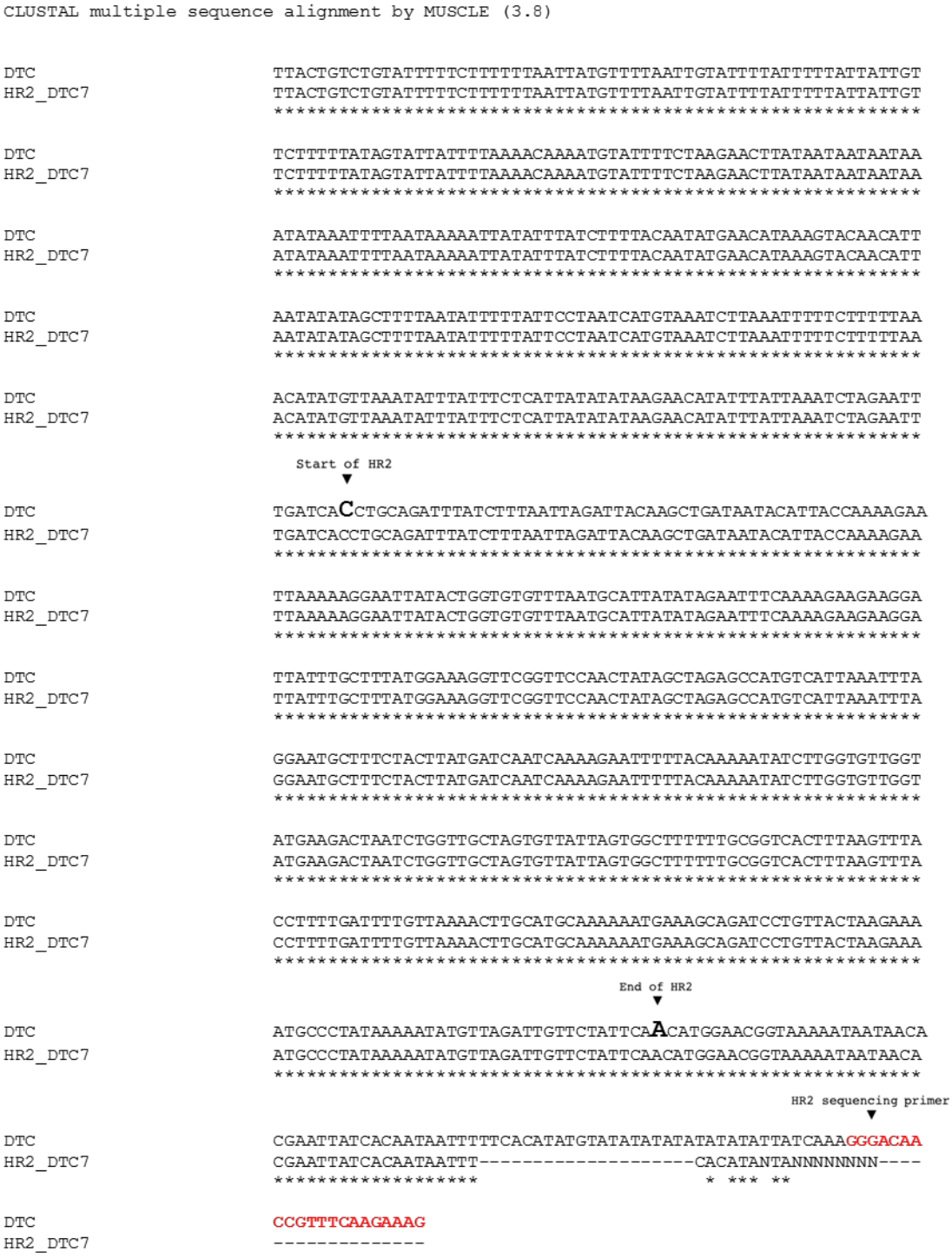



**Figure 3.13 Sequence alignment of HR2-DTC**
DNA sequencing of the selected clone was compared with the DTC sequence from PlasmoDB (Gene ID: PF3D7_0823900) using Clustal multiple sequence alignment. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.


**Figure 3.14 Sequence alignment of the sgRNA-DTC**
DNA sequencing of two clones were compared with the predicted sequence of sgRNA-DTC using Clustal multiple sequence alignment. The predicted sequence was generated by using APE plasmid editor. The red bold sequence represents the sequence and the position of the primer used for DNA sequencing.

## 3.3 Preparation of the linear plasmids for transfection

According to the transfection technique reported by Ghorbal et al. (2014), the integrant parasites using a linear plasmid can be detected at the same time as for the circular plasmid. However, the advantage of using a linear plasmid is that the negative selection does not require the removal of the episomal parasites, which is a time-saving approach. In this study, three plasmids (pL7-UMP-CMP kinase, pL7-DTC, pL7-KAHRP) were digested with the restriction enzyme *Hinc*II to produce the linear plasmids and the digested plasmids were then verified on agarose gel electrophoresis. The results showed that the linearization produced the expected bands but undigested plasmids were also found (Figure 3.15). Then, 10 µg of linear plasmid was used for co-transfection along with 60 µg of pUF1-Cas9 circular plasmid (plasmid size 11096 bp shown in Figure 3.15, lane 1).

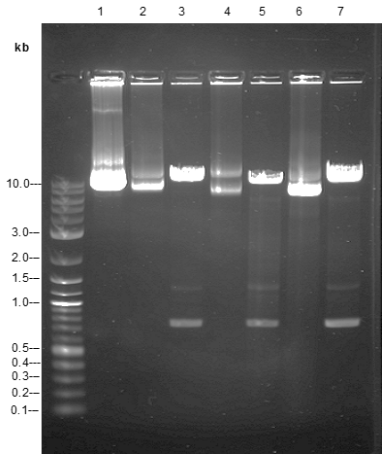



**Figure 3.15 The circular and linearized plasmids for transfection**
For the co-transfection of pUF1-Cas9 plasmid and linear pL7-GOI plasmid, all plasmids after extraction and linearization by cutting with the restriction enzyme *HinC*II (double cutting in the yFCU cassette) were analysed on 0.8% agarose gel electrophoresis. Lane 1: pUF1-Cas9 circular plasmid (11096 bp), Lane 2: pL7-KAHRP circular plasmid, Lane 3: pL7-KAHRP linearized plasmid, Lane 4: pL7-UMP-CMP kinase circular plasmid, Lane 5: pL7-UMP-CMP kinase linearized plasmid (expected bands: 9168 bp and 747 bp), Lane 6: pL7-DTC circular plasmid and Lane 7: pL7-DTC linearized plasmid (expected bands: 9241 bp and 747 bp).

## 3.4 Generation of transgenic parasites

In this study, four plasmids, pL7-UMP-CMP kinase, pL7-DTC, pL7-KAHRP, and pL7-KAHRP2 were used for establishing the CRISPR-Cas9 knockout system. The transfection experiment was performed in four individual transfection experiments. The first experiment was conducted for the co-transfection of the pUF1-Cas9 plasmid and the linear plasmid of pL7-GOIs (UMP-CMP kinase and DTC) with the non-essential KAHRP gene (pL7-KAHRP), obtained from the Pasteur Institute, as a positive control for the CRISPR-Cas9 system. The other three experiments were performed using the pUF1-Cas9 plasmid co-transfected with the circular plasmids of pL7-GOIs and used pL7-KAHRP2 (1.1 kb smaller than the pL7-KAHRP) as a positive control for the CRISPR-Cas9 system. In each transfection, 60 µg of pUF1-Cas9 and 60 µg of pL7-GOI circular plasmid/10 µg of pL7-GOI linearized plasmid were mixed in the cytomix buffer and then the pelleted infected RBC with 8–10% parasitaemia at the early ring stage parasites was added, mixed gently, and then the electroporation was performed. After transfection, the parasites were cultured in 6 ml of complete RPMI medium (0.5% albumax) in 25 $cm^2$ tissue culture flasks. Subsequently, 6 hours after transfection the inhibitors 1.5 nM

WR99210 and 300 μM DSM265 were applied to the culture flasks for selection of the parasites that only contained both plasmids. The drug pressure was applied for 5 days post-electroporation. The transfected parasites were then monitored by Giemsa-stained thin blood smear. On day 6 of post-transfection, no parasites were detected in the blood smears. The parasite cultures were continuously monitored for 60 days post-transfection and a routine blood smear was performed every week. From the thin blood smears analysis, after 2 months of transfection, no viable parasites were detected from the transgenic parasites with the UMP-CMP kinase and DTC, whereas for the non-essential gene, KAHRP2, only a small number of ring parasites and unhealthy trophozoite parasites were observed on day 20 post-transfection (as depicted in Figure 3.16). Ring stage parasites are considered a reliable sign of a healthy parasite culture amongst malariologists. To rescue the unhealthy parasites, the complete RPMI medium, supplemented with 10% human serum, was used instead of albumax; however, the results showed the transgenic parasites failed to recover and there was no progress in the parasite development after 2 months.

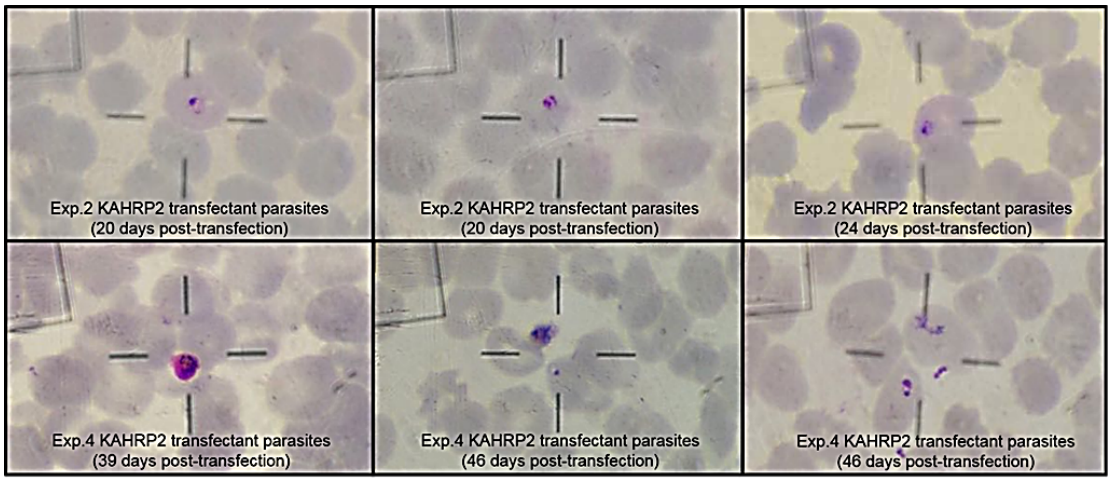



**Figure 3.16 Images of transgenic parasites of KAHRP2**
The viable parasites were detected from two batches of the transgenic parasites of KAHRP2 using Giemsa-stained thin blood smear.

## 3.5 Analysis of the transgenic parasites

### 3.5.1 Primers for genotyping

After the parasites were transfected with the two plasmids, the site-specific double strand DNA was broken by the Cas9 nuclease, and consequently, the cleaved target locus was repaired by homology recombination (using HR1 and HR2 flanking *hdhfr*), allowing *hdhfr* to be incorporated into the genome (Figure 3.17 A). According to the four individual transfection experiments, the genotyping was performed for screening the integrant parasites after 7 days of transfection. In each experiment, genomic DNA was isolated from 3 ml of parasite cultures, then used for PCR analysis. Primers were designed to amplify the 5' and 3' integrations (the left and right homology arms) for evaluation of the donor DNA that was integrated into the target loci of each transgenic parasite as showed in Figure 3.17. Three sets of primers (KAHRP_GT_FW1/ GT_RV1, GT_FW2/ KAHRP_GT_RV2, and GT_FW2/ KAHRP_GT_RVb) were generated for evaluation of the KAHRP and KAHRP2 mutant parasites. By using 5' integration primers (KAHRP_GT_FW1/ GT_RV1) and the 3' integration primers (GT_FW2/ KAHRP_GT_RV2), the expected size of the products of the KAHRP mutant were 1324 bp and 1043 bp for 5' integration and 3' integration, respectively. The positions and directions of the primers, including the expected product sizes, are illustrated in Figure 3.17 B. The same primer sets for the KAHRP mutant were also used for evaluation of the KAHRP2 mutant but were used to amplify the different PCR products (as shown in Figure 3.17 C). The expected bands from the KAHRP2 mutant are 1205 bp and 380 bp for the 5' integration and 3' integration, respectively. In addition, a new set of primers (GT_FW2 and KAHRP_GT_RVb) was designed for assessment of the 3' integration of the KAHRP2 mutant with a product size of 789 bp. It is a fact that the reverse primer of KAHRP_GT_RV2 is able to bind to the HR2 of the pL7-KAHRP2 plasmid (except for the first 3 bases on the 5' end). Ideally, the primer should not bind to any HR sequence to avoid any plasmid amplification. Therefore, a new reverse primer KAHRP_GT_RVb was created, allowing the primer to only bind to the genomic sequence. The schematic for the primer design is demonstrated in Figure 3.17 C.

To evaluate the integrated parasites of the UMP-CMP kinase, 3 sets of primers were created with the positions, directions of the primers, and the expected

size of the PCR products as presented in Figure 3.17 D. Two sets of primers (UMP-CMP_GT_FW1/GT_RV1, and U3/ P1) were used for diagnostic PCR of the 5' integration with products of 466 bp and 524 bp, respectively. Due to the AT-rich genomic sequence, the set of primers U3/P1 was designed to have a melting temperature at 59°C, while another set of primers (UMP-CMP_GT_FW1/GT_RV1) was designed to have a melting temperature around 54°C. It was reasoned that the higher melting temperature could provide greater specificity of target-binding. For the amplification of the 3' integration of the mutant parasite, primers GT_FW2/UMP-CMP_GT_RV2 were used for diagnosis, generating a product size of 748 bp.

In the assessment of DTC integrated parasites, the primer sets D3/P1 and P2/D4 were designed for verification of the 5' and 3' integration in the mutant parasite with the expected product sizes of 688 bp and 650 bp, respectively (Figure 3.17 E). These sets of primers were generated to have melting temperatures of 59°C. A summary of the primers and their sequences for genotyping are shown in the appendix.

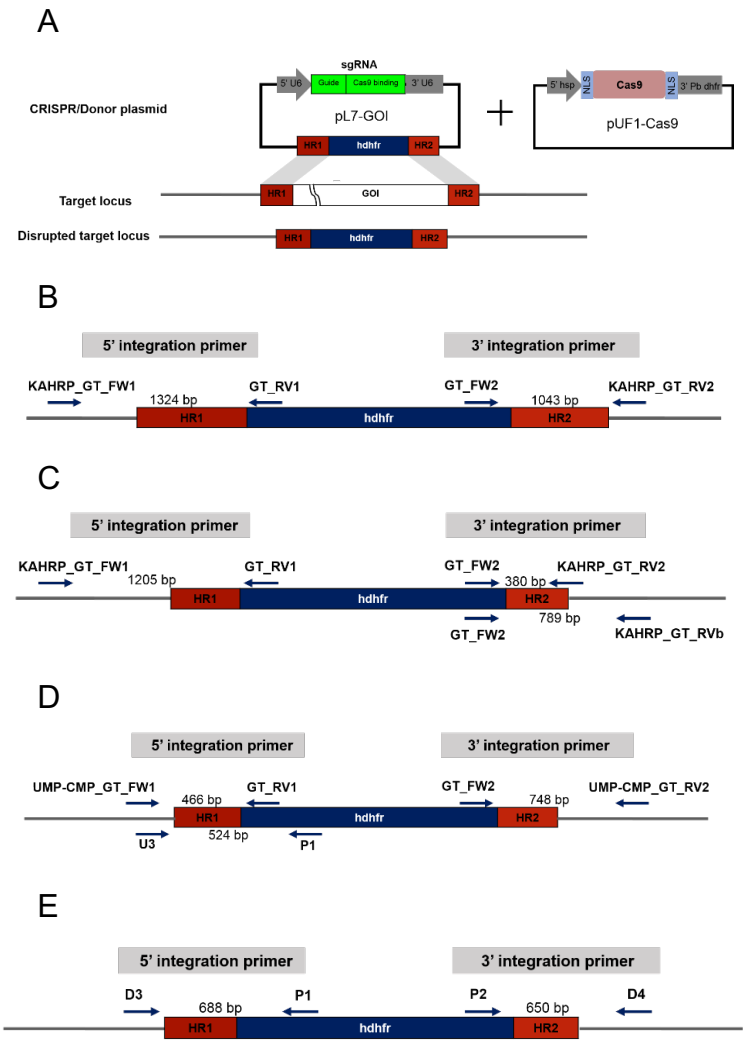



**Figure 3.17 CRISPR-Cas9 system and an illustration of diagnostic primers used for genotyping**

(A) Schematic of the CRISPR-Cas9 system for gene disruption in *P. falciparum*. Two plasmids are required for transfection into the *P. falciparum* infected erythrocytes. The pL7-GOI contains sgRNA and the donor DNA template (*hdhfr*) is flanked with two homology regions for homology repair after a double- strand break. The site-specific double strand DNA is broken by the Cas9 protein, which is expressed from the plasmid pUF1-Cas9. The subsequent homology directed repair uses the sequence surrounding the cutting site from the donor DNA template, allowing the *hdhr* gene to integrate into the target locus. (B) Genotyping primers for verification of the integrant parasites of KAHRP (positive control taken from a published paper). (C) Genotyping primers for verification of the integrant parasites of KAHRP2. (D) Genotyping primers for verification of the integrant parasites of the UMP-CMP kinase. (E) Genotyping primers for verification of the integrant parasites of DTC. The outer primers are genomic DNA specific and the inner primers are *hdhfr* gene specific. The positions and directions of the primers are indicated by the small blue arrows and the expected size of the PCR products are presented between the blue arrows.

### 3.5.2 Genotyping

Genomic DNA was isolated from the transgenic parasites of four individual transfection experiments and 10–20 ng of gDNA was then used for the analysis of gene disruption by genotyping. The first batch of transfections was the co-transfection of the pUF1-Cas9 plasmid and the linear plasmid of pL7-GOIs (UMP-CMP kinase and DTC) with pL7-KAHRP as the control for the CRISPR-Cas9 system. The PCR products from the genotyping were checked by agarose gel electrophoresis and were visualized under a UV transilluminator. As the results show in Figure 3.18, among 3 transgenic parasites, only 5' and 3' integrations from the UMP-CMP kinase and 3' integration from DTC were detected as PCR products with the expected sizes of 542 bp, 748 bp and 650 bp, consecutively; however, non-specific bands were also observed from the PCR product while the control KAHRP parasites had a negative genotyping result.

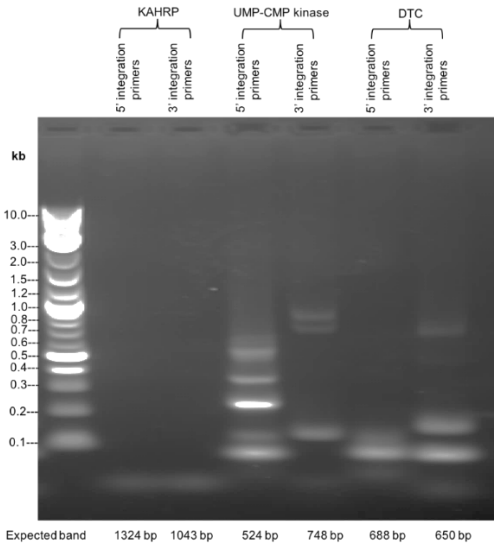



**Figure 3.18 Diagnostic PCR detection from the transfection experiment I**
For the first transfection experiment, the linear CRISPR plasmids were co-transfected with pUF1 Cas9 plasmid. After 7 days of culturing, the parasite DNA was isolated and evaluated by genotyping, the PCR products were then analysed on 1.5% agarose gel electrophoresis. The primers were designed for detection of the integrated donor DNA template (*hdhfr* gene) in the 5' and 3' integrations from transgenic parasites and the expected size of each PCR products is shown below the figure.

For the second transfection experiment, co-transfection of the pUF1-Cas9 plasmid and the circular plasmid of pL7-GOIs (UMP-CMP kinase and DTC) were performed, using pL7-KAHRP2 as a positive control for the CRISPR-Cas9 system. In Figure 3.19, the integration of *hdhfr* in the target loci in 5' and 3' integrations were detected as the expected bands (1205 bp and 380 bp) from the KAHRP2 transgenic parasites but non-specific bands were observed from the use of diagnostic primers for the 5' integration (KAHRP_GT_FW1/ GT_RV1). Unexpectedly, the non-specific PCR products were also found in the wild-type genomic DNA sample. The DNA from the band with the expected size of 5' integration in the genotyping test was subjected to Sanger sequencing to confirm successful integration (as the result was shown in Figure 3.22 A) whereas the 3' integration it should be noted that the diagnostic primer of KAHRP_GT_RV2 for detection of the 3' integration of KAHRP2 is able to bind to the HR2 of the pL7-KAHRP2 plasmid. Hence, the new set of primers (GT_FW2/ KAHRP_GT_RVb) as described in 3.5.1 Primers for genotyping will be used for diagnosis of the 3' integration of KAHRP2 trangenic parasites.

The diagnostic PCR of integrated mutant from UMP-CMP kinase clearly shows a PCR product of the expected size (748bp) was obtained for the 3' integrations whereas the 5' integrations showed a product size that was slightly higher than the expected size that might cause by the use of non-specific diagnostic primers (Figure 3.19). DNA sequencing was performed to confirm the incorporation of the *hdhfr* gene in the 3' integration (the result showed in Figure 3.23 B). In addition, the unexpected result of PCR products were obtained from wild-type genomic DNA samples by using both pairs of the diagnostic primers.

From the genotyping result of the DTC transgenic parasites, the incorporation of *hdhfr* in the 3' integration was obtained as a PCR product with the expected size (650 bp) while the 5' integration presented a negative genotyping result. Nevertheless, the positive genotyping of the integrated mutant in the 3' integration from the DTC transgenic parasites was confirmed by Sanger sequencing as the result was depicted in Figure 3.24 B (the multiple sequence alignment was presented in the appendix, Figure A7 clone 3_DTC1 and clone 3_DTC2). The 5' integration of the DTC was not re-checked by diagnostic PCR due to running out of the gDNA template.

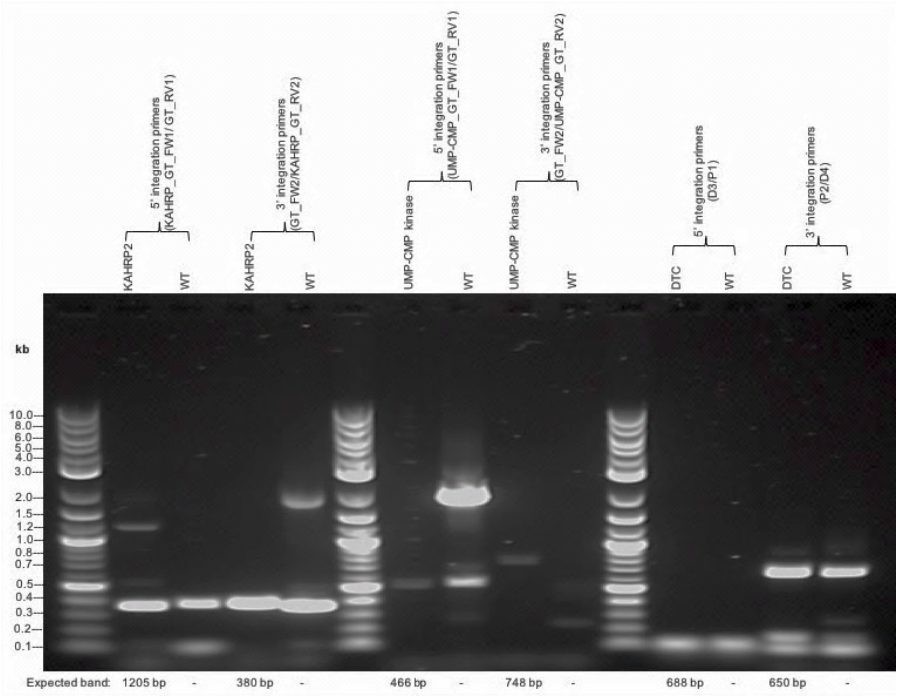



**Figure 3.19 Diagnostic PCR detection from the transfection experiment II**

For the second batch of transfection experiment, the circular CRISPR plasmids were co-transfected with pUF1 Cas9 plasmid. After 7 days of culturing, the parasite DNA was isolated and evaluated by genotyping, the PCR products were then analysed on 1.5% agarose gel electrophoresis. Genotyping was performed by using diagnostic primer that targeted integration of the *hdhfr* cassette in 5' and 3' integrations from the KAHRP2, UMP-CMP kinase and DTC transgenic parasites. The plasmid pL7-KAHRP2 was used as a positive control for the CRISPR system. The expected size of each PCR products is shown below the figure.

Genotyping results from the third transfection experiment are depicted in Figure 3.20. Despite the presence of non-specific bands including in the wild-type gDNA sample, the incorporation of the *hdhfr* gene in the 5' integration and the 3' integration for all transgenic parasites were detected at the expected sizes with the exception of the 5' integration of the UMP-CMP kinase. The 5' integration of the UMP-CMP kinase showed a PCR product that was larger than the expected size.

Therefore, the diagnostic PCR was repeated by using a new set of primers (U3/P1) designed to have an annealing temperature of 59°C. These results showed a PCR product at the expected size 524 bp as shown in Figure 3.20 B, The incorporation of the *hdhfr* gene in the 5' integration of the UMP-CMP kinase transgenic parasites was confirmed by DNA sequencing ( as the result

was demonstrated in Figure 3.23 A). It should be noted that the diagnostic primers for detection of the 3' integration of KAHRP2 in this genotyping were GT_FW2/KAHRP_GT_RVb and generated a product size of 789 bp as expected size, then the integration was verified by Sanger sequencing as shown in Figure 3.22 B. For the transgenic parasites of DTC showed a positive genotyping result with 688 bp and 650 bp as expected bands for the 5' integration and the 3' integration, respectively. The positive genotyping products of DTC were confirmed the integration mutants by DNA sequencing as the results were depicted in Figure 3.24 A and B for 5' integration and 3' integration, respectively.

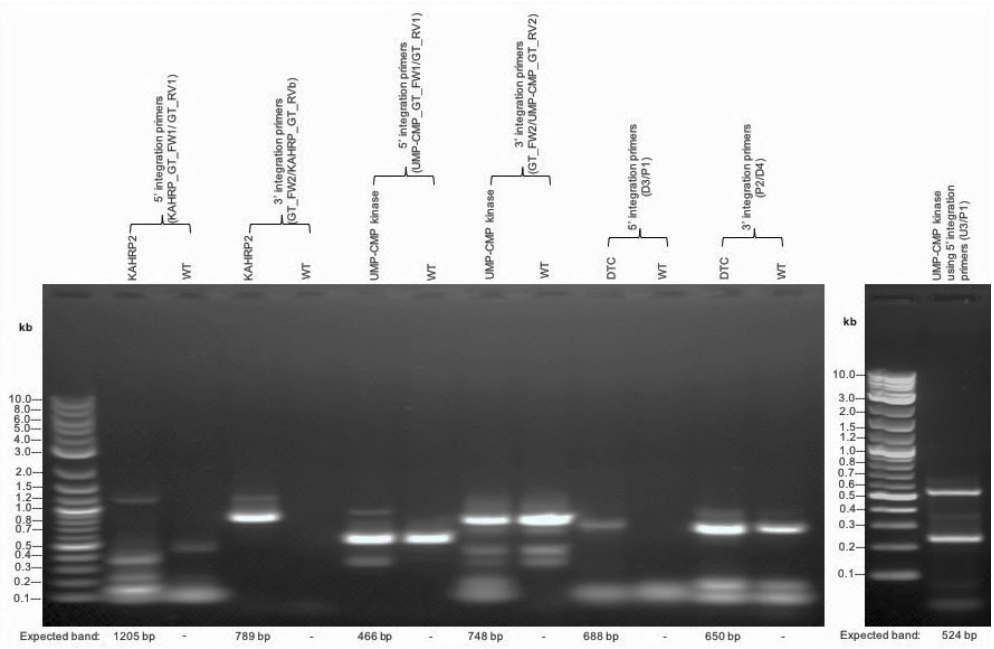



**Figure 3.20 Diagnostic PCR detection from the transfection experiment III**

For the third batch of transfection experiment, the circular CRISPR plasmids were co-transfected with pUF1 Cas9 plasmid. After 7 days of culturing, the parasite DNA was isolated and evaluated by genotyping, the PCR products were then analysed on 1.5% agarose gel electrophoresis. (A) Genotyping was performed by using diagnostic primer that targeted integration of the *hdhfr* cassette in 5' and 3' integrations from the KAHRP2, UMP-CMP kinase and DTC transgenic parasites. The plasmid pL7-KAHRP2 was used as a positive control for the CRISPR system. (B) Re-diagnostic PCR product from 5' integrations of UMP-CMP kinase. The expected sizes of genotyping are shown below the figure.

For the genotyping of the last transfection experiment, the results are illustrated in Figure 3.21. It can be seen that the targeted integration of the *hdhfr* cassette in the 5' integration and 3' integration for all transgenic parasites were detected at the expected sizes (as these positive genotyping results from the previous transfections were confirmed via DNA sequencing) although non-specific binding of primers was observed.

In summary, the genotyping results demonstrated a successful diagnostic PCR product. In addition, the obtained PCR products at the expected sizes were cloned into bacterial cells using the TOPO® TA cloning® kit or purified directly from the agarose gels and the samples were then subjected to Sanger sequencing to confirm the integration of the *hdhfr* gene in the target loci.

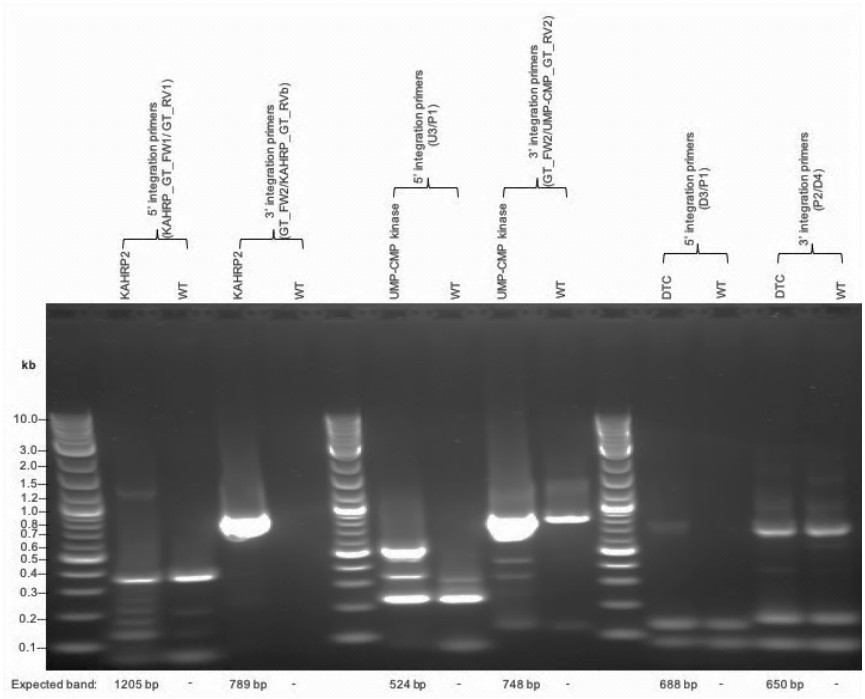



**Figure 3.21 Diagnostic PCR detection from the transfection experiment IV**

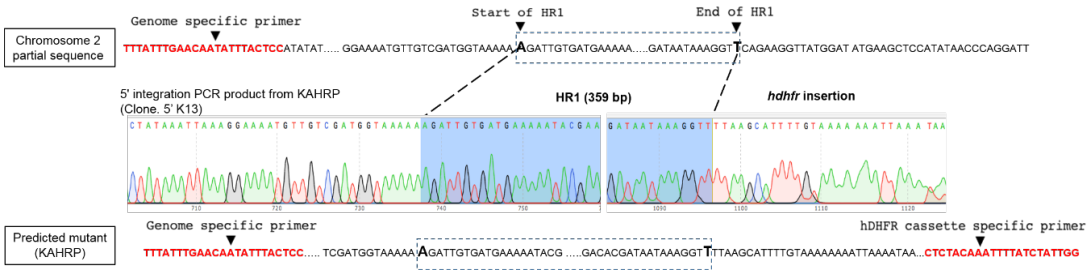
For the fourth batch of transfection experiment, the circular CRISPR plasmids were co-transfected with pUF1 Cas9 plasmid. After 7 days of culturing, the parasite DNA was isolated and evaluated by genotyping, the PCR products were then analysed on 1.5% agarose gel electrophoresis. Genotyping was performed by using diagnostic primer that targeted integration of the *hdhfr* cassette in 5' and 3' integrations from the KAHRP2, UMP-CMP kinase and DTC transgenic parasites. The plasmid pL7-KAHRP2 was used as a positive control for the CRISPR system. The expected sizes are shown below the figure.

### 3.5.3 Sequencing confirmed the successful integration of the *hdhfr* cassette into the target loci.

The diagnostic PCR products, amplified from the gDNA of each transgenic parasite, were subjected to Sanger sequencing to confirm the incorporation of the *hdhfr* cassette in the 5' and 3' integration (left and right homology arms) at specific sites guided by sgRNA. The sequencing result of the obtained PCR products from the KAHRP2 transgenic parasites is illustrated in Figure 3.22. The diagnostic PCR product of the 5' integration (referred to as clone 5'K13) shows the presence of the *hdhfr* cassette that was sequenced after the HR1 (highlighted in the blue dashed box) as the predicted mutant sequence (Figure 3.22 A). Moreover, 2 clones (3' K4 and 3' K6) obtained from the PCR analysis of the 3' integration showed the sequence of the *hdhfr* cassette was inserted before the sequence of HR2 (as depicted in Figure 3.22 B). The DNA analysis by multiple sequence alignment of DNA sequencing compared with the wild-type sequence and the predicted mutant sequence are shown in the appendix (Figure A2 and A3).

For the verification of the integrated mutant UMP-CMP kinase and DTC, DNA sequencing confirmed successful integration of donor DNA into the targeted loci in the left and right homology arms as shown in Figure 3.23 and Figure 3.24 for the UMP-CMP kinase and DTC, respectively. The sequence surrounding the homology region (between HR1 and HR2) presents the sequence of the *hdhfr* cassette as matched with the sequence of the predicted mutant. In addition, the DNA analysis by multiple sequence alignment confirmed the insertion of the *hdhfr* cassette into the targeted loci as shown in the appendix (Figure A4-A7).

A)

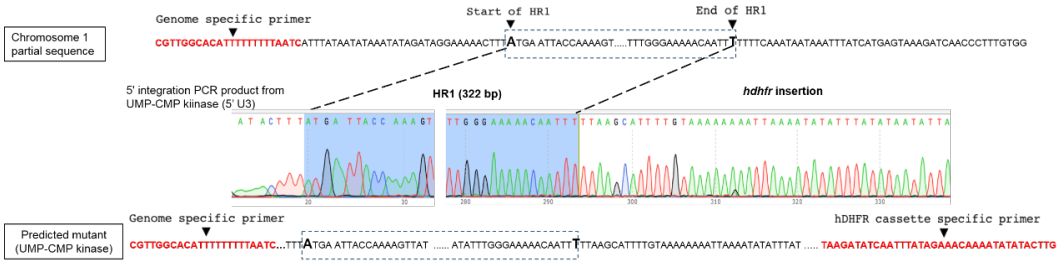



B)


**Figure 3.22 DNA sequencing confirms the *hdhfr* cassette was integrated into the target locus (KAHRP2) through use of the CRISPR-Cas9 system**

The partial sequence of chromosome 2 from the wild-type *P. falciparum* strain 3D7 was extracted from PlasmoDB as shown in the upper row. The predicted mutant sequence, generated by using APE plasmid editor, is shown in the lower row. (A) Sequencing result of the donor DNA (*hdhfr*) inserted into the left homology arm (5' integration). (B) Sequencing results of 2 clones (3'K4 and 3'K8) show the donor DNA (*hdhfr*) integrated into the right homology arm (3' integration). The primers were designed to amplify each fragment represented in bold red.

A)

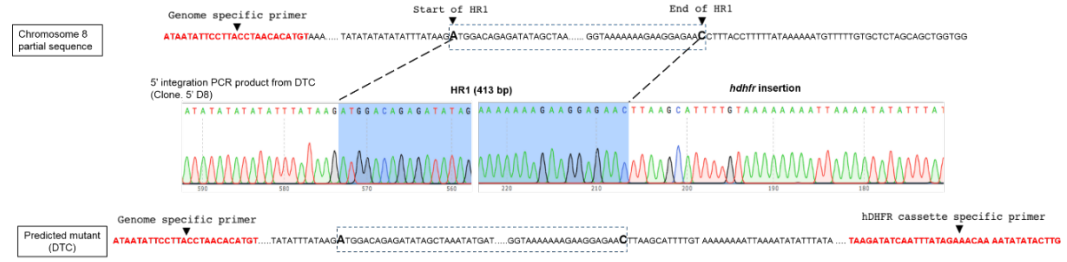



B)


**Figure 3.23 DNA sequencing confirms the *hdhfr* cassette was integrated into the target locus (UMP-CMP kinase) through use of the CRISPR-Cas9 system**

The partial sequence of chromosome 1 from the wild-type *P. falciparum* strain 3D7 was extracted from PlasmoDB shown in the upper row. The predicted mutant sequence, generated by using APE plasmid editor, is shown in the lower row. (A) Sequencing result of the donor DNA (*hdhfr*) that was inserted into the left homology arm (5' integration). (B) Sequencing result of the donor DNA (*hdhfr*) was integrated into the right homology arm (3' integration). The primers were designed to amplify each fragment represented in bold red.

A)




B)


**Figure 3.24 DNA sequencing confirms the *hdhfr* cassette was integrated into the target locus (DTC) through use of the CRISPR-Cas9 system**

The partial sequence of chromosome 8 from the wild-type *P. falciparum* strain 3D7 was extracted from PlasmoDB as shown in the upper row. The predicted mutant sequence, generated by using APE plasmid editor, is shown in the lower row. (A) Sequencing resultof the donor DNA (*hdhfr*) that was inserted into the left homology arm (5' integration). (B) Sequencing results of 3 clones (3'D1, 3'D2 and 3'D3) show the donor DNA (*hdhfr*) that was integrated into the right homology arm (3' integration). The primers were designed to amplify each fragment as represented in bold red.

## Chapter 4 Discussion

The main objective of this study was to utilise a CRISPR-Cas9 system to verify the essential biological functions of two potential drug targets in malaria parasites predicted by the metabolic model of *Plasmodium falciparum*. The set-up of a CRISPR-Cas9 system here could be applied for testing other predicted malaria drug targets. CRISPR-Cas9 gene editing of the human malaria parasite *P. falciparum* was first reported by Ghorbal et al. (2014). It is based on the co-transfection of two plasmids, one being the pUF1-Cas9 plasmid that provides a Cas9 endonuclease for making the double strand break (DSB) in a site-specific of the target DNA, guided by sgRNA and another plasmid (pL7-GOI) that contains the sgRNA and the donor DNA template for homology recombination. This system demonstrated that the stably transfected parasites could be detected 8 days after transfection. Thus, the well-established CRISPR-Cas9 system in *Plasmodium* spp. provides a robust and high efficiency gene modification approach, that could bring about the application of basic biological research into new treatments and vaccine development against malaria.

In this research, two genes, UMP-CMP kinase and dicarboxylate/ tricarboxylate carrier protein, were subjected to drug target validation via the CRISPR-Cas9 system. The original plasmid pL6 eGFP, obtained from Ghorbal et al. (2014), allows for genetic modification of other target loci in *P. falciparum* by changing the homology sequences and introducing new oligonucleotides of the targeted sgRNA into the plasmid. Construction of plasmids in this project started from the amplification of HR1 and HR2 of GOIs, then both fragments were cloned into the original plasmid via replacing the GFP 5' and GFP 3' with HR1 and HR2, respectively. The optimised PCR condition was achieved by using the PCR extension temperature at 62°C for all amplification of homology regions except the HR2 of UMP-CMP kinase. The optimal temperature for extension of that fragment was found to be 64°C. It has been reported that a reduction of the PCR extension temperature is required for the amplification of extremely A+T rich DNA, particularly in *P. falciparum* genomic DNA (Su *et al.* 1996). This finding demonstrated that fragments with ~90% A+T rich content were successfully amplified by using the extension temperature at 60°C while a failure to achieve the correct PCR product occurred at 65°C and 72°C, suggesting that using low extension temperatures prevents the melting of the A+T rich DNA templates during the extension step. Therefore, the successful

amplification of homology regions in this study is very encouraging for re-considering the optimized temperature for amplifying DNA fragments from *P. falciparum* genomic DNA.

After obtaining the HR fragments, the insertion of those fragments into the plasmid pL6 eGFP was performed through the use of an In-Fusion cloning kit. In the beginning, the positive clones were screened by digestion with the restriction enzyme *Bcl*I since the restriction site of this enzyme was added into the forward primer, allowing for positive clone screening. The results showed a slightly shifted band after digestion, which was difficult to distinguish among the positive clones carrying the *Bcl*I restriction site and the backbone plasmid. It is important to note that the fragment sizes of the homology region are around 300–500 bp, not very different from the fragments of GFP 5' (341 bp) and GFP 3' (269 bp) where the homology regions were replaced. Thus, screening of the positive clones by enzyme digestion is not a proper technique for screening and this step was skipped for the further experiments.

All clones were confirmed to have the correct sequence by Sanger sequencing. The introduction of annealed oligonucleotides of targeting sgRNA into the plasmid via replacement of the *Btg*ZI adapter region (between the U6 promoter and terminator) resulted in the seamless insertion of sgRNA, as confirmed by Sanger sequencing. It is important that the sgRNA has to insert into the corresponding site to ensure that the sgRNA was transcribed under U6 promoter. Then, the subsequent plasmid constructions of UMP-CMP kinase and DTC were used for the transfection into *P. falciparum* parasites.

The transfection of the two circular plasmids into the parasites was performed by electroporation using 0.31 kV, 25 μF and 200 Ω resistance with 2 mm cuvettes on a BioRad Gene Pulser II and the synchronized ring-stage parasites (8–10% parasitaemia) were used for co-transfection. Similar to the previous successful transfection in *P. falciparum*, the intracellular ring-stage is the predominant stage of parasites used for routine transfection by electroporation because the ring parasites can be cultured *in vitro* and rapidly purified as a synchronous ring form. In addition, the ring-stage can tolerate the electroporation that causes damage to both the host cells and the parasites (Wu *et al.* 1995, Deitsch *et al.* 2001, Crabb *et al.* 2004). For the transfection method, electroporation is the only method that has been developed for introducing exogenous DNA into *P. falciparum* infected erythrocytes with consistent reliability and this method has been widely used. The transfection of plasmid DNA into the ring stage of *P. falciparum* using electroporation was

first successfully demonstrated by Wu et al. (1995). The conditions of 2.5 kV and 25 μF with 4 mm cuvettes was used for the routine electroporation. Consequently, in 1997, a new condition for electroporation by using 0.31 kV and 960 μF with 2 mm cuvettes showed more efficiency than the previous methods (Fidock and Wellems 1997). In this experiment, the conditions as described by Fidock, D.A., and Wellems, T.E (1997) were used for electroporation but boiled parasites (with burning) were observed after electroporation. Therefore, a modified condition with 0.31 kV, 25 μF and 200 Ω resistance with 2 mm cuvettes was applied for all transfections of circular plasmids into the parasites. When comparing this with the previous conditions, using a high voltage/low capacitance electric pulse, presented by Wu et al. (1995) or a low voltage/high capacitance electric pulse from the protocol of Fidock, D.A., and Wellems, T.E (1997), both conditions were successful in transfecting ring-stage parasites *P. falciparum,* while the conditions used in this experiment involved a low voltage/low capacitance electric pulse. It is possible that the different conditions used here may have resulted in the low transfection efficiency because the results showed only a small number of parasites were detected among the non-essential (KAHRP2) transgenic parasites. Furthermore, the failure of recovery of the KAHRP2 transgenic parasites after 2 months may result from the highly inefficient transfection.

Evaluation of the integrant parasites was performed by PCR analysis after 7 days of transfection. As the genotyping results showed the repeatedly successful integrations of the donor DNA (*hdhfr* cassette) into the target locus of the transgenic parasites KAHRP2, UMP-CMP kinase and DTC while KAHRP, obtained for the Pasteur Institute showed a negative genotyping result (the experiment was repeated more than 4 times in our group). A possible reason for this is the vector length could contribute to transfection efficiency by electroporation. A previous study reported the transfection efficiency per moles of transfected DNA decreased in response to increasing vector length when DNA was transfected into HeLa cells by electroporation (Hornstein *et al.* 2016). Additionally, the entry of DNA into the nucleus by passive diffusion showed that the movement of the longer pieces of DNA is slower than for shorter pieces of DNA (Ludtke *et al.* 1999). Moreover, it should be highlighted that the two plasmids of the CRISPR-Cas9 system need to reach the parasite nucleus for gene modification, meaning that the plasmids have to pass through four membranes: the RBC membrane, the parasitophorous membrane, the parasite cytoplasm membrane and the parasite nucleus membrane. Hence, larger sized vectors might not end up in

the parasite nucleus after electroporation (Gopalakrishnan *et al.* 2013). When compared with the plasmid size used in this research, KAHRP (10971 bp), KAHRP2 (9818 bp), UMP-CMP kinase (9915 bp) and DTC (9999 bp) and among the four constructs, three of them (KAHRP2, UMP-CMP kinase and DTC) had a plasmid size smaller than 10 kb and showed the positive genotyping results and DNA sequencing confirmed successful integration of the mutant. Therefore, for further investigation of drug target validation in *P. falciparum* using the CRISPR-Cas9 system, the size of the plasmid should be taken into account.

For the knockout UMP-CMP kinase and DTC using the CRISPR-Cas9 system, as established in this study it could be applied to other predicted malaria drug targets with an efficient and site-specific model for genome editing. The results as presented the *hdhfr* cassette was successfully incorporated into the target locus only 7 days post-transfection according to the integrated mutants that were detected by genotyping. By using diagnostic PCR primers, the presence of the PCR products from the wild-type genomic DNA was unexpected and presented the same bands with the expected sizes for the integrated mutants. The presence of unexpected bands in the wild-type sample could be caused by the contamination between the PCR reactions during preparation. In addition, non-specific binding of diagnostic primers was observed since various sizes of PCR products were obtained. In addition to this, optimization of the PCR conditions including preparation of PCR reactions carefully or designing new diagnostic primers may prove useful in genotyping.

However, the PCR products obtained from genotyping with the expected sizes confirmed the insertion of the *hdhfr* cassette into the target loci by Sanger sequencing. The DNA sequences of the left and right homology arms showed the *hdhfr* cassette was correctly inserted into the cutting site of the target locus. These results indicate that the guide RNAs that were designed in this study can guide the Cas9 endonuclease to the site-specific target before the double-strand DNA is cleaved by the endonuclease and the 300–500 bp of homologous templates were efficient in mediating the homology-directed recombination.

During phenotype analysis, after drug selection the transgenic parasites were continuously monitored for 60 days post-transfection. Then, if there is an appearance of transgenic parasites, the negative selection (40 μm of 5-Fluorocytosine) will be applied to select the parasites free of episomal mutants. Due to the presence of suicide gene *yfcu* (yeast cytosine deaminase and uridyl

phosphoribosyl transferase) in the plasmid pL7-GOI, 5-Fluorocytosine kills parasites that carry the transient plasmid. However, as transfectants were cultured for two months, episomes will not be maintained and are lost from the parasites during propagation in the absence of drug pressure (van Dijk *et al.* 1997, O'Donnell *et al.* 2001). Thus, the negative selection is not required in this case.

The failure of recovery of the non-essential KAHRP2 parasites could be explained by the transfection efficiency among *P. falciparum* being extremely low (~1 × $10^{-6}$) because the exogenous DNA has to traverse 4 membranes (O'Donnell *et al.* 2002). Some protocols presented that parasites were typically detected by Giemsa smear within 4-6 weeks (Wagner et al 2015). Moreover, according to the previous publication in "method in malaria research" recommends for growing the transfectants up to 75 days (Moll *et al*., 2008). Therefore, the appearance of transfectants in this study may also take longer than 2 months. Additionally, the conditions used for electroporation also affects the transfection efficiency as mentioned above. For the transgenic parasites of UMP-CMP kinase and DTC, no viable parasites were detected after 2 months. In principle, for gene knockout by the CRISPR-Cas9 system, the successful integration of the *hdhfr* cassette into the target locus will result in the disruption of the gene function of the target. Hence there will be no parasite survivors if the GOI is an essential gene. The results from this study suggest that UMP-CMP kinase and DTC could have a crucial biological function for parasites.

A recent publication used a different approach to identify essential genes throughout the genome. Zhang et al. (2018) used genome-wide saturation mutagenesis in *P. falciparum* using the piggyBac system. It is based on the AT-rich *P. falciparum* genome that allows for piggyBac transposons (integrating into sequences of TTAA) that can insert into targets, generating around 38,000 mutants with nonmutable (essential genes) and mutable (dispensable genes). UMP-CMP kinase and DTC were identified as essential genes in their study (Zhang *et al.* 2018). Thus, the findings in this project are consistent with this published data on non-viable parasites of both transgenes. In this study, however, the viability of each transgenic parasite needs to be further investigated using the quantitative approach via RT-qPCR to confirm the essentiality of UMP-CMP kinase and DTC. RT-qPCR is considered to be an effective approach for detection and quantification of mRNA transcripts of interest in complex environmental samples (Smith and Osborn 2009). Using such a technique may be able to confirm the living or dead parasites of

transgene UMP-CMP kinase and DTC by comparing them with the non-essential gene (KAHRP2), leading to validation of the drug target.

## Chapter 5 Conclusions and recommendation

In this research, two CRISPR plasmids (containing sgRNA and donor DNA for homology recombination) of pL7-UMP-CMP kinase and pL7-DTC were successfully constructed and co-transfected with the Cas9 plasmid into *P. falciparum*-infected erythrocytes. When the targeting gene was disrupted by the donor DNA (*hdhfr*) it allowed for investigation into its biological functions. Genotyping and DNA sequencing confirmed that the *hdhfr* cassette was integrated into the targeted locus, indicating the gene disruption of UMP-CMP kinase and DTC were successfully achieved through the CRISPR-Cas9 system. For drug target evaluation, no visible viable parasites were detected for the transgenic parasites of UMP-CMP and DTC by Giemsa-stained thin blood smears. This result suggests that UMP-CMP kinase and DTC could be important for parasites survival. To confirm the essential function of both genes, quantitative RT-qPCR data is required and needs to be compared with a non-essential gene (KAHRP2). This information could validate both genes as therapeutic targets.

For the non-essential transgenic parasites (KAHRP2), the genotyping result and DNA sequencing confirmed that the donor DNA was integrated into the KAHRP locus, allowing for gene disruption. The transgenic parasites were detected 20 days post-transfection and after that, there was no parasite development. This is contradictory to a previous study that reported that KAHRP gene disrupted parasites can be detected 8 days after transfection. In further investigation, the method for introducing DNA into the malaria parasites must be revised to improve the transfection efficiency. A previous study compared transfection efficiency of three electroporation techniques in *P. falciparum*. The transfection efficiency of the pre-loaded erythrocytes technique is about 180-fold higher than the direct electroporation of ring stage infected erythrocytes while the combined technique (using the direct electroporation of ring stage infected erythrocytes and parasites that were allowed to mature for 24 hours, then followed by the pre-loaded erythrocytes) was found to be 6-fold less efficient than the pre-loaded erythrocytes technique (Hasenkamp *et al.* 2012).

The pre-loaded erythrocytes technique transfers DNA into erythrocytes by electroporation prior to the parasites transfection by the invasion of DNA loaded erythrocytes. DNA is then taken up spontaneously from the host cell cytoplasm into the parasite's nuclei (Deitsch *et al.* 2001). The advantage of this technique is that the parasites tend to be healthier because they are not electroporated, resulting in an improved viability rate. Moreover, the pre-loaded erythrocytes technique has been successfully used in CRISPR-Cas9 transfections with nearly 100 transfections (Ribeiro *et al.* 2018). Furthermore, human serum may help parasites recovery faster. A 20% pooled-human serum has been used for cloning transgenic parasites (Mogollon *et al.* 2016). Therefore, using a pre-loaded erythrocytes technique may maximize the transfection efficiency in *P. falciparum* and when applying the 20% human serum to the parasites culture, the transgenic parasites may be detected earlier.

Another transfection technique that can improve the transfection efficiency in *P. falciparum* is plasmid-free CRISPR/Cas9 genome editing. This technique uses the recombinant Cas9 protein complex with synthetic guide RNAs (ribonucleoprotein, RNP) and single-stranded oligodeoxynucleotide (ssODN). This complex solution is then transfected into the synchronised ring-stage parasites. The introduction of the protein complex into the parasites is far easier than the use of circular plasmids. Based on the CRISPR-Cas9 system, the success of gene editing is subjected to the co-existence of two plasmids in the nucleus of parasites and the size of plasmids also has an effect on the transfection efficiency. Thus, a plasmid-free CRISPR/Cas9 genome editing system may lead to an improvement in the transfection of malaria parasites (Crawford *et al.* 2017).

Finally, it has been recently reported that UMP-CMP kinase and DTC have been identified as essential genes in *P. falciparum* using the piggyBac system; however, there is a lack of biochemical data of a novel target of UMP-CMP kinase. Regarding DTC, although this protein has been purified and characterized, the available drug discovery and drug development information are still limited. Future investigation of both proteins should focus on their biochemistry, for example, recombinant protein production, enzyme

characterization, and crystallisation to understand their enzyme properties. This information can be used to guide rational-drug design and lead to optimization of anti-malarial approaches in the future.

## References


Ahmed, M. A. and Cox-Singh, J. (2015) 'Plasmodium knowlesi - an emerging pathogen', *ISBT Sci Ser*, 10(Suppl 1), 134-140.

Al-Maktari, M. T., Bassiouny, H., Al-Hamd, Z., Assabri, A., El-Massry, A. and Shatat, H. (2003) 'Malaria status in Al-Hodeidah Governorate, Yemen: malariometric parasitic survey & chloroquine resistance P. falciparum local strain', *Journal of the Egyptian Society of Parasitology*, 33(2), 361-372.

Anderson, T. J., Nair, S., Qin, H., Singlam, S., Brockman, A., Paiphun, L. and Nosten, F. (2005) 'Are transporter genes other than the chloroquine resistance locus (pfcrt) and multidrug resistance gene (pfmdr) associated with antimalarial drug resistance?', *Antimicrobial agents and chemotherapy*, 49(6), 2180-2188.

Ariey, F., Witkowski, B., Amaratunga, C., Beghain, J., Langlois, A.-C., Khim, N., Kim, S., Duru, V., Bouchier, C. and Ma, L. (2014) 'A molecular marker of artemisinin-resistant Plasmodium falciparum malaria', *Nature*, 505(7481), 50.

Armstrong, C. M. and Goldberg, D. E. (2007) 'An FKBP destabilization domain modulates protein levels in Plasmodium falciparum', *Nature methods*, 4(12), 1007.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D. A. and Horvath, P. (2007) 'CRISPR provides acquired resistance against viruses in prokaryotes', *science*, 315(5819), 1709-1712.

Brouns, S. J., Jore, M. M., Lundgren, M., Westra, E. R., Slijkhuis, R. J., Snijders, A. P., Dickman, M. J., Makarova, K. S., Koonin, E. V. and Van Der Oost, J. (2008) 'Small CRISPR RNAs guide antiviral defense in prokaryotes', *science*, 321(5891), 960-964.

Cojean, S., Noël, A., Garnier, D., Hubert, V., Le Bras, J. and Durand, R. (2006) 'Lack of association between putative transporter gene polymorphisms in Plasmodium falciparum and chloroquine resistance in imported malaria isolates from Africa', *Malaria journal*, 5(1), 24.

Collins, C. R., Das, S., Wong, E. H., Andenmatten, N., Stallmach, R., Hackett, F., Herman, J. P., Müller, S., Meissner, M. and Blackman, M. J. (2013) 'Robust inducible Cre recombinase activity in the human malaria parasite P lasmodium falciparum enables efficient gene deletion within a single asexual erythrocytic growth cycle', *Molecular microbiology*, 88(4), 687-701.

Cowman, A. F., Healer, J., Marapana, D. and Marsh, K. (2016) 'Malaria: biology and disease', *Cell*, 167(3), 610-624.

Crabb, B. S. and Cowman, A. F. (1996) 'Characterization of promoters and stable transfection by homologous and nonhomologous recombination in Plasmodium falciparum', *Proceedings of the National Academy of Sciences*, 93(14), 7289-7294.

Crabb, B. S., Rug, M., Gilberger, T.-W., Thompson, J. K., Triglia, T., Maier, A. G. and Cowman, A. F. (2004) 'Transfection of the human malaria parasite Plasmodium falciparum' in *Parasite Genomics Protocols*Springer, 263-276.

Crawford, E. D., Quan, J., Horst, J. A., Ebert, D., Wu, W. and DeRisi, J. L. (2017) 'Plasmid-free CRISPR/Cas9 genome editing in Plasmodium falciparum confirms mutations conferring resistance to the dihydroisoquinolone clinical candidate SJ733', *PloS one*, 12(5), e0178163.

D'alessandro, U. and Buttiens, H. (2001) 'History and importance of antimalarial drug resistance', *Tropical Medicine & International Health*, 6(11), 845-848.

de Koning-Ward, T. F., Gilson, P. R. and Crabb, B. S. (2015) 'Advances in molecular genetic systems in malaria', *Nature Reviews Microbiology*, 13(6), 373.

Deitsch, K. W., Driskill, C. L. and Wellems, T. E. (2001) 'Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes', *Nucleic acids research*, 29(3), 850-853.

Dondorp, A. M., Nosten, F., Yi, P., Das, D., Phyo, A. P., Tarning, J., Lwin, K. M., Ariey, F., Hanpithakpong, W. and Lee, S. J. (2009) 'Artemisinin resistance in Plasmodium falciparum malaria', *New England Journal of Medicine*, 361(5), 455-467.

Fidock, D. A. and Wellems, T. E. (1997) 'Transformation with human dihydrofolate reductase renders malaria parasites insensitive to WR99210 but does not affect the intrinsic activity of proguanil', *Proceedings of the National Academy of Sciences*, 94(20), 10931-10936.

Ghorbal, M., Gorman, M., Macpherson, C. R., Martins, R. M., Scherf, A. and Lopez-Rubio, J.-J. (2014) 'Genome editing in the human malaria parasite Plasmodium falciparum using the CRISPR-Cas9 system', *Nature biotechnology*, 32(8), 819.

Gopalakrishnan, A. M., Kundu, A. K., Mandal, T. K. and Kumar, N. (2013) 'Novel nanosomes for gene delivery to Plasmodium falciparum-infected red blood cells', *Scientific reports*, 3, 1534.

Hasenkamp, S., Russell, K. T. and Horrocks, P. (2012) 'Comparison of the absolute and relative efficiencies of electroporation-based transfection protocols for Plasmodium falciparum', *Malaria journal*, 11(1), 210.

Hornstein, B. D., Roman, D., Arévalo-Soliz, L. M., Engevik, M. A. and Zechiedrich, L. (2016) 'Effects of circular DNA length on transfection efficiency by electroporation into HeLa cells', *PloS one*, 11(12), e0167537.

Janse, C. J., Kroeze, H., van Wigcheren, A., Mededovic, S., Fonager, J., Franke-Fayard, B., Waters, A. P. and Khan, S. M. (2011) 'A genotype and phenotype database of genetically modified malaria-parasites', *Trends in parasitology*, 27(1), 31-39.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A. and Charpentier, E. (2012) 'A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity', *science*, 1225829.

Kafsack, B. F., Rovira-Graells, N., Clark, T. G., Bancells, C., Crowley, V. M., Campino, S. G., Williams, A. E., Drought, L. G., Kwiatkowski, D. P. and Baker, D. A. (2014) 'A transcriptional switch underlies commitment to sexual development in malaria parasites', *Nature*, 507(7491), 248.

Ke, H., Lewis, I. A., Morrisey, J. M., McLean, K. J., Ganesan, S. M., Painter, H. J., Mather, M. W., Jacobs-Lorena, M., Llinás, M. and Vaidya, A. B. (2015) 'Genetic investigation of tricarboxylic acid metabolism during the Plasmodium falciparum life cycle', *Cell reports*, 11(1), 164-174.

Lee, A. H., Symington, L. S. and Fidock, D. A. (2014) 'DNA repair mechanisms and their biological roles in the malaria parasite Plasmodium falciparum', *Microbiology and Molecular Biology Reviews*, 78(3), 469-486.

Lee, M. C. and Fidock, D. A. (2014) 'CRISPR-mediated genome editing of Plasmodium falciparum malaria parasites', *Genome medicine*, 6(8), 63.

Liou, J.-Y., Dutschman, G. E., Lam, W., Jiang, Z. and Cheng, Y.-C. (2002) 'Characterization of human UMP/CMP kinase and its phosphorylation of D-and L-form deoxycytidine analogue monophosphates', *Cancer research*, 62(6), 1624-1631.

López-Barragán, M. J., Lemieux, J., Quiñones, M., Williamson, K. C., Molina-Cruz, A., Cui, K., Barillas-Mury, C., Zhao, K. and Su, X.-z. (2011) 'Directional gene expression and antisense transcripts in sexual and asexual stages of Plasmodium falciparum', *BMC genomics*, 12(1), 587.

Ludtke, J. J., Zhang, G., Sebestyén, M. G. and Wolff, J. A. (1999) 'A nuclear localization signal can enhance both the nuclear transport and expression of 1 kb DNA', *Journal of cell science*, 112(12), 2033-2041.

Mbengue, A., Bhattacharjee, S., Pandharkar, T., Liu, H., Estiu, G., Stahelin, R. V., Rizk, S. S., Njimoh, D. L., Ryan, Y. and Chotivanich, K. (2015) 'A molecular mechanism of artemisinin resistance in Plasmodium falciparum malaria', *Nature*, 520(7549), 683.

Meshnick, S. R. and Dobson, M. J. (2001) 'The history of antimalarial drugs' in *Antimalarial chemotherapy*Springer, 15-25.

Miotto, O., Amato, R., Ashley, E. A., MacInnis, B., Almagro-Garcia, J., Amaratunga, C., Lim, P., Mead, D., Oyola, S. O. and Dhorda, M. (2015) 'Genetic architecture of artemisinin-resistant Plasmodium falciparum', *Nature genetics*, 47(3), 226.

Mita, T., Tanabe, K. and Kita, K. (2009) 'Spread and evolution of Plasmodium falciparum drug resistance', *Parasitology international*, 58(3), 201-209.

Mogollon, C. M., van Pul, F. J., Imai, T., Ramesar, J., Chevalley-Maurel, S., de Roo, G. M., Veld, S. A., Kroeze, H., Franke-Fayard, B. M. and Janse, C. J. (2016) 'Rapid generation of marker-free P. falciparum fluorescent reporter lines using modified CRISPR/Cas9 constructs and selection protocol', *PloS one*, 11(12), e0168362.

Naing, C., Whittaker, M. A., Wai, V. N. and Mak, J. W. (2014) 'Is Plasmodium vivax malaria a severe malaria?: a systematic review and meta-analysis', *PLoS neglected tropical diseases*, 8(8), e3071.

Naito, Y., Hino, K., Bono, H. and Ui-Tei, K. (2014) 'CRISPRdirect: software for designing CRISPR/Cas guide RNA with reduced off-target sites', *Bioinformatics*, 31(7), 1120-1123.

Noedl, H., Socheat, D. and Satimai, W. (2009) 'Artemisinin-resistant malaria in Asia', *New England Journal of Medicine*, 361(5), 540-541.

Nozawa, A., Fujimoto, R., Matsuoka, H., Tsuboi, T. and Tozawa, Y. (2011) 'Cell-free synthesis, reconstitution, and characterization of a mitochondrial dicarboxylate–tricarboxylate carrier of Plasmodium falciparum', *Biochemical and biophysical research communications*, 414(3), 612-617.

Nzila, A. M., Nduati, E., Mberu, E. K., Hopkins Sibley, C., Monks, S. A., Winstanley, P. A. and Watkins, W. M. (2000) 'Molecular evidence of greater selective pressure for drug resistance exerted by the long-acting antifolate pyrimethamine/sulfadoxine compared with the shorter-acting chlorproguanil/dapsone on Kenyan Plasmodium falciparum', *Journal of Infectious Diseases*, 181(6), 2023-2028.

O'Donnell, R. A., Freitas-Junior, L. H., Preiser, P. R., Williamson, D. H., Duraisingh, M., McElwain, T. F., Scherf, A., Cowman, A. F. and Crabb, B. S. (2002) 'A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of Plasmodium falciparum chromosomes', *The EMBO journal*, 21(5), 1231-1239.

O'Donnell, R. A., Preiser, P. R., Williamson, D. H., Moore, P. W., Cowman, A. F. and Crabb, B. S. (2001) 'An alteration in concatameric structure is associated with efficient segregation of plasmids in transfected Plasmodium falciparum parasites', *Nucleic acids research*, 29(3), 716-724.

Phyo, A. P. and Nosten, F. (2018) 'The Artemisinin Resistance in Southeast Asia: An Imminent Global Threat to Malaria Elimination' in *Towards Malaria Elimination-A Leap Forward*IntechOpen.

Plasmodium Genomic Resource [online], available: http://plasmodb.org/plasmo/ [accessed

Ribeiro, J. M., Garriga, M., Potchen, N., Crater, A. K., Gupta, A., Ito, D. and Desai, S. A. (2018) 'Guide RNA selection for CRISPR-Cas9 transfections in Plasmodium falciparum', *International journal for parasitology*.

Schlitzer, M. (2007) 'Malaria chemotherapeutics part I: History of antimalarial drug development, currently used therapeutics, and drugs in clinical development', *ChemMedChem*, 2(7), 944-986.

Segura-Peña, D., Sekulic, N., Ort, S., Konrad, M. and Lavie, A. (2004) 'Substrate-induced conformational changes in human UMP/CMP kinase', *Journal of Biological Chemistry*, 279(32), 33882-33889.

Smith, C. J. and Osborn, A. M. (2009) 'Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology', *FEMS microbiology ecology*, 67(1), 6-20.

Snounou, G., Viriyakosol, S., Jarra, W., Thaithong, S. and Brown, K. N. (1993) 'Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections', *Molecular and biochemical parasitology*, 58(2), 283-292.

Straimer, J., Lee, M. C., Lee, A. H., Zeitler, B., Williams, A. E., Pearl, J. R., Zhang, L., Rebar, E. J., Gregory, P. D. and Llinás, M. (2012) 'Site-specific genome editing in Plasmodium falciparum using engineered zinc-finger nucleases', *Nature methods*, 9(10), 993.

Su, X.-z., Wu, Y., Sifri, C. D. and Wellems, T. E. (1996) 'Reduced extension temperatures required for PCR amplification of extremely A+ T-rich DNA', *Nucleic acids research*, 24(8), 1574-1575.

Ta, T. H., Hisam, S., Lanza, M., Jiram, A. I., Ismail, N. and Rubio, J. M. (2014) 'First case of a naturally acquired human infection with Plasmodium cynomolgi', *Malaria journal*, 13(1), 68.

Totanes, F. I. G. (2017) *Improved computational model of the malaria metabolic network and flux analysis for drug target prediction*, unpublished thesis University of Leeds.

Trape, J.-F. (2001) 'The public health impact of chloroquine resistance in Africa', *The American journal of tropical medicine and hygiene*, 64(1_suppl), 12-17.

Triglia, T., Menting, J. G., Wilson, C. and Cowman, A. F. (1997) 'Mutations in dihydropteroate synthase are responsible for sulfone and sulfonamide resistance in Plasmodium falciparum', *Proceedings of the National Academy of Sciences*, 94(25), 13944-13949.

van Dijk, M. R., Vinkenoog, R., Ramesar, J., Vervenne, R. A., Waters, A. P. and Janse, C. J. (1997) 'Replication, expression and segregation of plasmid-borne DNA in genetically transformed malaria parasites', *Molecular and biochemical parasitology*, 86(2), 155-162.

Waterkeyn, J., Crabb, B. and Cowman, A. (1999) 'Transfection of the human malaria parasite Plasmodium falciparum', *International journal for parasitology*, 29(6), 945-955.

White, N. (2008) 'Plasmodium knowlesi: the fifth human malaria parasite'.

World Health Organization (2015) *Guidelines for the treatment of malaria 3rd Ed. Geneva: WHO Press* [online], available: http://www.who.int/malaria/publications/atoz/9789241549127/en/ [accessed

World Health Organization (2017) *World Malaria Report 2017 Geneva* [online], available: http://www.who.int/mediacentre/factsheets/fs094/en/ [accessed

Wu, Y., Sifri, C. D., Lei, H.-H., Su, X.-z. and Wellems, T. E. (1995) 'Transfection of Plasmodium falciparum within human red blood cells', *Proceedings of the National Academy of Sciences*, 92(4), 973-977.

Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D. and Mayho, M. (2018) 'Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis', *science*, 360(6388), eaap7847.

# Appendix

**Table A1.** Oligonucleotides used in this study

| Oligo name | Sequence |
|---|---|
| UMP-CMP_HR1_FW | CTTTCCGCGGGGAGGACTAGTGATCAATGAATTACCAAAAGTTATT |
| UMP-CMP_HR1_RV | TTTTTTTACAAAATGCTTAAAAATTGTTTTTCCCAAATA |
| UMP-CMP_HR2_FW | TATTTATTAAATCTAGAATTTGATCAGGATGGATTAATATAATAGG |
| UMP-CMP_HR2_RV | ATTATTTTTACCGTTCCATGTTACATATTTGTAAAAACATAC |
| DTC_HR1_FW | CTTTCCGCGGGGAGGACTAGTGATCAATGGACAGAGATATAGCTAA |
| DTC_HR1_RV | TTTTTTTACAAAATGCTTAAGTTCTCCTTCTTTTTTTACC |
| DTC_HR2_FW | TATTTATTAAATCTAGAATTTGATCACCTGCAGATTTATCTTTAA |
| DTC_HR2_RV | ATTATTTTTACCGTTCCATGTTGAATAGAACAATCTAACAT |
| sgRNA_oligo1_UMP-CMP kinase | CATATTAAGTATATAATATTGAGTAAAGATCAACCCTTTGTGGGTTTTAGAGCTAGAAATAGC |
| sgRNA_oligo2_UMP-CMP kinase | GCTATTTCTAGCTCTAAAACCCACAAAGGGTTGATCTTTACTCAATATTATATACTTAATATG |
| sgRNA_oligo1_DTC | CATATTAAGTATATAATATTGGTGGATTAGGAGCCTTTATTGGGTTTTAGAGCTAGAAATAGC |
| sgRNA_oligo2_DTC | GCTATTTCTAGCTCTAAAACCCAATAAAGGCTCCTAATCCACCAATATTATATACTTAATAT |
| KAHRP_GT_FW1 | TTTATTTGAACAATATTTACTCC |
| GT_RV1 | CCAATAGATAAAATTTGTAGAG |
| GT_FW2 | ATAAAGTACAACATTAATATATAGC |
| KAHRP_GT_RV2 | AGATTATTTAACCACAGC |
| KAHRP_GT_RVb | ACTCTTTTCTGTATAAACG |
| U3 | CGTTGGCACATTTTTTTTTAATC |
| UMP-CMP_GT_FW1 | TTTTTTTACCGTTGGC |
| UMP-CMP_GT_RV2 | AAACTATAAAAAGGGTACG |
| D3 | ATAATATTCCTTACCTAACACATGT |
| D4 | AATTACTTGTGGTTGTGTAATCA |
| P1 | CAAGTATATATTTTGTTTCTATAAATTGATATCTTA |
| P2 | ATTCCTAATCATGTAAATCTTAAATTTTTC |
| HR1 sequencing primer | CTAAATATATATCCAATGGCCC |
| HR2 sequencing primer | CTTTCTTGAAACGGTTGTCCC |
| sgRNA sequencing primer | TAGGAAATAATAAAAAAGCACC |

```
CLUSTAL multiple sequence alignment by MUSCLE (3.8)

                                                                    Genome specific primer
                                                                              ▼
5'KAHRP (predicted mutant)  ----------------------------------------TTTATTTGAACAATATTTACT
Chro.2 (partial seq.)       TTTTTTACCTAATATTTTATATATATATTTATATATTAGTTTATTTGAACAATATTTACT
5'K13                       ------------------------------NNNNNNTTTTTTATNNNAANAATATTNNNN
                                                            *****   ** ******

5'KAHRP (predicted mutant)  CCATATATTTATTTCATATATATATATATATATATATATATATGTGATGTTTAAAAAATA
Chro.2 (partial seq.)       CCATATATTTATTTCATATATATATATATATATATATATATATGTGATGTTTAAAAAATA
5'K13                       NNANNNNNNNNNNNNANNNNNNNNNNNNNNTATATATATATATGTGATGTTTAAAAAATA
                              *           *              ******************************

5'KAHRP (predicted mutant)  TGAAATATATATTTATATATTATTATTTTCTTATTTTATTTTTATTTTATTTTATTTTAT
Chro.2 (partial seq.)       TGAAATATATATTTATATATTATTATTTTCTTATTTTATTTTTATTTTATTTTATTTTAT
5'K13                       TGAAATATATATTTATATATTATTATTTTCTTATTTTATTTTTATTTTATTTTATTTTAT
                            ************************************************************

5'KAHRP (predicted mutant)  TTTAATTTTATTTTATTTTATTTTAATTTTATTTTAATTTTATTTTAATTTTATTTTAAT
Chro.2 (partial seq.        TTTAATTTTATTTTATTTTATTTTAATTTTATTTTAATTTTATTTTAATTTTATTTTAAT
5'K13                       TTTAATTTTATTTTATTTTATTTTAATTTTATTTTAATTTTATTTTAATTTTATTTTAAT
                            ************************************************************

5'KAHRP (predicted mutant)  TTTATTTTAATTTTATTTTATTTTAATTTTATTTTATTTTATTTTAATTTTATTTTATTT
Chro.2 (partial seq.)       TTTATTTTAATTTTATTTTATTTTAATTTTATTTTATTTTATTTTAATTTTATTTTATTT
5'K13                       TTTATTTTAATTTTATTTTATTTTAATTTTATTTTATTTTATTTTAATTTTATTTTATTT
                            ************************************************************

5'KAHRP (predicted mutant)  TATTTTAATTTTATTTTATTTTATTTTAATTTAATTTTATTTTTTTTTTCATATAGTGCA
Chro.2 (partial seq.)       TATTTTAATTTTATTTTATTTTATTTTAATTTAATTTTATTTTTTTTTTCATATAGTGCA
5'K13                       TATTTTAATTTTATTTTATTTTATTTTAATTTAATTTTATTTTTTTTTTCATATAGTGCA
                            ************************************************************

5'KAHRP (predicted mutant)  ATAATGGAAACGGATCCGGTGACTCCTTCGATTTCAGAAATAAGAGAACTTTAGCACAAA
Chro.2 (partial seq.)       ATAATGGAAACGGATCCGGTGACTCCTTCGATTTCAGAAATAAGAGAACTTTAGCACAAA
5'K13                       ATAATGGAAACGGATCCGGTGACTCCTTCGATTTCAGAAATAAGAGAACTTTAGCACAAA
                            ************************************************************

5'KAHRP (predicted mutant)  AGCAACATGAACACCATCACCACCATCACCATCAACATCAACACCAACACCAAGCTCCAC
Chro.2 (partial seq.)       AGCAACATGAACACCATCACCACCATCACCATCAACATCAACACCAACACCAAGCTCCAC
5'K13                       AGCAACATGAACACCATCACCACCATCACCATCAACATCAACACCAACACCAAGCTCCAC
                            ************************************************************

5'KAHRP (predicted mutant)  ACCAAGCACACCACCATCATCATCATGGAGAAGTAAATCACCAAGCACCACAGGTTCACC
Chro.2 (partial seq.)       ACCAAGCACACCACCATCATCATCATGGAGAAGTAAATCACCAAGCACCACAGGTTCACC
5'K13                       ACCAAGCACACCACCATCATCATCATGGAGAAGTAAATCACCAAGCACCACAGGTTCACC
                            ************************************************************

5'KAHRP (predicted mutant)  AACAAGTACATGGTCAAGACCAAGCACACCATCACCATCATCACCACCATCATCAATTAC
Chro.2 (partial seq.)       AACAAGTACATGGTCAAGACCAAGCACACCATCACCATCATCACCACCATCATCAATTAC
5'K13                       AACAAGTACATGGTCAAGACCAAGCACACCATCACCATCATCACCACCATCATCAATTAC
                            ************************************************************

5'KAHRP (predicted mutant)  AACCTCAACAACCCCAGGGAACAGTTGCTAATCCTCCTAGTAATGAACCAGTTGTAAAAA
Chro.2 (partial seq.)       AACCTCAACAACCCCAGGGAACAGTTGCTAATCCTCCTAGTAATGAACCAGTTGTAAAAA
5'K13                       AACCTCAACAACCCCAGGGAACAGTTGCTAATCCTCCTAGTAATGAACCAGTTGTAAAAA
                            ************************************************************

5'KAHRP (predicted mutant)  CCCAAGTATTCAGGGAAGCAAGACCAGGTGGAGGTTTCAAAGCATATGAAGAAAAATACG
Chro.2 (partial seq.)       CCCAAGTATTCAGGGAAGCAAGACCAGGTGGAGGTTTCAAAGCATATGAAGAAAAATACG
5'K13                       CCCAAGTATTCAGGGAAGCAAGACCAGGTGGAGGTTTCAAAGCATATGAAGAAAAATACG
                            ************************************************************

                                                                  Start of HR1
                                                                        ▼
5'KAHRP (predicted mutant)  AATCAAAACACTATAAATTAAAGGAAAATGTTGTCGATGGTAAAAAAGATTGTGATGAAA
Chro.2 (partial seq.)       AATCAAAACACTATAAATTAAAGGAAAATGTTGTCGATGGTAAAAAAGATTGTGATGAAA
5'K13                       AATCAAAACACTATAAATTAAAGGAAAATGTTGTCGATGGTAAAAAAGATTGTGATGAAA
                            ************************************************************

5'KAHRP (predicted mutant)  AATACGAAGCTGCCAATTATGCTTTCTCCGAAGAGTGCCCATACACCGTAAACGATTATA
Chro.2 (partial seq.)       AATACGAAGCTGCCAATTATGCTTTCTCCGAAGAGTGCCCATACACCGTAAACGATTATA
5'K13                       AATACGAAGCTGCCAATTATGCTTTCTCCGAAGAGTGCCCATACACCGTAAACGATTATA
                            ************************************************************
```


**Figure A2** Sequence alignment of KAHRP2 transgenic parasites at the 5' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

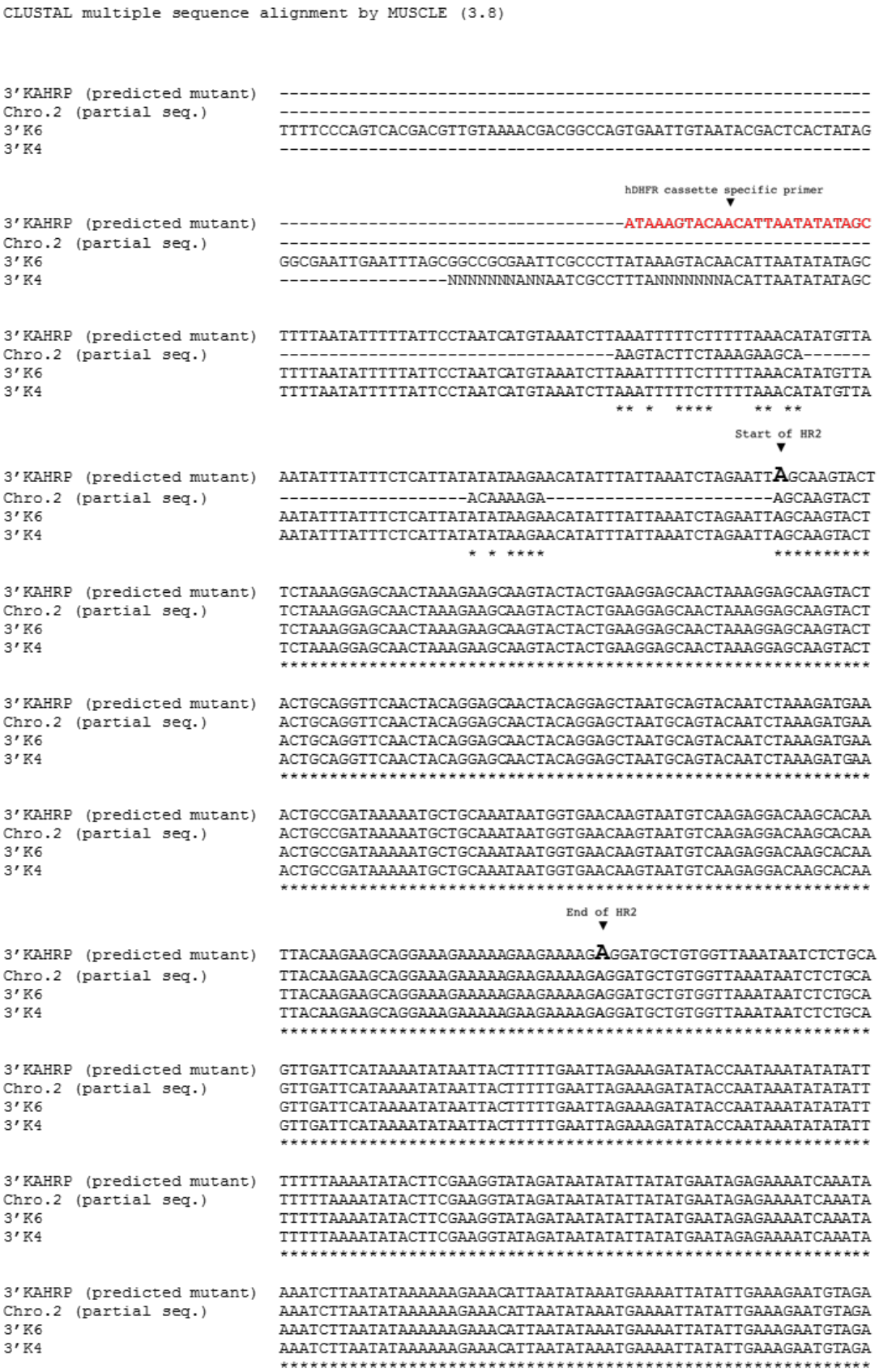

CLUSTAL multiple sequence alignment by MUSCLE (3.8)
3'KAHRP (predicted mutant)
Chro.2 (partial seq.)
3'K6
3'K4
hDHFR cassette specific primer
ATAAAGTACAACATTAATATATAGC
GGCGAATTGAATTTAGCGGCCGCGAATTCGCCCTTATAAAGTACAACATTAATATATAGC
NNNNNNNANNAATCGCCTTTANNNNNNNACATTAATATATAGC
TTTTAATATTTTTATTCCTAATCATGTAAATCTTAAATTTTTCTTTTTAAACATATGTTA
AAGTACTTCTAAAGAAGCA
Start of HR2
AATATTTATTTCTCATTATATATAAGAACATATTTATTAAATCTAGAATTAGCAAGTACT
ACAAAAGA
AGCAAGTACT
TCTAAAGGAGCAACTAAAGAAGCAAGTACTACTGAAGGAGCAACTAAAGGAGCAAGTACT
ACTGCAGGTTCAACTACAGGAGCAACTACAGGAGCTAATGCAGTACAATCTAAAGATGAA
ACTGCCGATAAAAATGCTGCAAATAATGGTGAACAAGTAATGTCAAGAGGACAAGCACAA
End of HR2
TTACAAGAAGCAGGAAAGAAAAAGAAGAAAAGAGGATGCTGTGGTTAAATAATCTCTGCA
GTTGATTCATAAAATATAATTACTTTTTGAATTAGAAAGATATACCAATAAATATATATT
TTTTTAAAATATACTTCGAAGGTATAGATAATATATTATATGAATAGAGAAAATCAAATA
AAATCTTAATATAAAAAAGAAACATTAATATAAATGAAAATTATATTGAAAGAATGTAGA


**Figure A3** Sequence alignment of KAHRP2 transgenic parasites at the 3' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

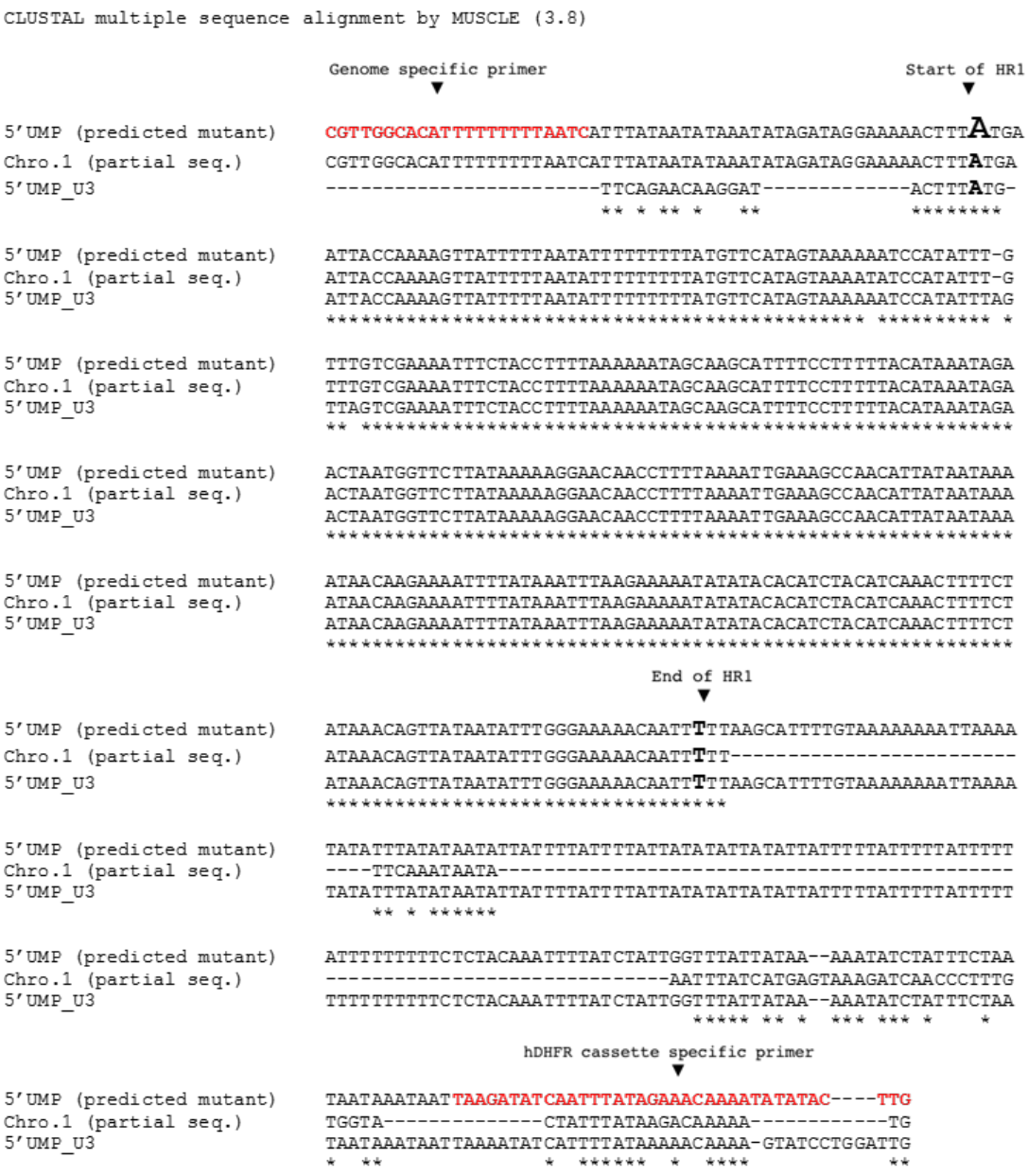



**Figure A4** Sequence alignment of UMP-CMP kinase transgenic parasites at the 5' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

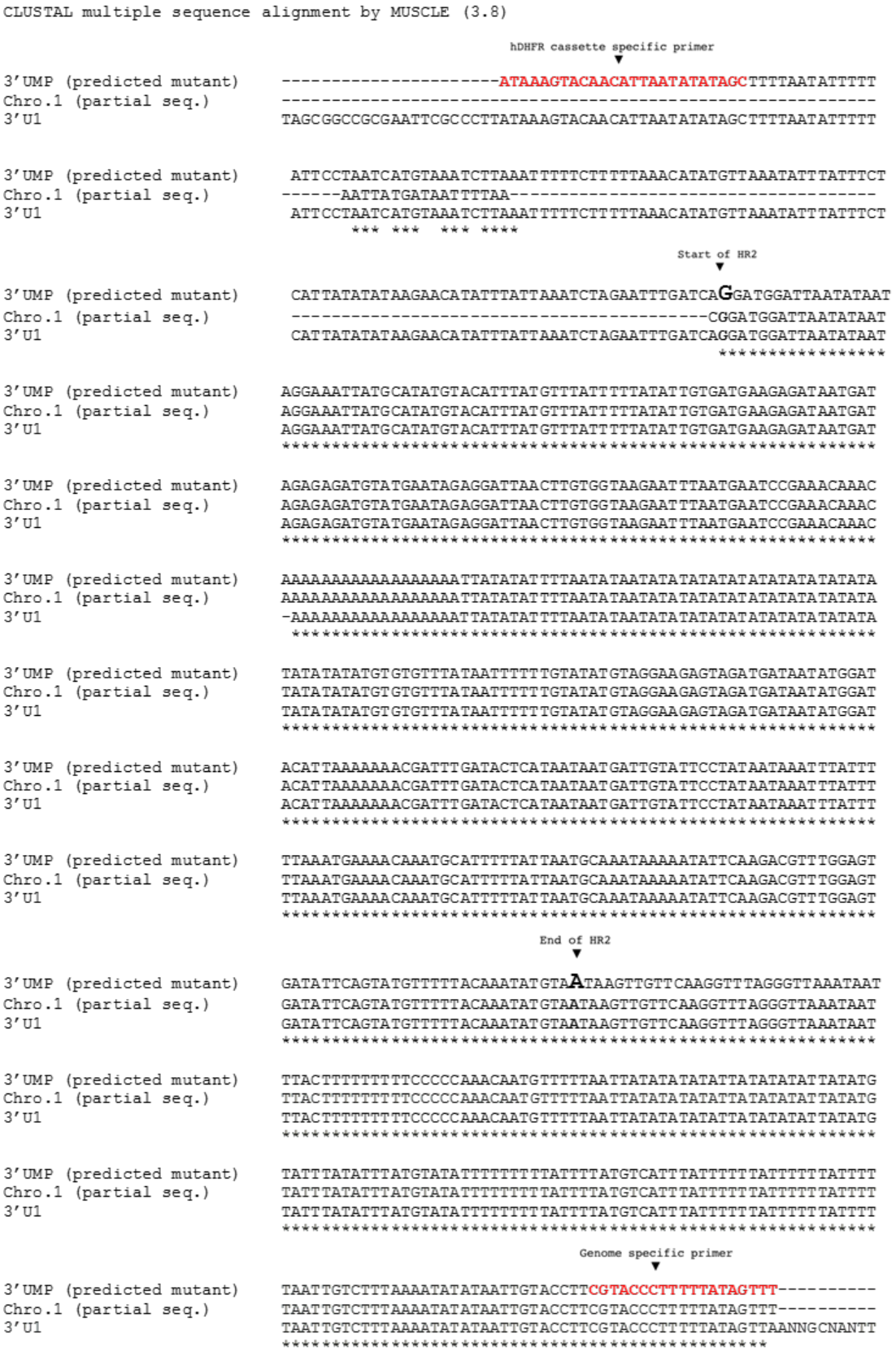



**Figure A5** Sequence alignment of UMP-CMP kinase transgenic parasites at the 3' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

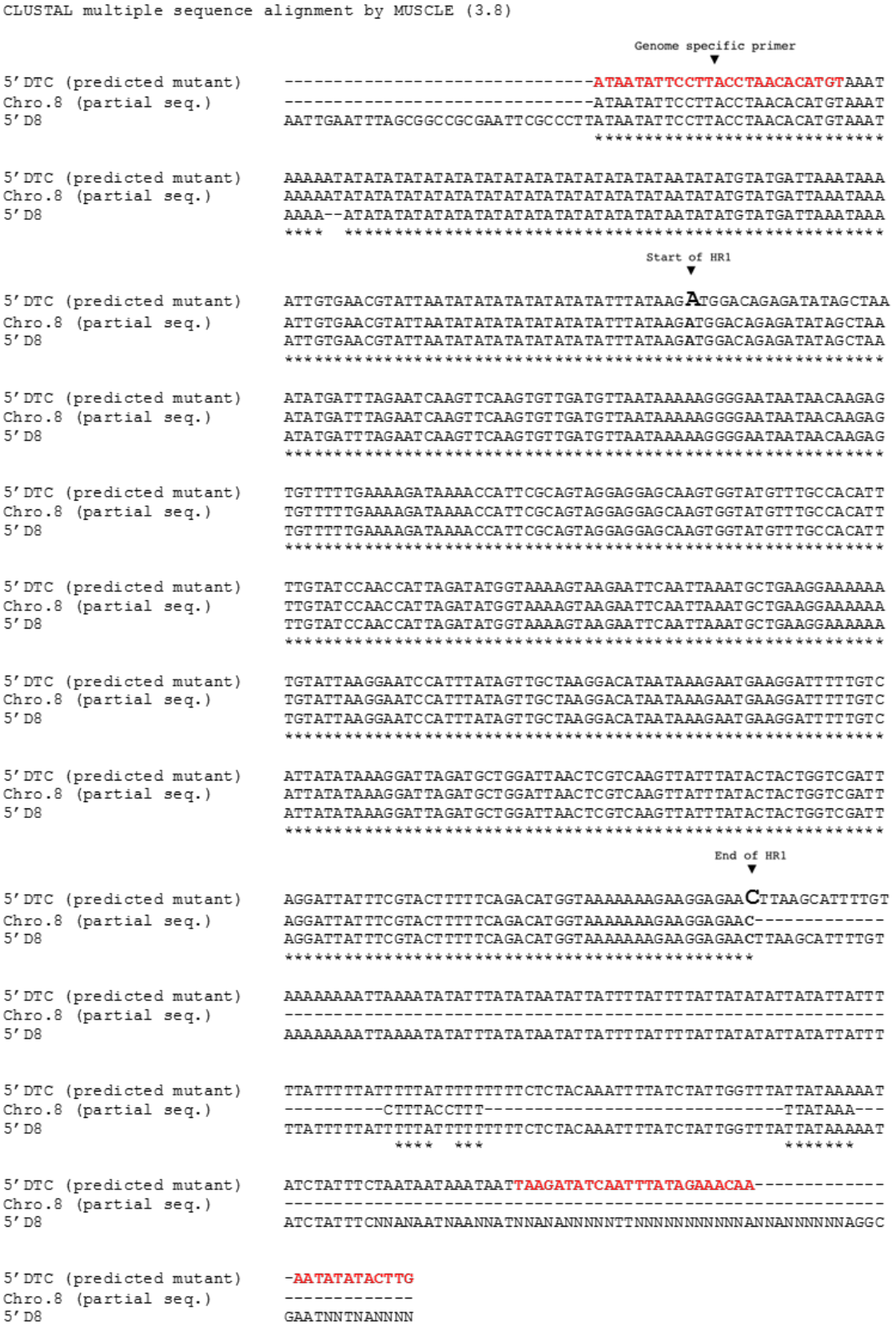



**Figure A6** Sequence alignment of DTC transgenic parasites at the 5' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

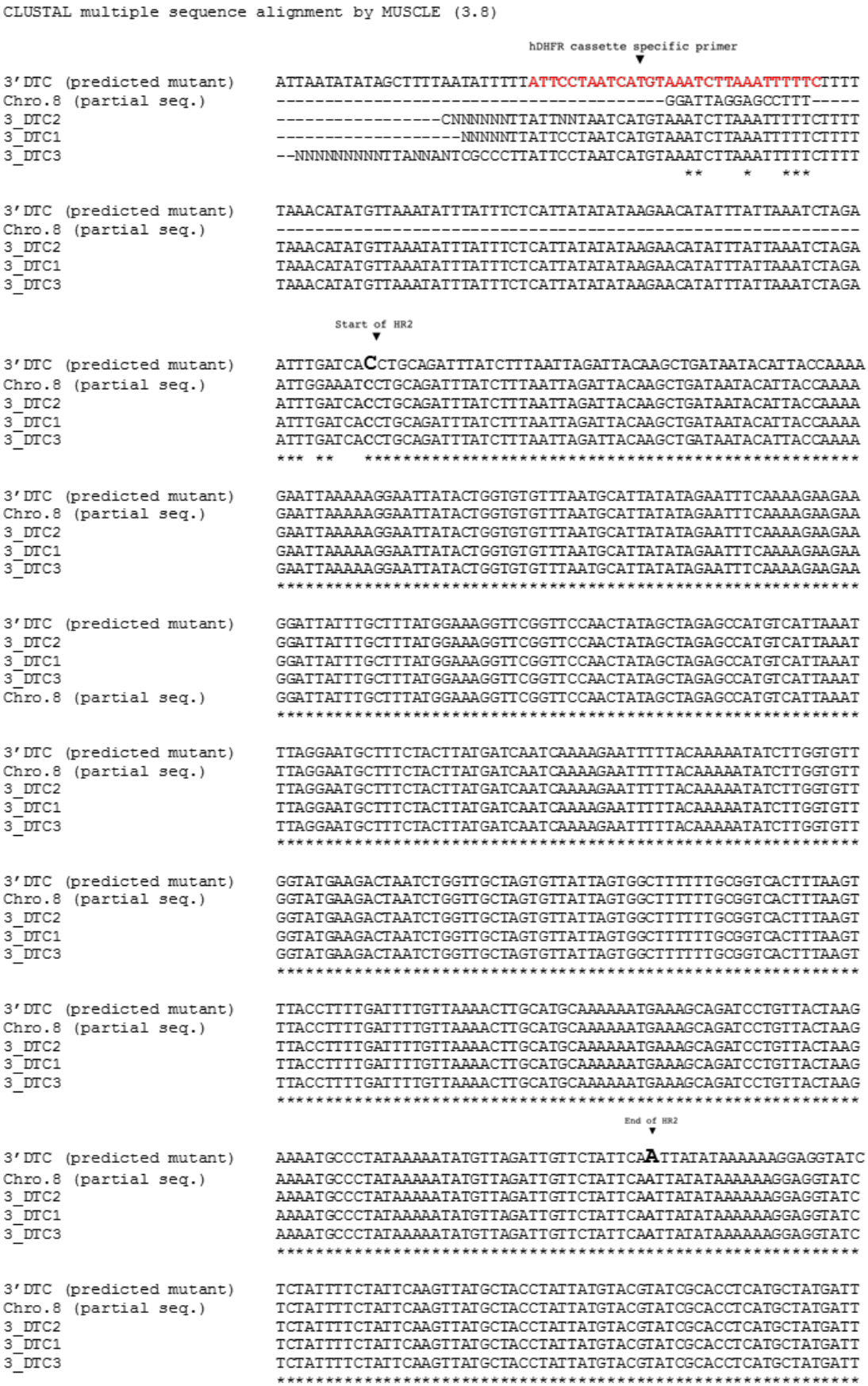

CLUSTAL multiple sequence alignment by MUSCLE (3.8)
hDHFR cassette specific primer
3'DTC (predicted mutant)
Chro.8 (partial seq.)
3_DTC2
3_DTC1
3_DTC3
Start of HR2
End of HR2


**Figure A7** Sequence alignment of DTC transgenic parasites at the 3' integration compare with the predicted mutant sequence and the wild-type sequence. The predicted mutant sequence was generated by using APE plasmid editor.

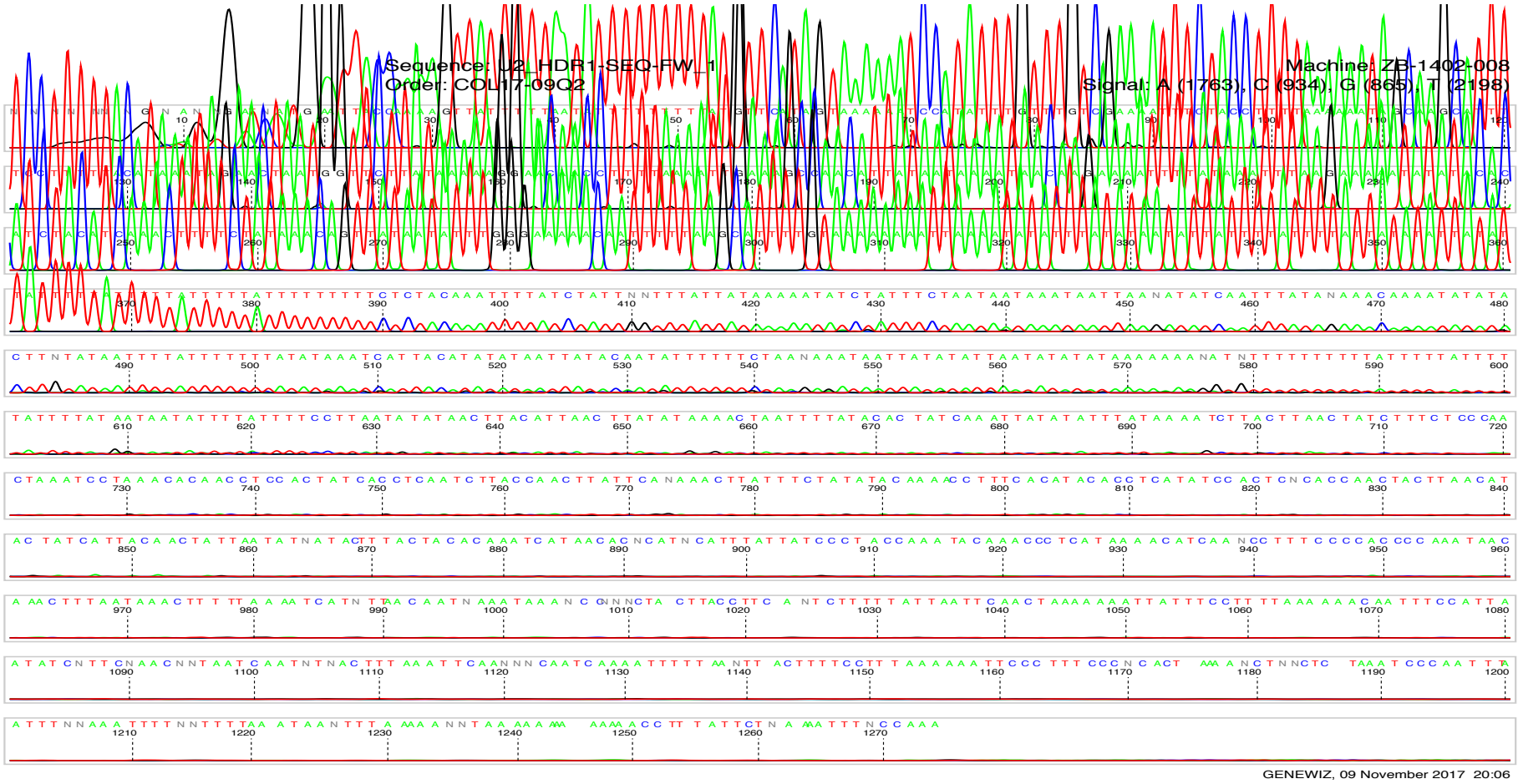



**Figure A8** DNA sequencing of homology region (HR1) of UMP-CMP kinase (clone. HR1_U2)

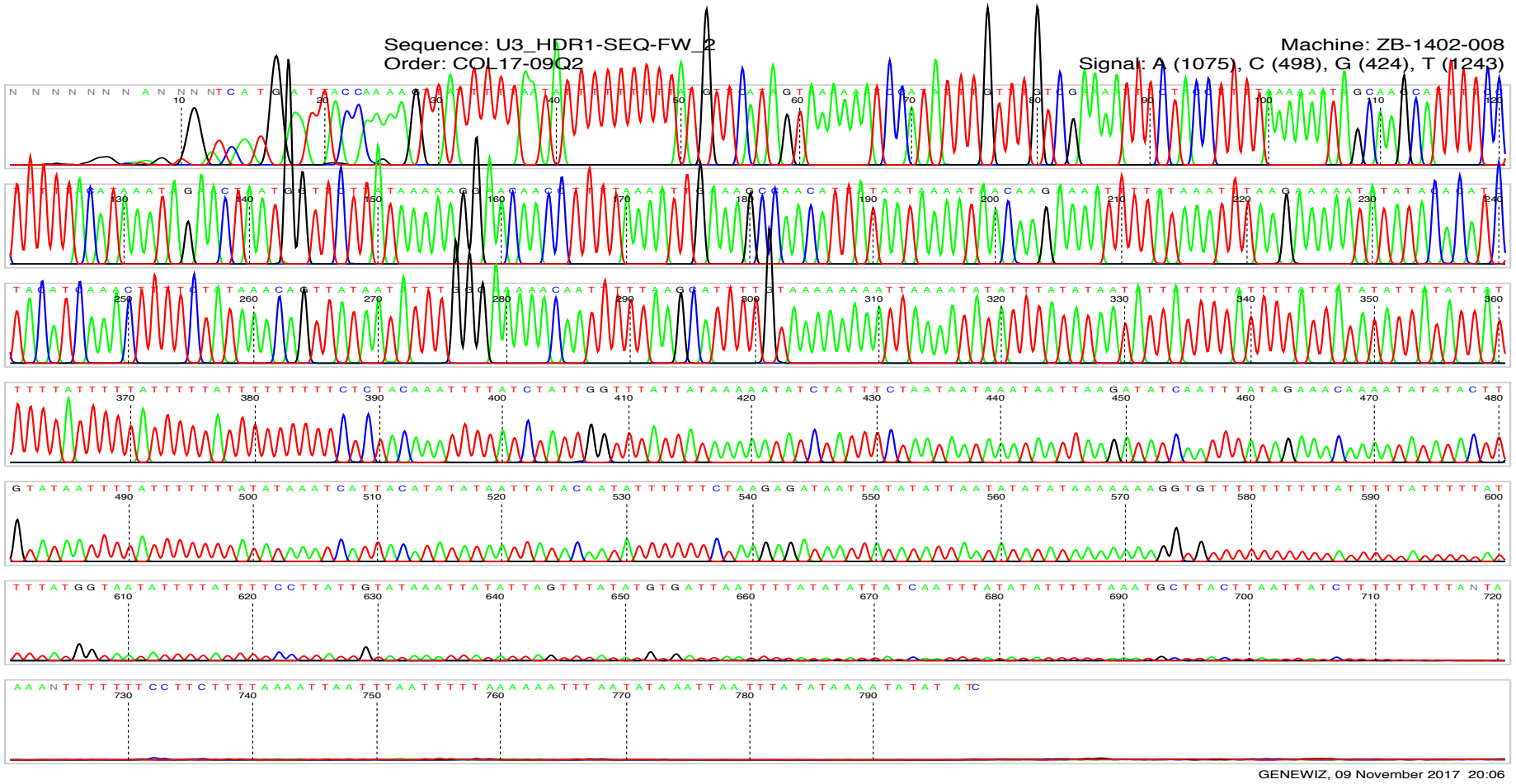



**Figure A9** DNA sequencing of homology region (HR1) of UMP-CMP kinase (clone. HR1_U3)

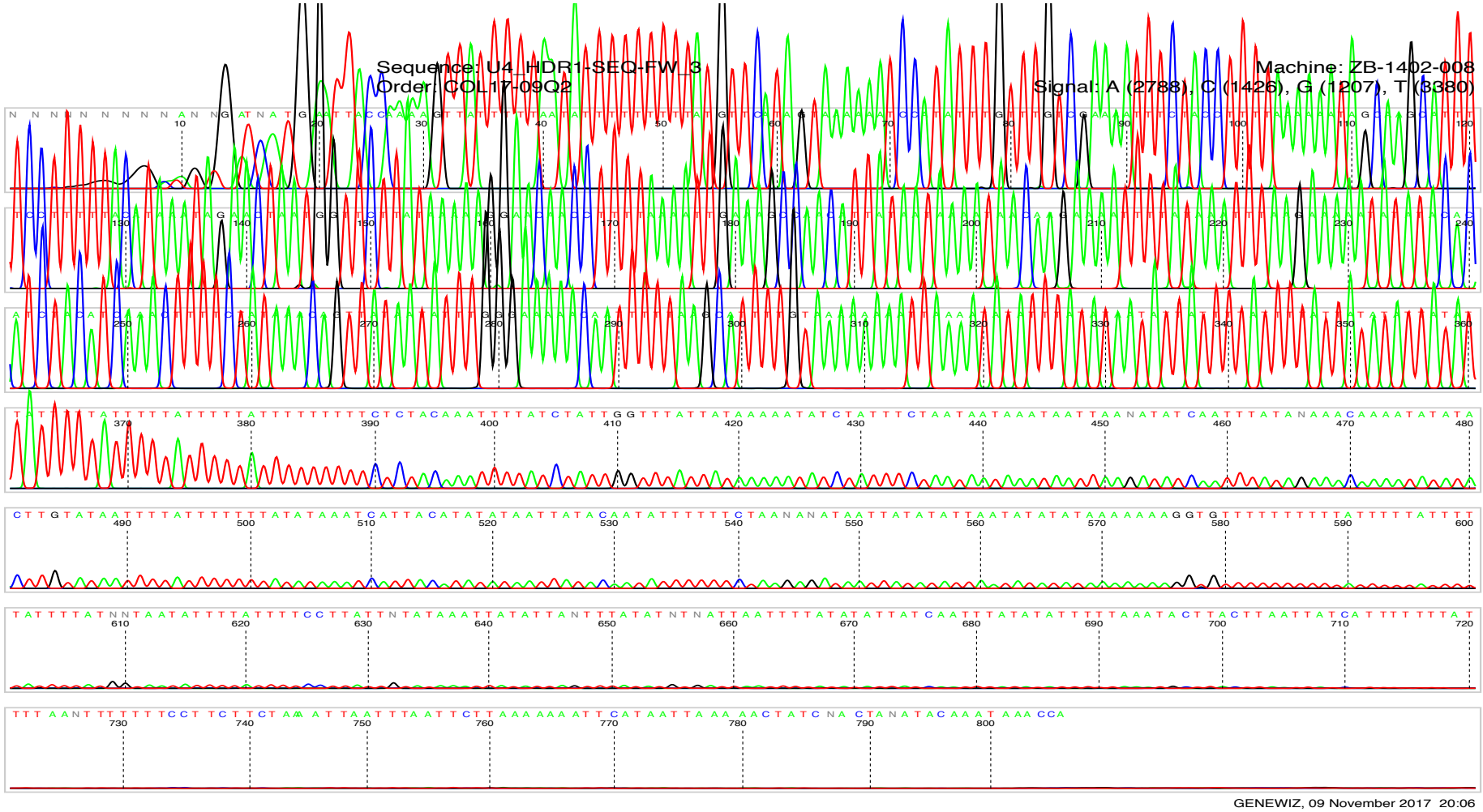



**Figure A10** DNA sequencing of homology region (HR1) of UMP-CMP kinase (clone. HR1_U4)

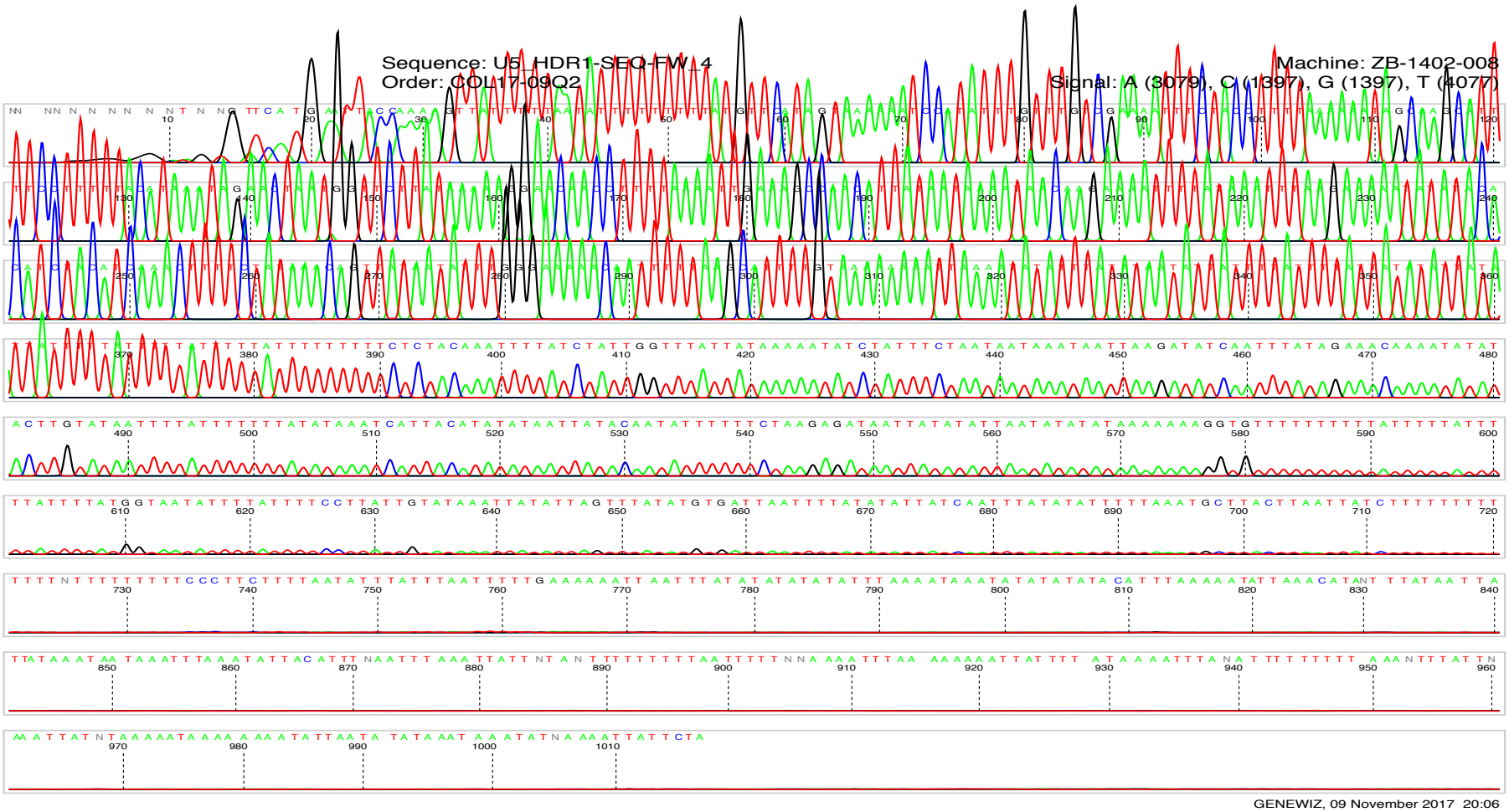



**Figure A11** DNA sequencing of homology region (HR1) of UMP-CMP kinase (clone. HR1_U5)

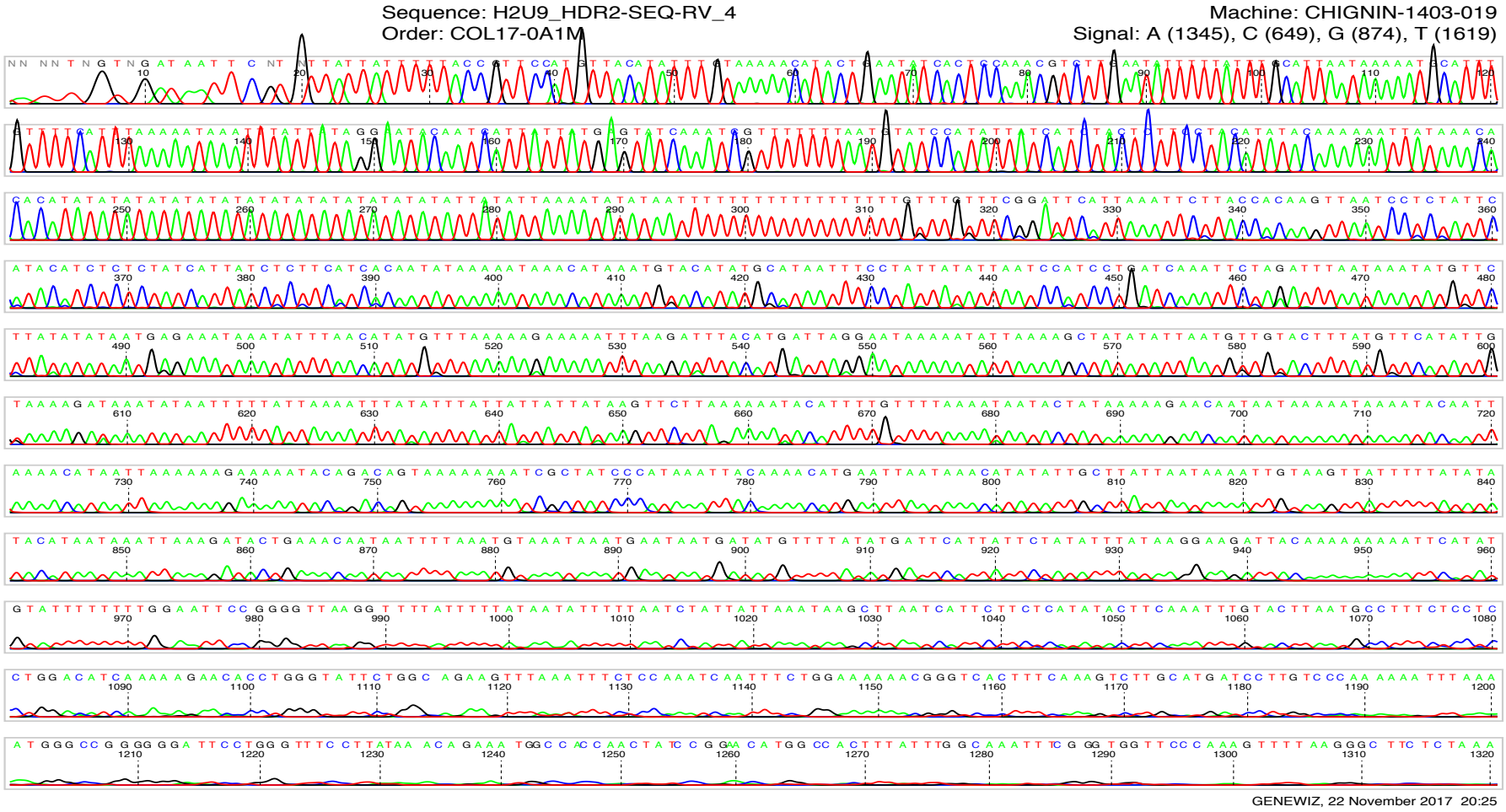



**Figure A12** DNA sequencing of homology region (HR2) of UMP-CMP kinase (clone. HR2_U9)

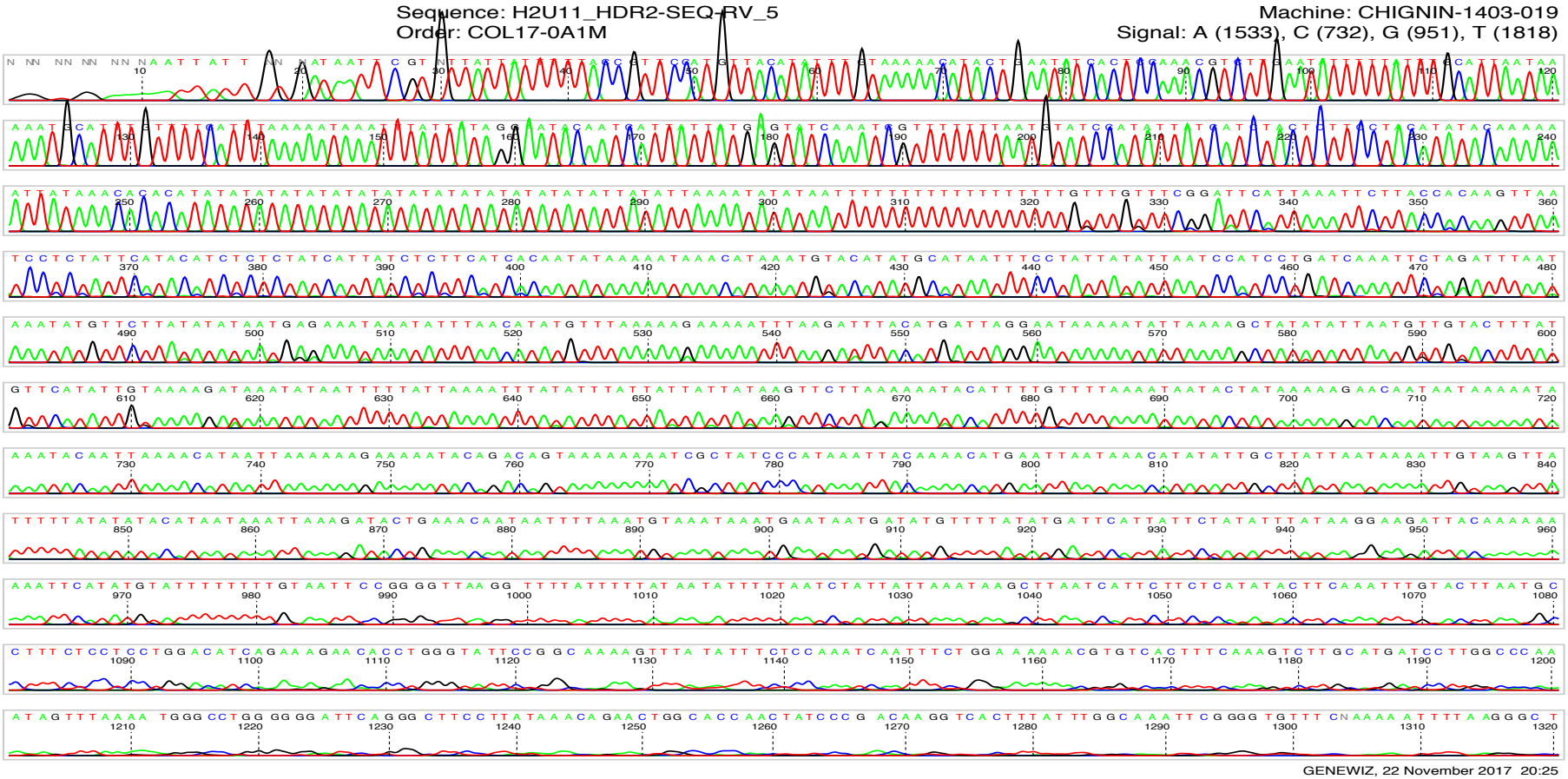



**Figure A13** DNA sequencing of homology region (HR2) of UMP-CMP kinase (clone. HR2_U11)

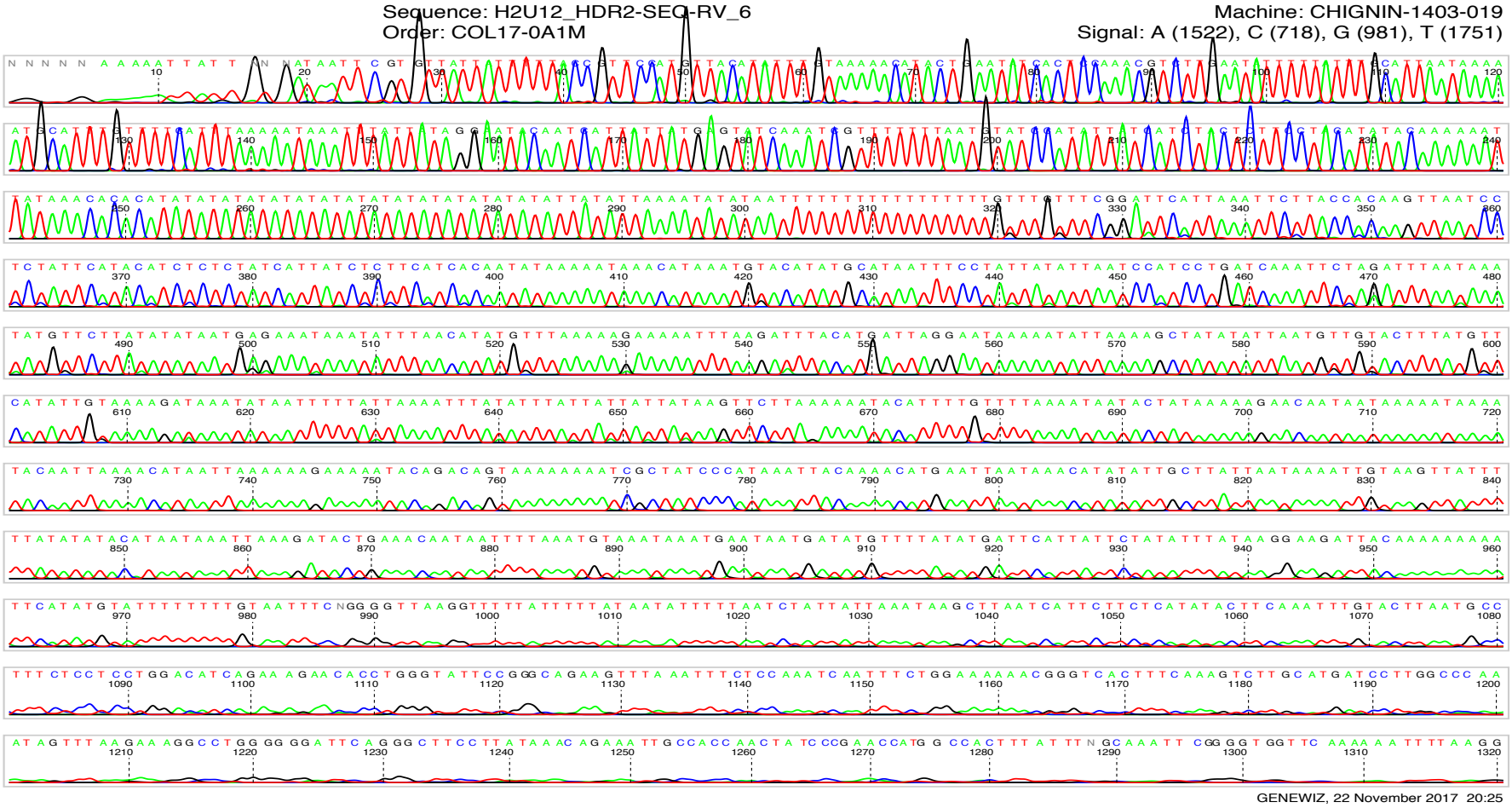



**Figure A14** DNA sequencing of homology region (HR2) of UMP-CMP kinase (clone. HR2_U12)

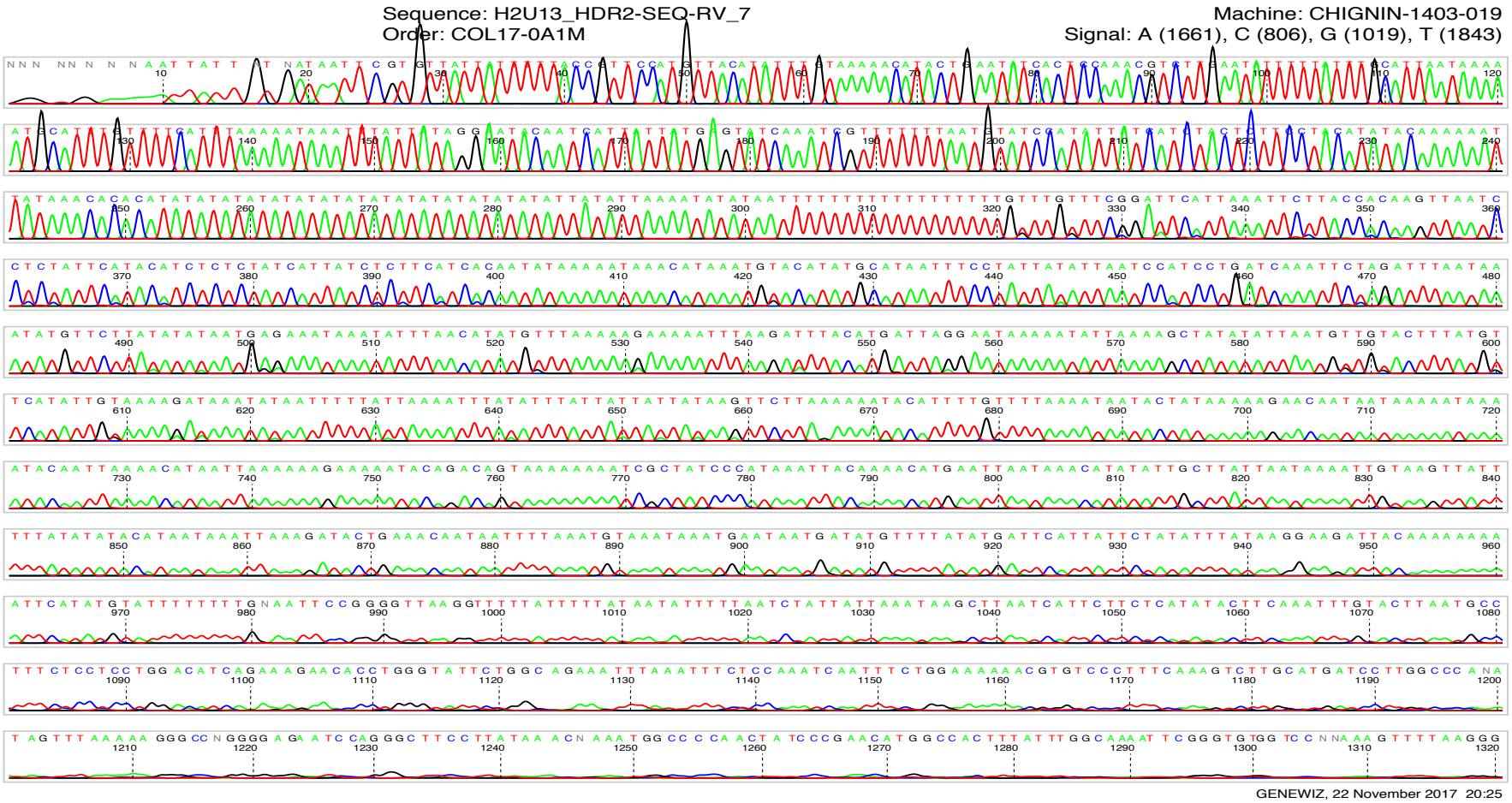



**Figure A15** DNA sequencing of homology region (HR2) of UMP-CMP kinase (clone. HR2_U13)

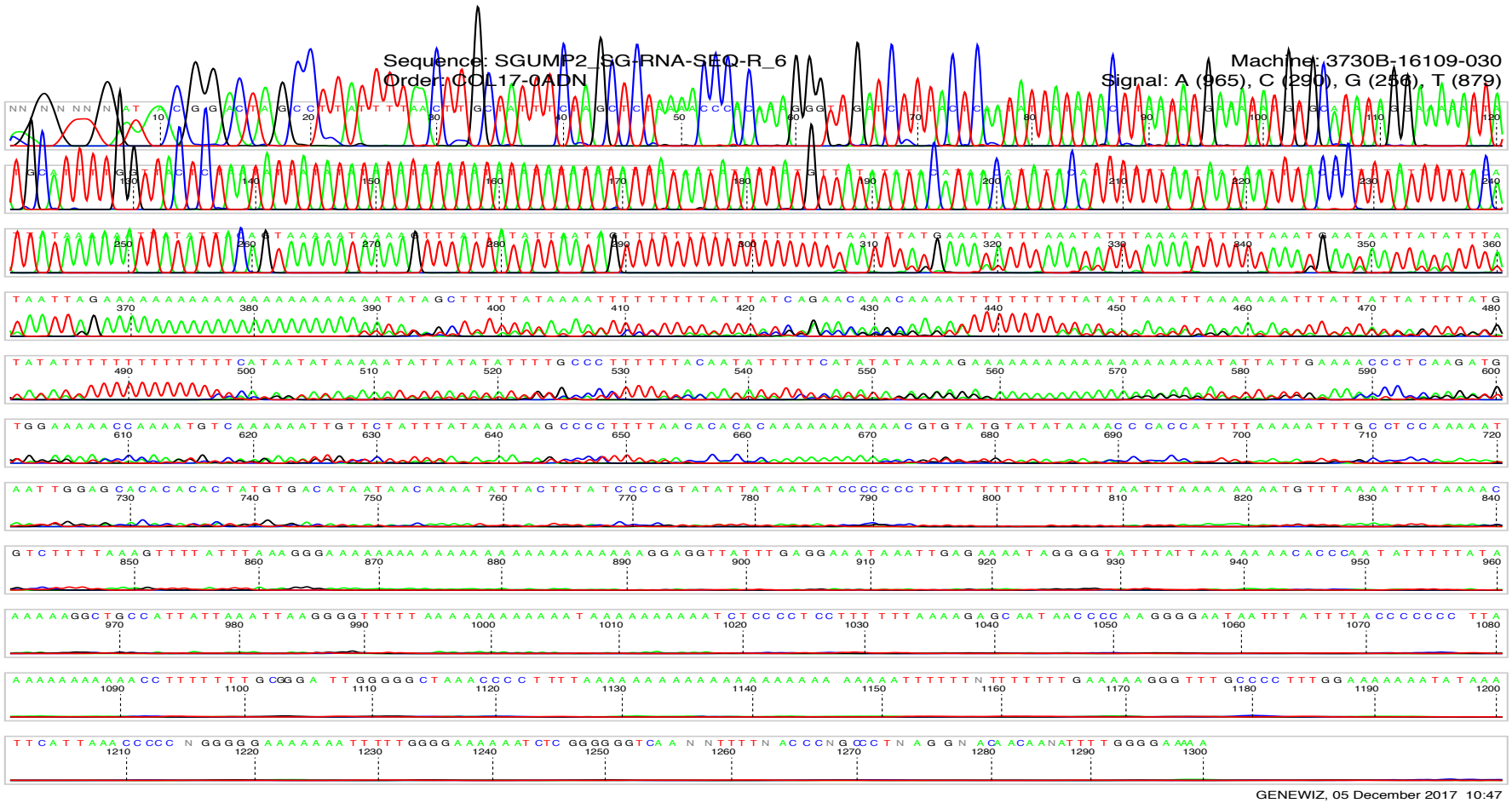



**Figure A16** DNA sequencing of sgRNA of UMP-CMP kinase (clone. sgUMP2)

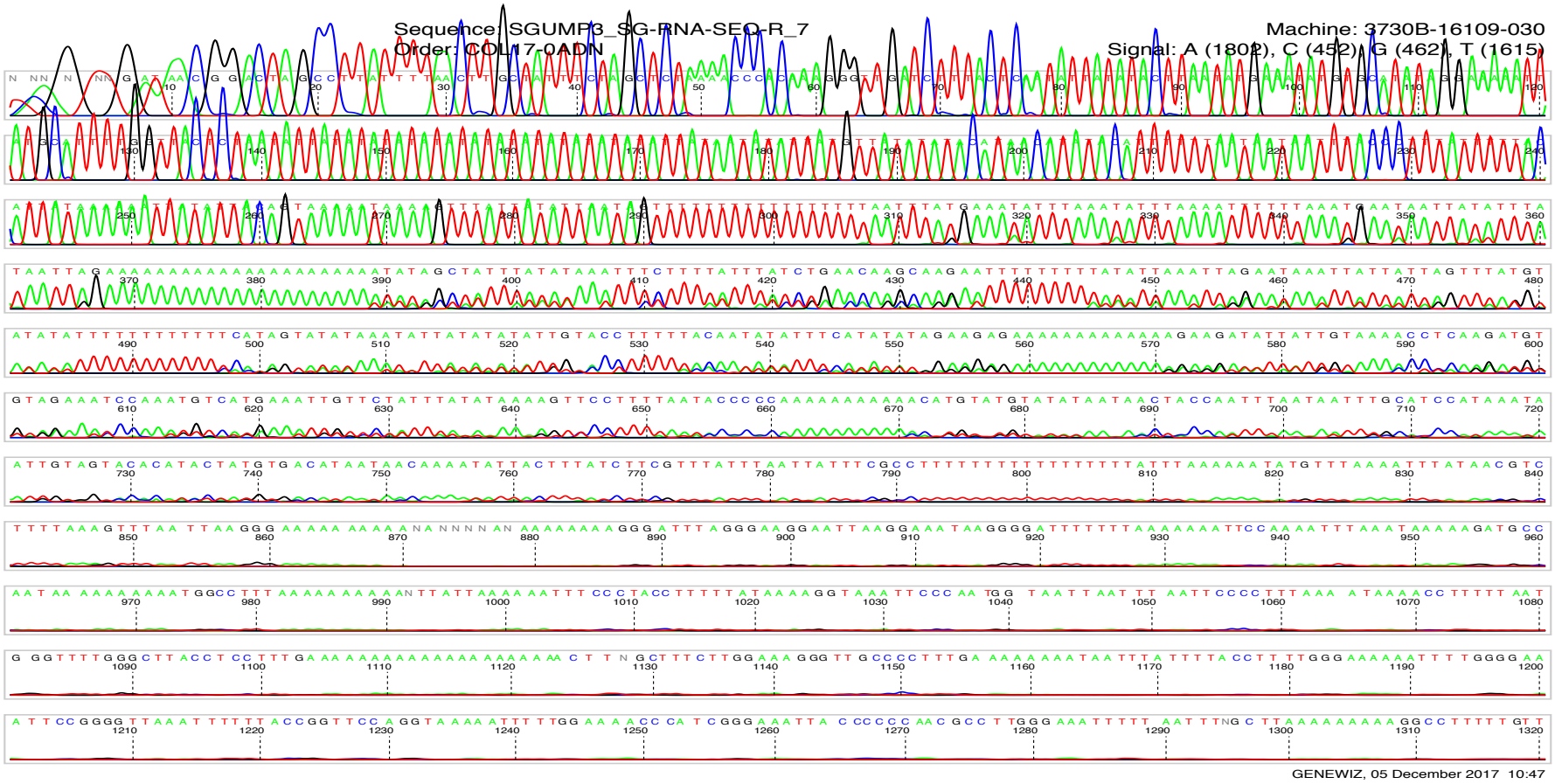



**Figure A17** DNA sequencing of sgRNA of UMP-CMP kinase (clone. sgUMP4)

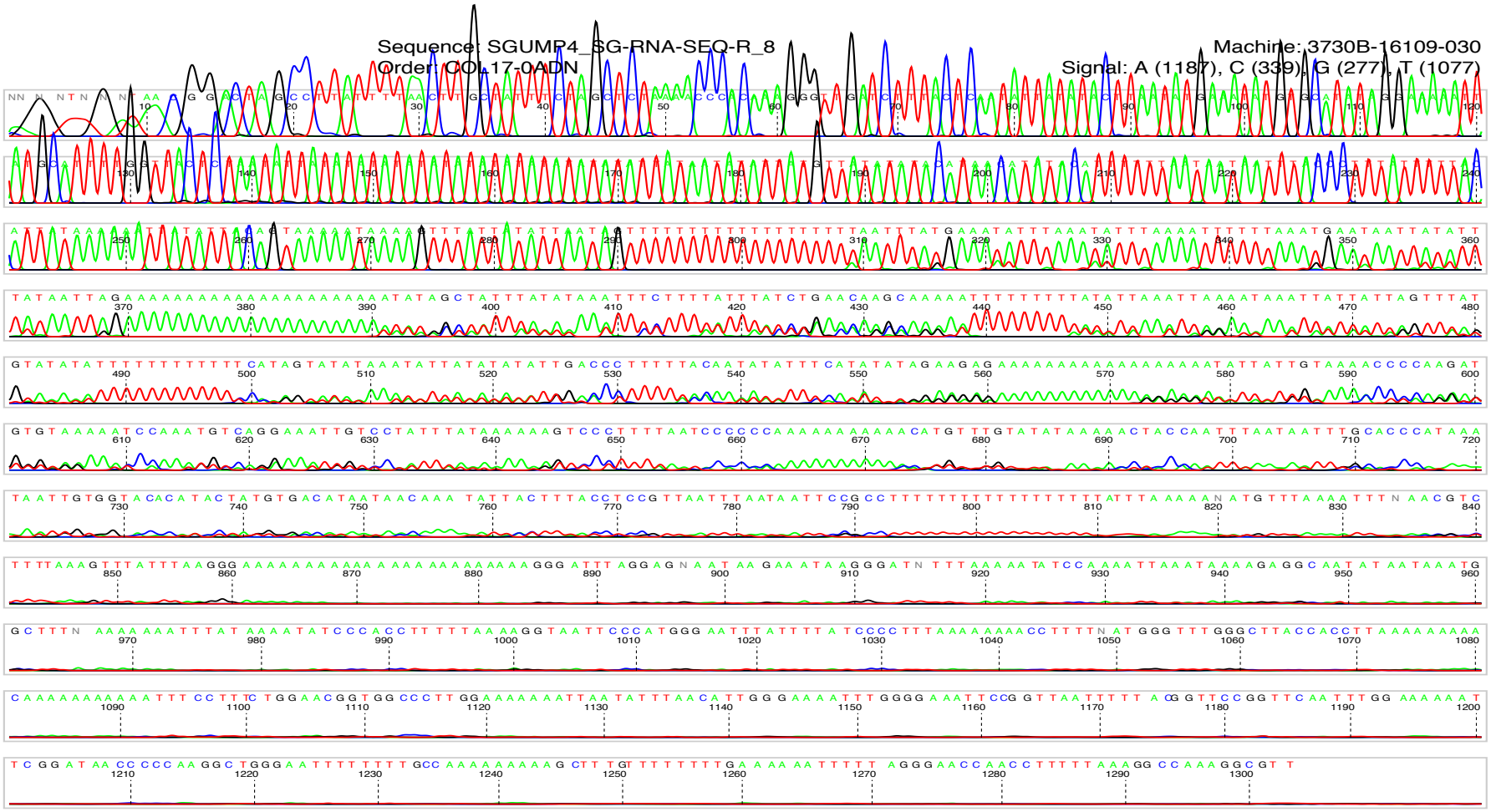



**Figure A18** DNA sequencing of sgRNA of UMP-CMP kinase (clone. sgUMP4)

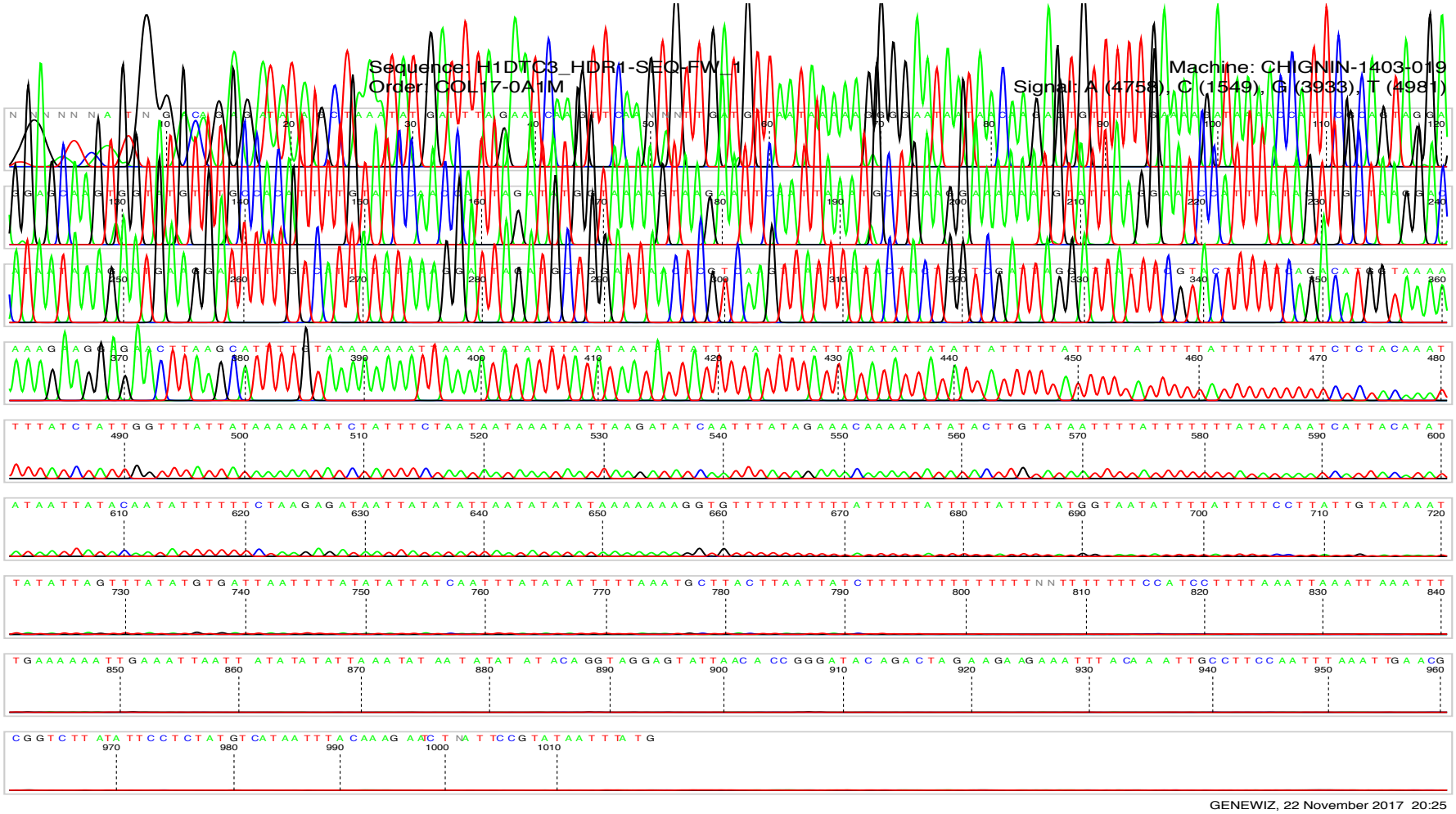



**Figure A19** DNA sequencing of homology region (HR1) of DTC (clone. HR1_DTC3)

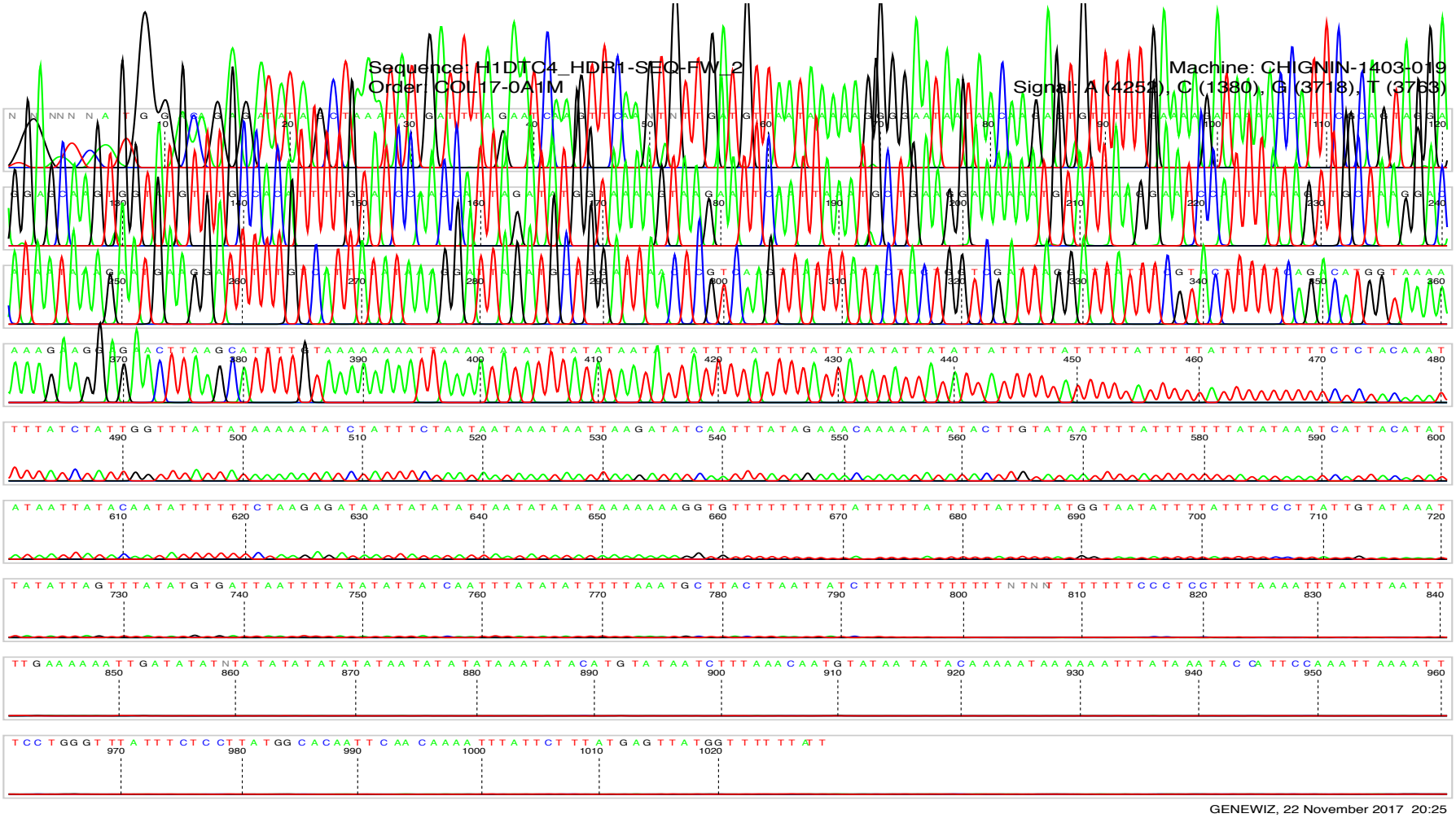



**Figure A20** DNA sequencing of homology region (HR1) of DTC (clone. HR1_DTC4)

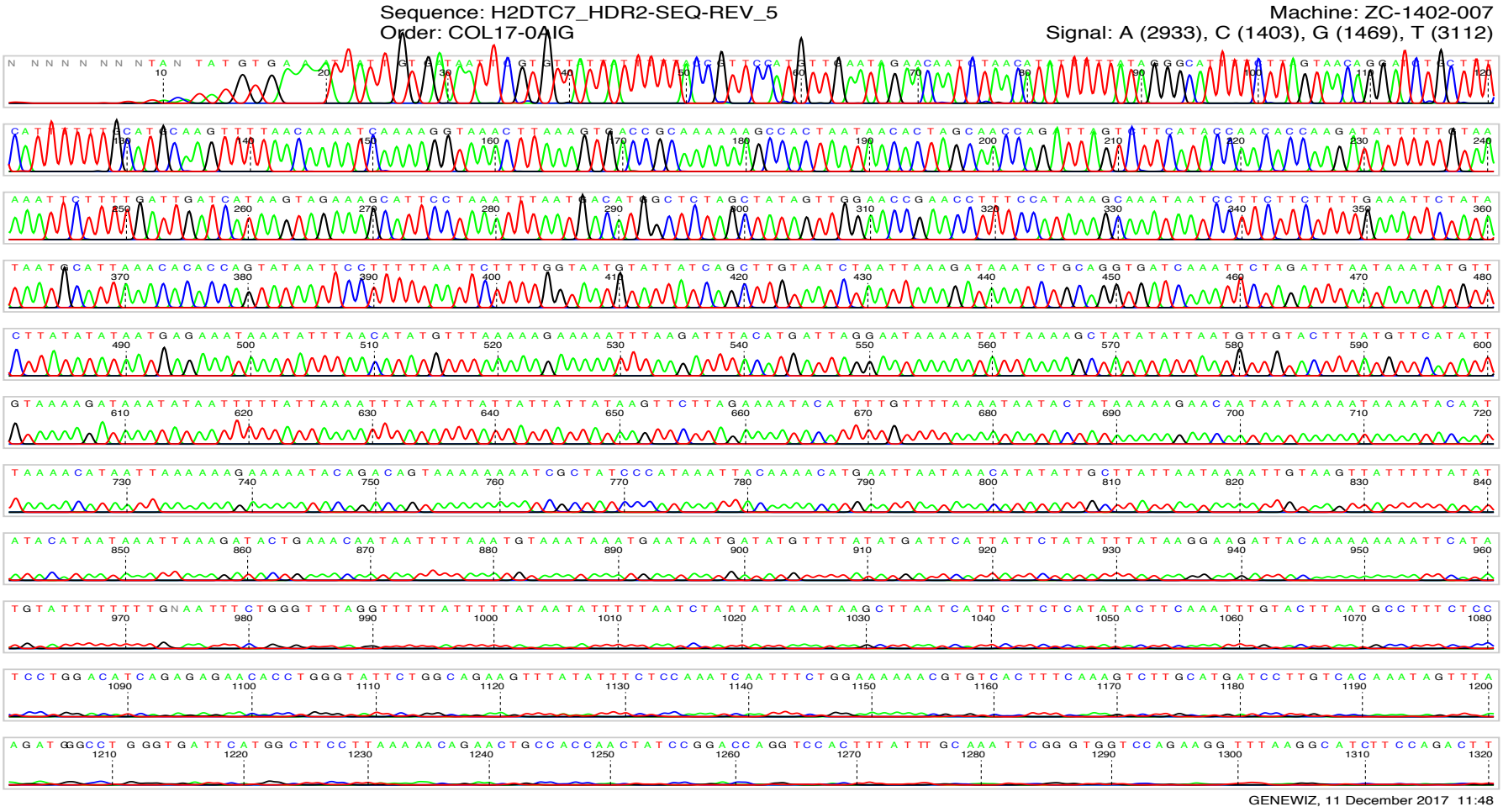



**Figure A21** DNA sequencing of homology region (HR2) of DTC (clone. HR2_DTC7)

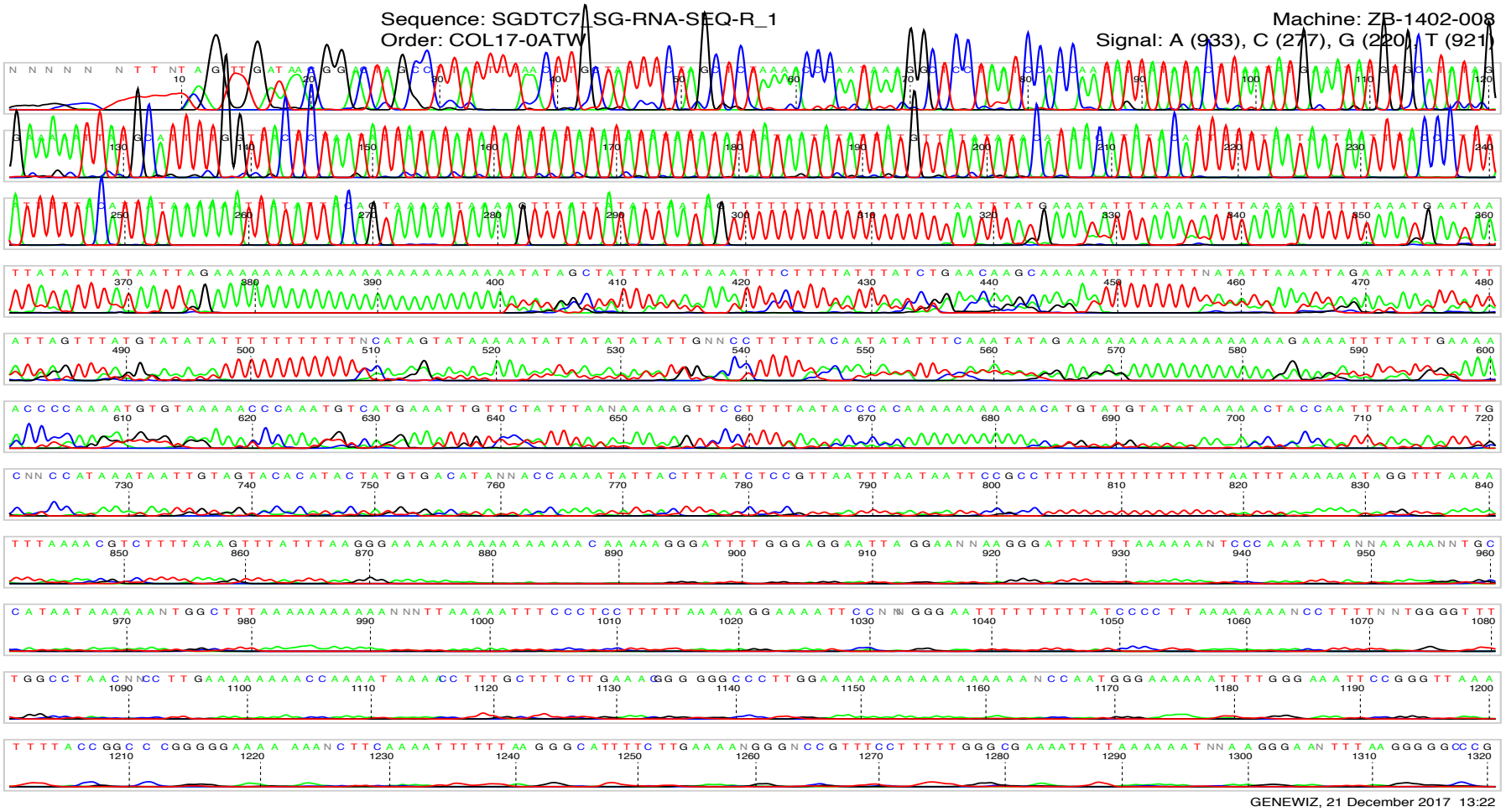



**Figure A22** DNA sequencing of sgRNA of DTC (clone. sgDTC7)

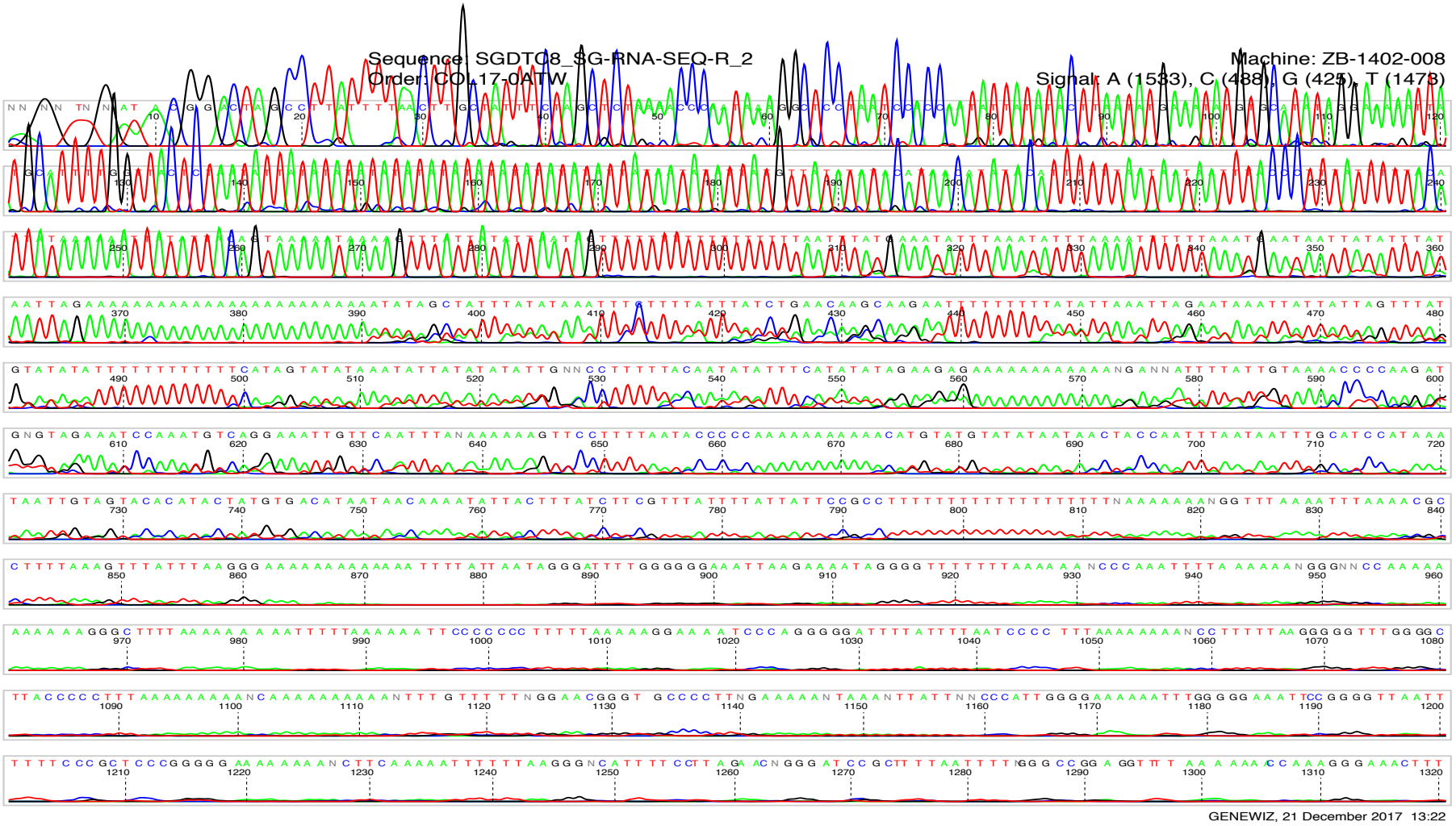



**Figure A23** DNA sequencing of sgRNA of DTC (clone. sgDTC8)

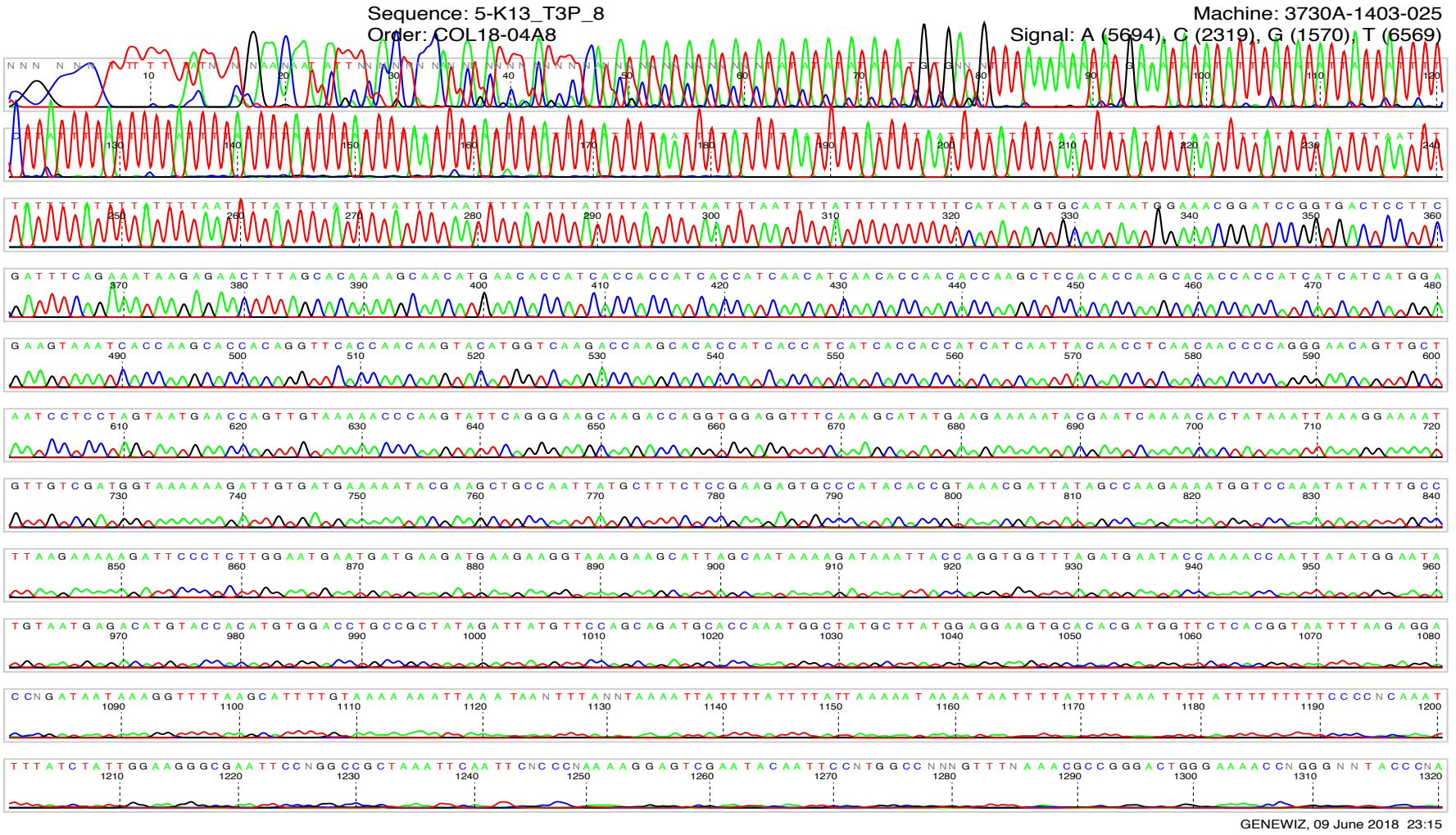



**Figure A24** DNA sequencing from genotyping of 5' integration of KAHRP2 transgenic parasites (clone. 5' K13)

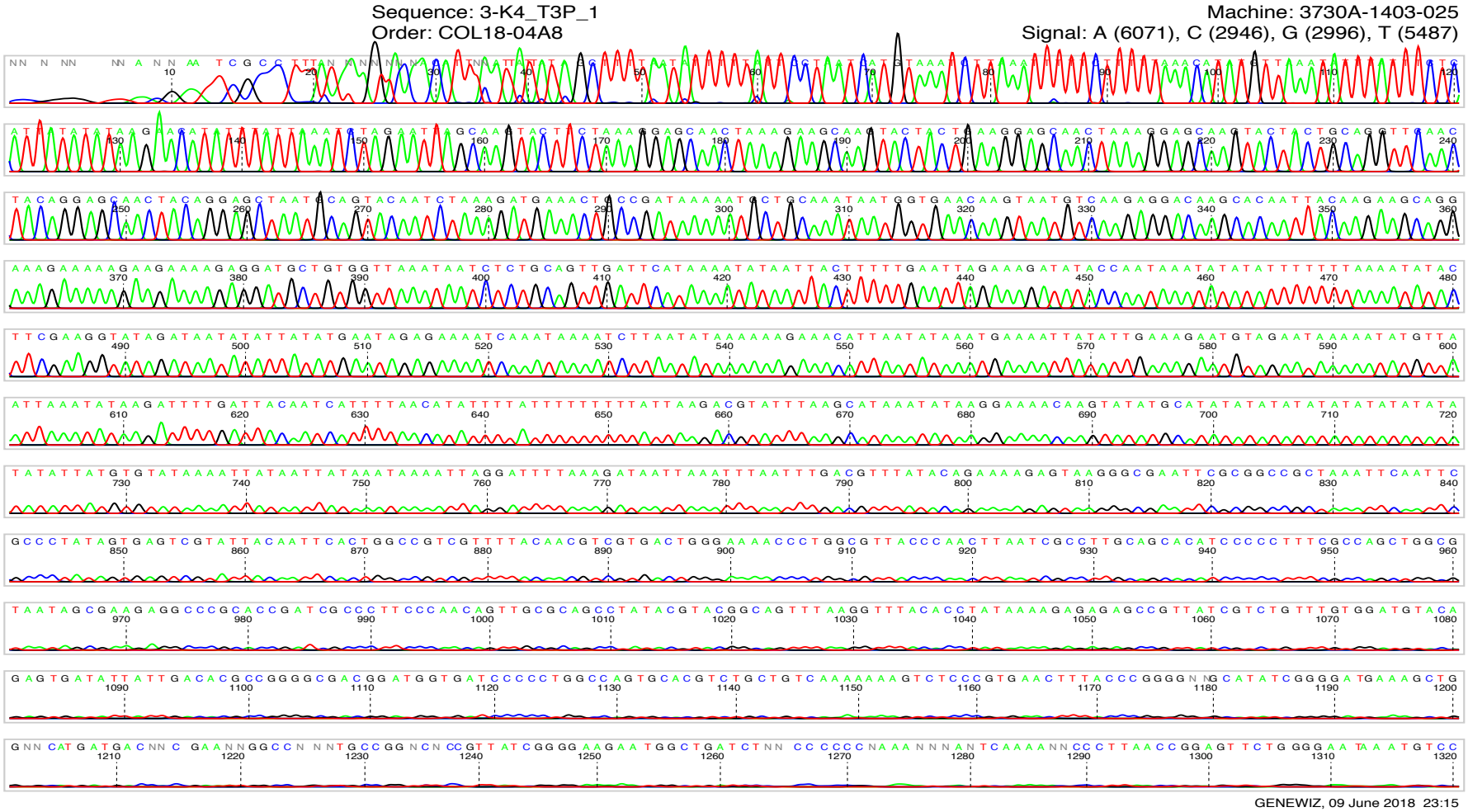



**Figure A25** DNA sequencing from genotyping of 3' integration of KAHRP2 transgenic parasites (clone. 3' K4)

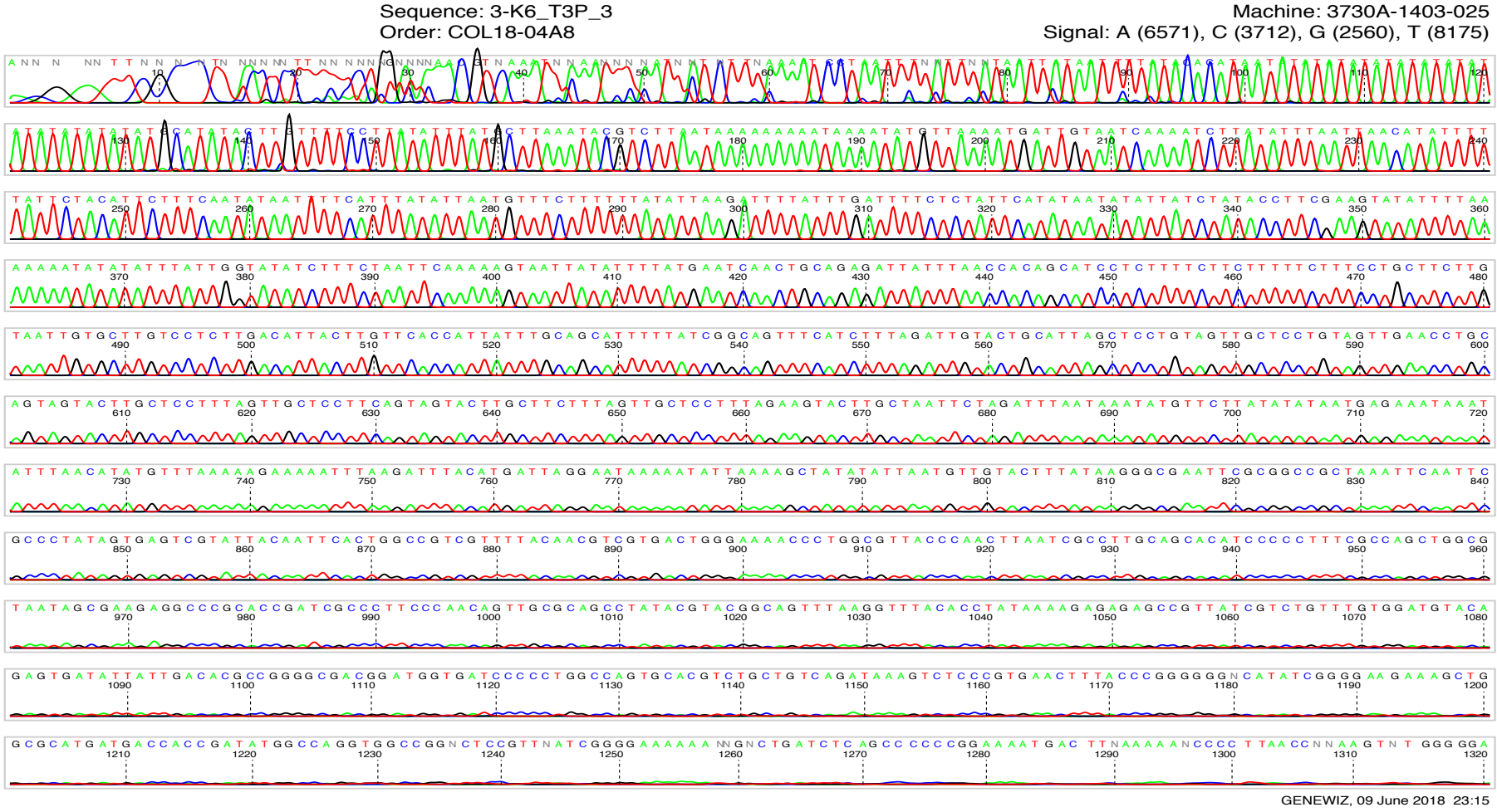



**Figure A26** DNA sequencing from genotyping of 3' integration of KAHRP2 transgenic parasites (clone. 3' K6)

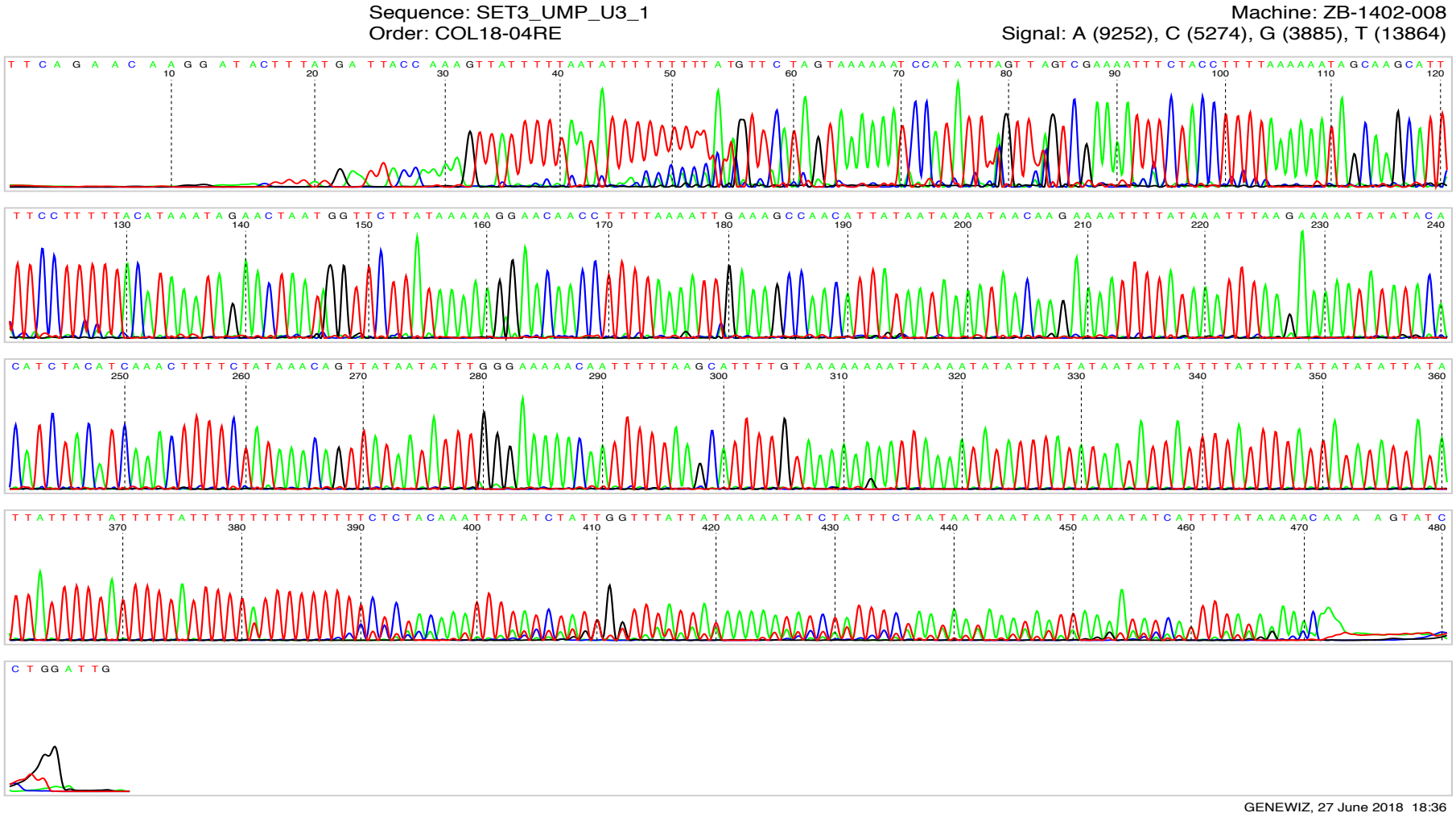



**Figure A27** DNA sequencing from genotyping of 5' integration of UMP-CMP kinase transgenic parasites (PCR product of 5' UMP)

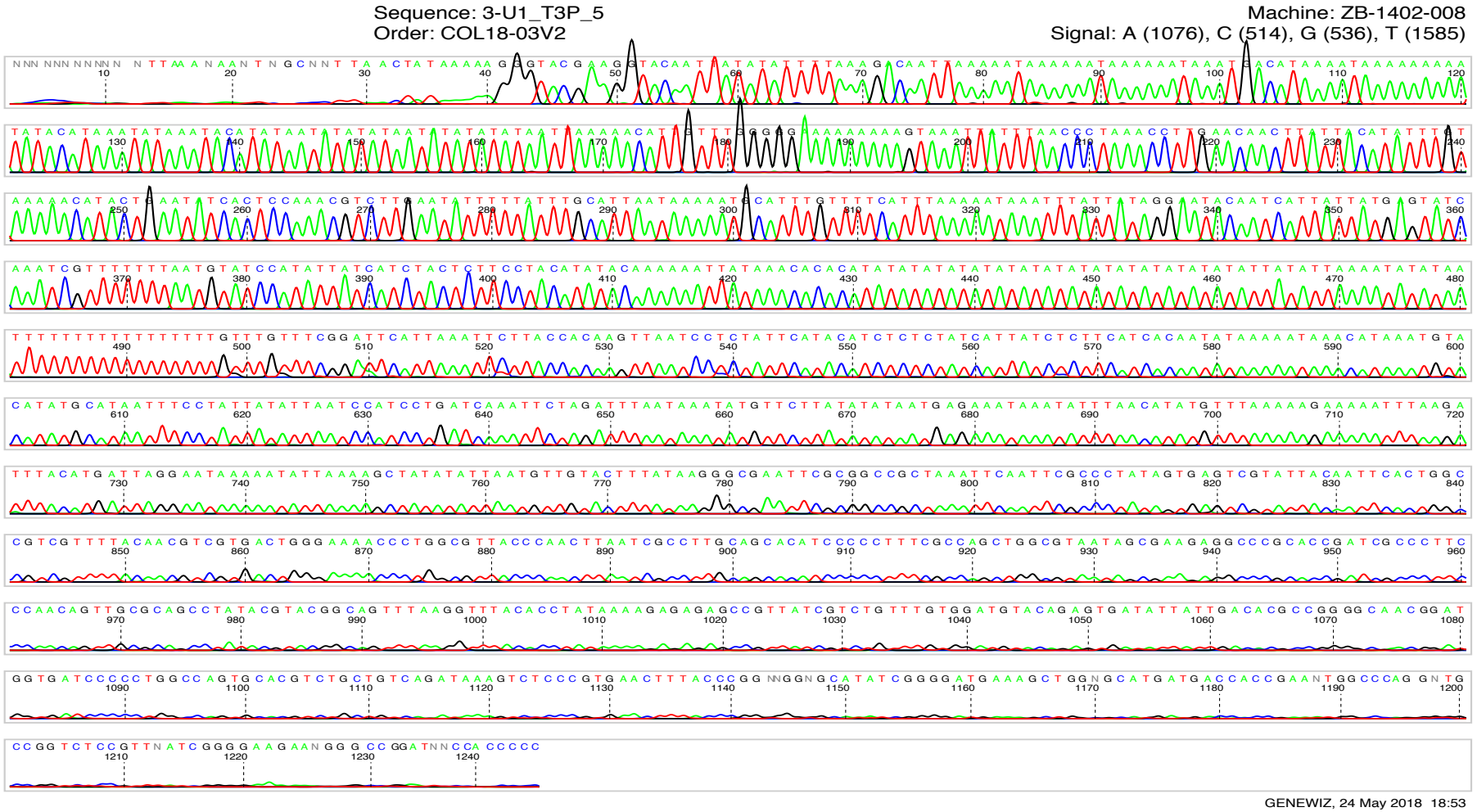



**Figure A28** DNA sequencing from genotyping of 3' integration of UMP-CMP kinase transgenic parasites (clone. 3' U1)

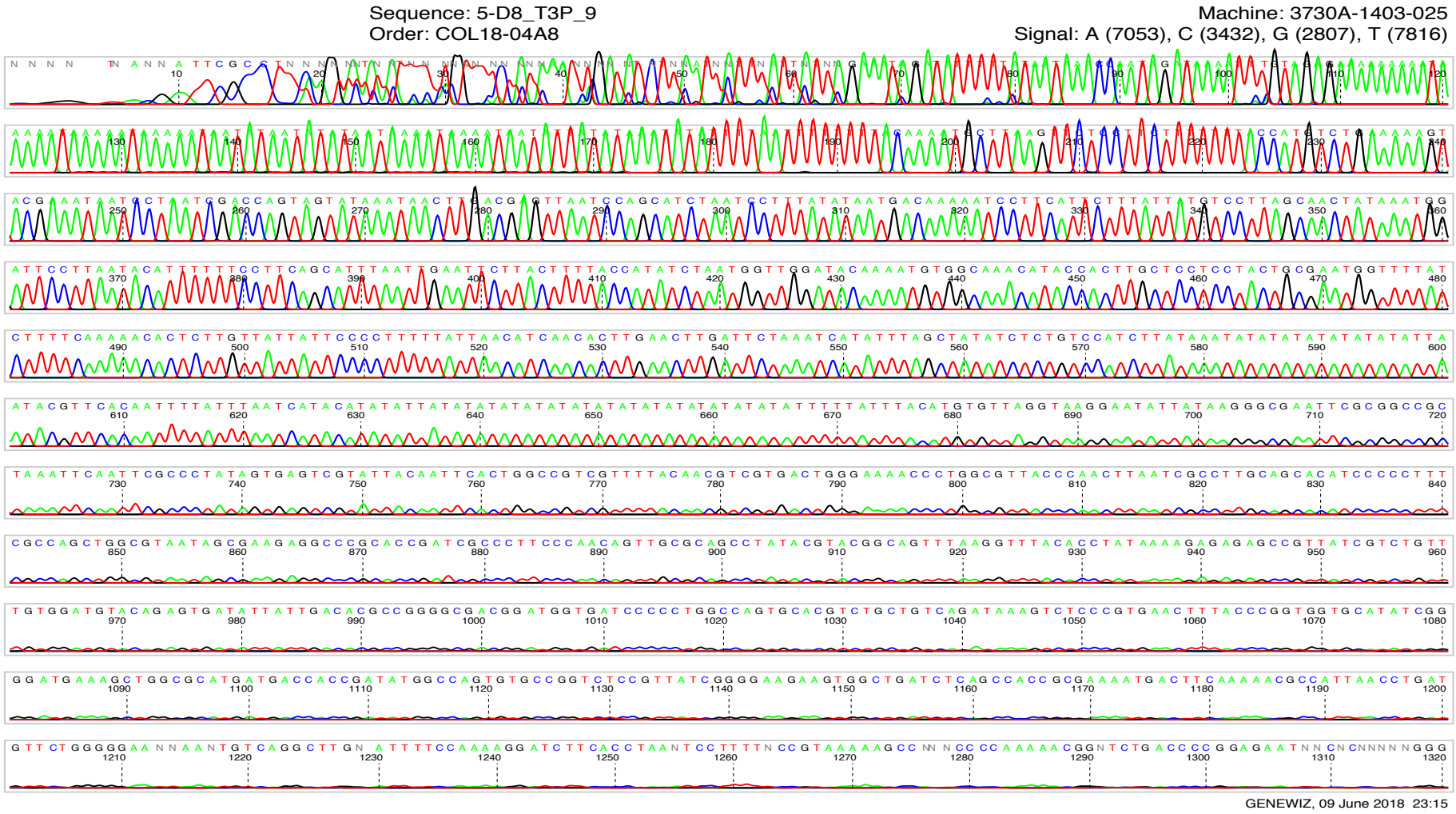



**Figure A29** DNA sequencing from genotyping of 5' integration of DTC transgenic parasites (clone. 5' D8)

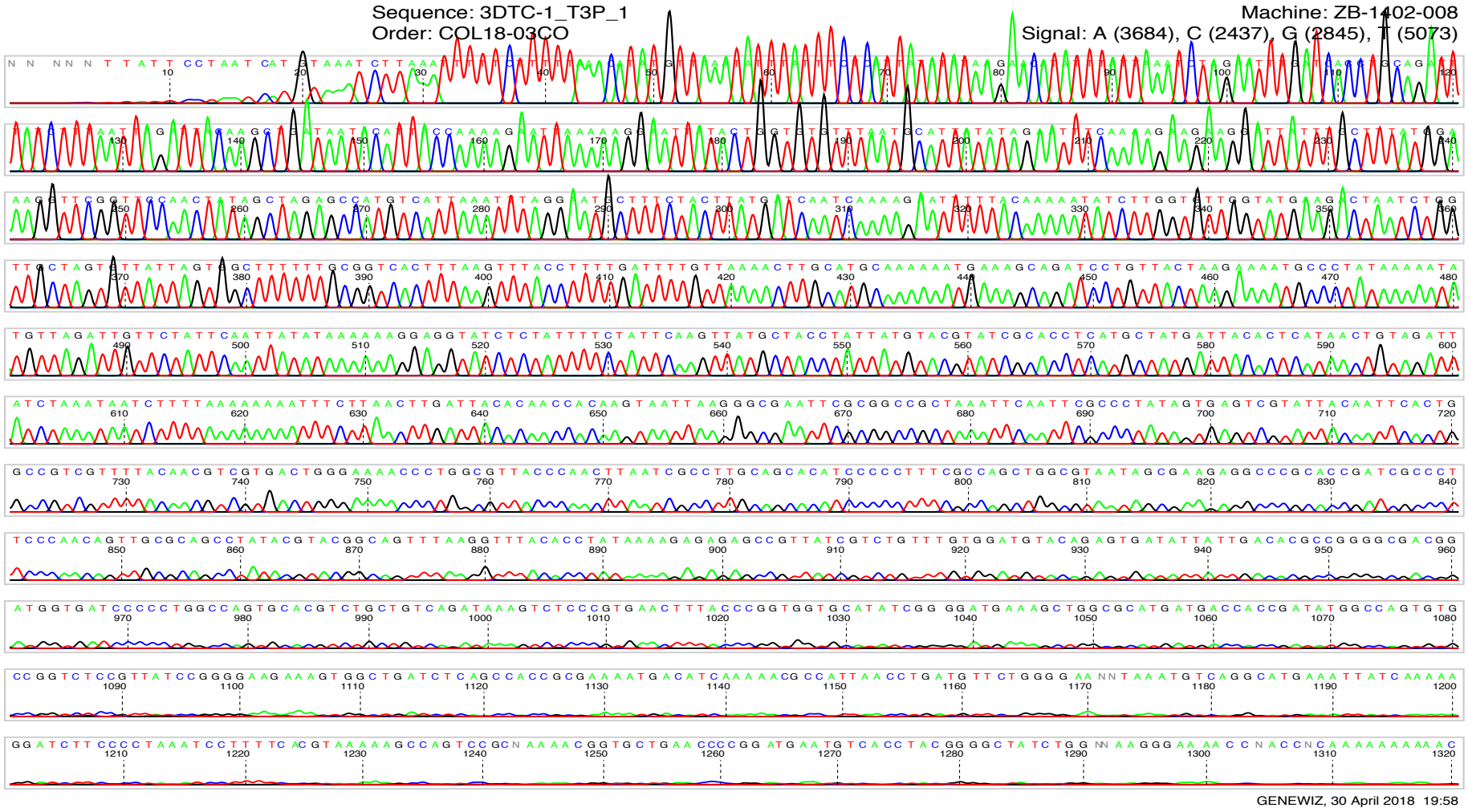



**Figure A30** DNA sequencing from genotyping of 3' integration of DTC transgenic parasites (clone. 3' DTC1)

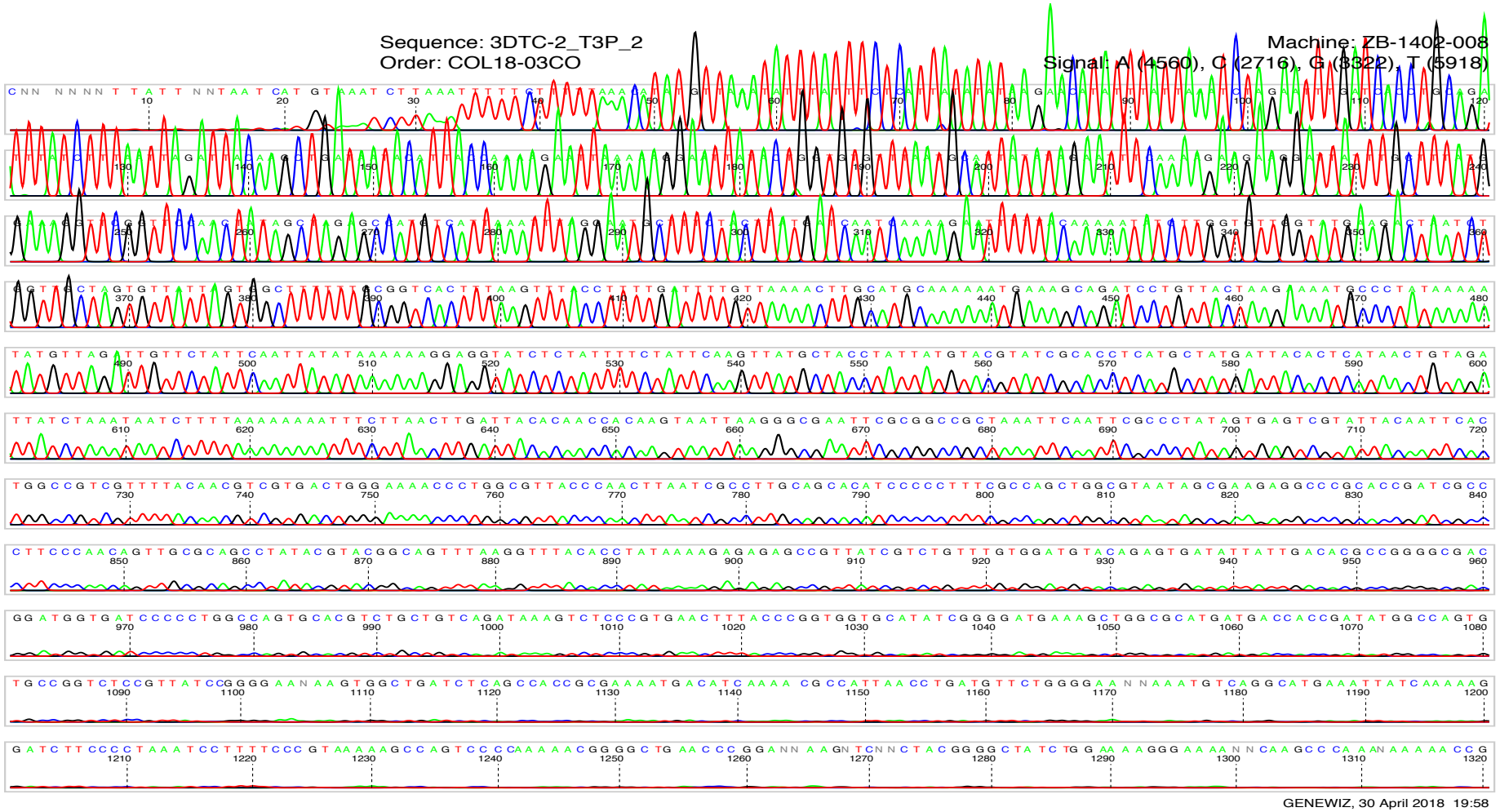



**Figure A31** DNA sequencing from genotyping of 3' integration of DTC transgenic parasites (clone. 3' DTC2)

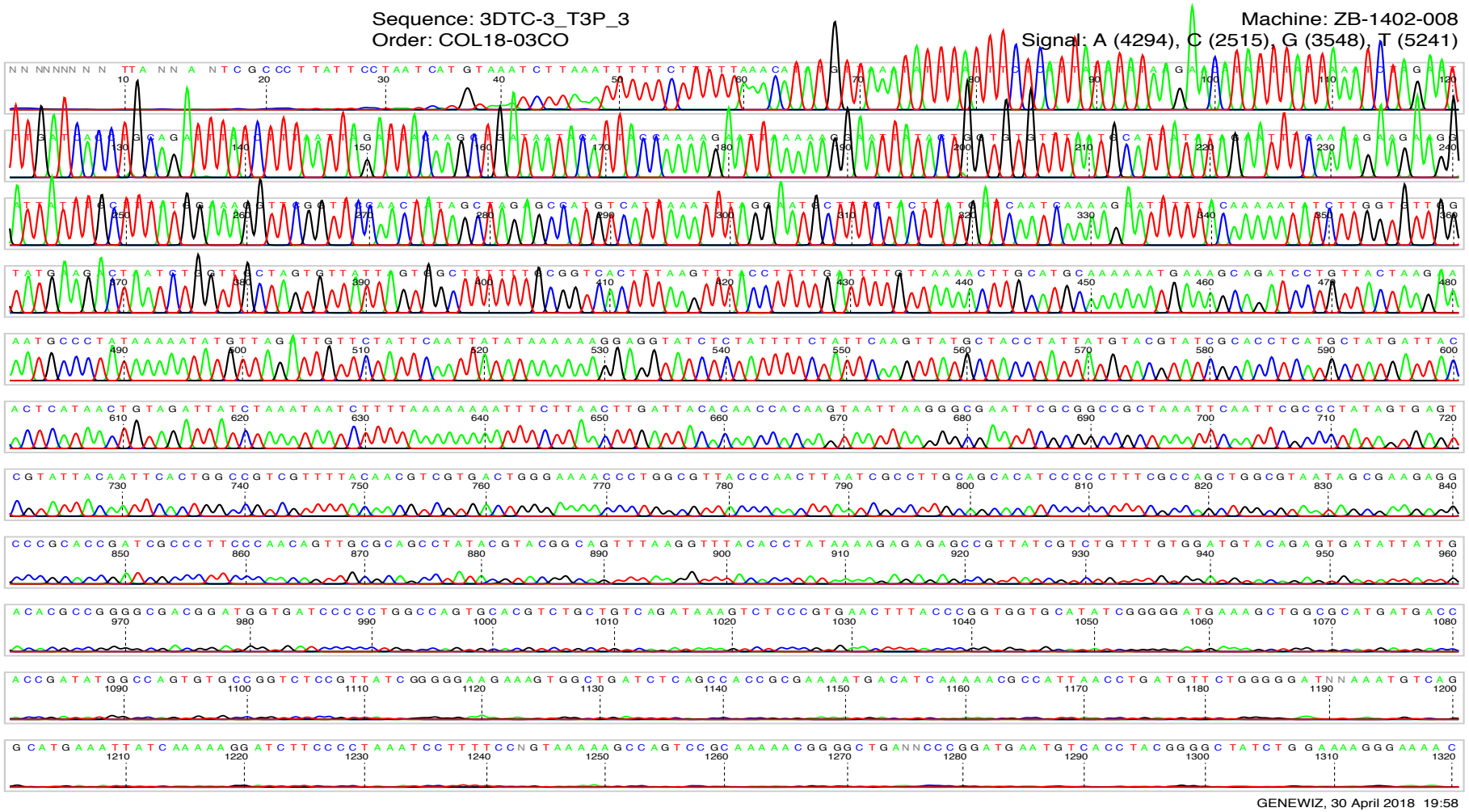



**Figure A32** DNA sequencing from genotyping of 3' integration of DTC transgenic parasites (clone. 3' DTC3)