UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

In vitro validation of chemo-transcriptomic fingerprints for the classification of the mode of action of antimalarial drugs

by

Natanya Venter

Submitted in partial fulfilment of the requirements for the degree

Magister Scientiae (Specialisation in Biochemistry)

In the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

(October 2024)

This work is based on research supported wholly by the National Research Foundation of South Africa. Any opinions expressed and conclusions arrived at in this work are those of the author, and the NRF accepts no liability for them.

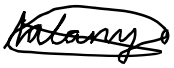
© University of Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Submission declaration

Declaration of originality:

I, Natanya Venter declare that the thesis/dissertation, which I hereby submit for the Degree *Magister Scientiae* in Biochemistry at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE: 

DATE: 31 October 2024

© University of Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Declaration of Originality

The Department of Biochemistry, Genetics and Microbiology, University of Pretoria, places great emphasis upon integrity and ethical conduct in the preparation of all written work submitted for academic evaluation. While academic staff teaches you about referencing techniques and how to avoid plagiarism, you too have a responsibility in this regard. If you are at any stage uncertain as to what is required, you should speak to your lecturer before any written work is submitted.

You are guilty of plagiarism if you copy something from another author's work (e.g. a book, an article or a website) without acknowledging the source and pass it off as your own. In effect you are stealing something that belongs to someone else. This is not only the case when you copy work word-for-word (verbatim), but also when you submit someone else's work in a slightly altered form (paraphrase) or use a line of argument without acknowledging it. You are not allowed to use work previously produced by another student. You are also not allowed to let anybody copy your work with the intention of passing if off as his/her work.

Students who commit plagiarism will not be given any credit for plagiarised work. The matter may also be referred to the Disciplinary Committee (Students) for a ruling. Plagiarism is regarded as a serious contravention of the University's rules and can lead to expulsion from the University.

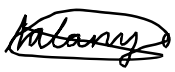
The declaration which follows must accompany all written work submitted while you are a student of the Department of Biochemistry, Genetics and Microbiology, University of Pretoria. No written work will be accepted unless the declaration has been completed and attached.

1. I understand what plagiarism is and am aware of the University's policy in this regard.
2. I declare that this dissertation is my own original work. Where other people's work has been used (either from a printed source, Internet, or any other source), this has been properly
3. acknowledged and referenced in accordance with departmental requirements.
4. I have not used work previously produced by another student or any other person to hand in as my own.
5. I have not allowed and will not allow anyone to copy my work with the intention of passing it off as his or her own work.

Full name of student: Natanya Venter

Student number: 19010126

Topic of work: In vitro validation of chemo-transcriptomic fingerprints for the classification of the mode of action of antimalarial drugs

Signature: 

Date: 31 October 2024

© University of Pretoria

Acknowledgements

First and foremost, I would like to thank our Heavenly Father for granting me the opportunity and resolve to pursue my purpose. He always has a much bigger plan for us than we had for ourselves.

To my whole family, it makes a huge difference to have someone believe in you, and all of you had faith in me throughout this process. You made a continuous effort to try and understand what I do and that means the world to me. Thank you to my mother, Henna, for all she is and all that she does for me. She has taught me what it is to work hard for what you have and what you want, raising me to be my own independent person. To my father, Frans, for sharing in my love of further education and showing me what it is to be a lifelong learner, it is an inspiration. You always support my pursuit of knowledge and encourage me to be curious.

To Suzelle and Tyrick, you guys are the friends everyone wishes they had. Thank you for making me feel understood and appreciated. You are what made every day bearable and changed tears into laughter. I will forever cherish the friendship we have, and I look forward to innumerable more memories and adventures to come. Thank you for always being willing to help, you have shown me what it is to have a true friend.

I would like to thank my supervisor, Prof Birkholtz, for being an inspiration and pushing me to be and do better. Thank you for helping me grow as a scientist as well as provide me with the opportunity to learn under your tutelage. Without you, I would not be here.

I would also like to thank M^{2}PL for the assistance and input they provided. Especially Dr Niemand, for putting up with all the random questions and dropping in as well as checking in to make sure I was doing okay.

I am blessed to have the most amazing people standing by my side and supporting me through it. The only thing I can say to them is thank you, and I am tremendously grateful to you.

Lastly, I would like to thank my funder, the NRF, for providing us with the funding to complete this research.

Someday when you get to where you are going, you'll look around and you will know it was you and the people who love you who put you there.

- Taylor Swift

© University of Pretoria

Table of Contents:

© University of Pretoria

© University of Pretoria

© University of Pretoria

List of Figures

© University of Pretoria

© University of Pretoria

List of Tables

© University of Pretoria

Table of Abbreviations

2D	Two-dimensional
ACT	Artemisinin combination therapy
ADME	Absorption, distribution, metabolism, and excretion
ASA-9	2-aminosuberic acid derivative
BLAST	Basic local alignment search tool
bp	Base pair
BSL2	Biosafety level 2
cDNA	Complementary DNA
CDPK	Calcium-dependent protein kinase
CETSA	Cellular thermal shift assay
CMap	Connectivity Map
CNV	Copy number variant
C_q	Quantification cycle
CRISPR	Clustered regularly interspaced short palindromic repeats
CRISPR/Cas9	CRISPR-associated endonuclease Cas9
DARTS	Drug Affinity Responsive Target Stability
DFMO	Difluoromethylornithine
DHA	Dihydroartemisinin
DHFR	Dihydrofolate reductase
DHODH	Dihydroorotate dehydrogenase
dNTP	Deoxynucleoside triphosphate
dsDNA	Double-stranded DNA
ECG	Endogenous control gene
EDC	Extraerythrocytic development cycle
EF2	Translation elongation factor 2
EtBr	Ethidium bromide
FC	Fold change
GOI	Gene of interest
gRNA	Guide RNA

© University of Pretoria

HDAC	Histone deacetylase
hpi	Hours post invasion
DHPS	Dihydropteroate synthase
hpt	Hours post-treatment
IC_{50}	Half maximal inhibitory concentration
IC_{90}	90 % of the maximum inhibition concentration
IDC	Intraerythrocytic development cycle
ITDR-CETSA	Isothermal dose response cellular thermal shift assay
IVIEWGA	In vitro evolution and whole genome analysis
MalDA	Malaria drug accelerator
MM	Molecular marker
MMV	Medicines for Malaria Venture
MoA	Mode of action
NTC	No-template control
ODC	Ornithine decarboxylase
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PCT	Parasite clearance time
PI4K	Phosphatidylinositol 4-kinase
PKG	cGMP-dependent protein kinase
PMSF	Phenylmethylsulfonyl fluoride
PRR	Parasite reduction ratio
PTM	Post-translation modification
PV	Parasitophorous vacuole
qPCR	Real-time quantitative polymerase chain reaction
SAHA	Suberoylanilide hydroxamic acid
SANBS	South African National Blood Service
SAR	Structure-activity relationship
SDS-PAGE	Sodium dodecyl-sulphate polyacrylamide gel electrophoresis
SNV	Single nucleotide variant

© University of Pretoria

SOM	Self-organising map
T_a	Annealing temperature
TAE	Tris-acetate-EDTA buffer
TCP	Target candidate profiles
TF	Transcription factor
T_m	Melting temperature
TPP	Target product profiles
TSA	Trichostatin A
UDG	Uracil-DNA Glycosylase
UHPLC-MS	Ultra-high-performance liquid chromatography-mass spectrometry
WHO	World Health Organisation

© University of Pretoria

Project summary

Malaria is an infectious disease brought on by *Plasmodium* parasites. Resistance development to current treatment measures is a significant challenge to the progress made in eradicating malaria. As a result, developing new drugs with novel targets and modes of action (MoA) is of utmost importance to address this resistance. Phenotypic screening is often used to identify active compounds, but it has a low success rate in identifying compound targets or MoAs. Direct and indirect methods can be employed to identify a compound's target or MoA. However, most of these methods are time-consuming and labour-intensive.

Machine learning techniques have not yet been widely used in antimalarial drug discovery to identify a compound's MoA. However, recently van Heerden *et al*. (2021) used classification-based machine learning to develop a rationally selected model capable of stratifying compounds into different MoA groups with a 77 % accuracy. This model used chemo-transcriptomic fingerprints to indicate that the variant expression of only 50 transcripts was sufficient to classify different compounds with similar MoA into the same subsets quickly and specifically. This classification could accelerate antimalarial drug discovery if only 50 transcripts can be used to evaluate the response of the parasite to any new antimalarial agent, thereby indicating MoA classification as either into a known class (with inherent increased risk of cross-resistance) or into an unknown, novel class. However, due to the predictive nature of machine learning models, the use of the 50 transcripts as 'biomarkers' for drug MoA must first be validated *in vitro* and their use in a streamlined platform must be evaluated before they can be used extensively for drug discovery purposes.

To address these needs, this study used real-time, quantitative PCR and the $2^{-\Delta\Delta Cq}$ relative quantification method to investigate the *in vitro* expression levels of the biomarkers that van Heerden *et al*. (2021) identified as responsive to compound treatment within parasite populations treated with antimalarial compounds. The qPCR-based platform was first evaluated using five control compounds, each with a known MoA. This established that the qPCR amplification of the biomarkers was sufficient to distinguish between the compounds and that compounds with specific targets clustered separately from those with non-specific targets. The platform was extended to four clinical candidates with dissimilar MoAs to evaluate if the biomarkers can distinguish between compounds not previously evaluated in the machine learning model. This solidified that the biomarkers could create distinctive chemo-transcriptomic fingerprints for each compound's MoA. Lastly, a clinical candidate that shares a target with a control compound was introduced, proving that compounds with overlapping biological activity showed similarities in their chemo-transcriptomic fingerprints. This data indicated that the biomarkers identified using machine learning could be predictive biomarkers for compound MoA classification. The limited number of biomarkers, as well as the established qPCR-based platform parameters in this study, provides a rapid and scalable means to determine a compound's MoA, therefore greatly benefiting antimalarial drug discovery by allowing drug candidates to be evaluated for unfavourable MoAs and ensuring that the MoA remains unchanged during lead optimisation.

© University of Pretoria

Chapter 1: Literature Review

1.1 The impact of malaria

Malaria is a serious infectious disease caused by a protozoan parasite of the *Plasmodium* genus that has a potentially lethal effect on the human population. In 2022, there were approximately 249 million malaria cases and over 608,000 deaths, as stated by the World Malaria Report 2023 (1). A large proportion of cases is attributed to the African region as defined by the World Health Organisation (WHO) (1). Malaria is considered a disease of poverty due to the severe economic burden caused by attempting to control malaria (2). Notably, the lack of sufficient resources to combat malaria in Africa contributes to the disease susceptibility of the continent (3).

Young children under the age of five are especially vulnerable due to their lack of acquired immunity, while the placentas of pregnant women are exploited by the parasite, leading to an increase in miscarriages (3). Malaria has varying severity and can be asymptomatic, uncomplicated, or severe (also referred to as complicated). Common symptoms of malaria include fever, sweating, chills, headaches, and symptoms of digestive upset. Additionally, severe malaria can lead to anaemia, organ damage, and coma due to parasite sequestration in the microvasculature (3, 4).

Plasmodium species that can infect humans include *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, Plasmodium ovale curtisi, Plasmodium ovale wallikeri*, *Plasmodium knowlesi* and more recently *Plasmodium cynomolgi* (2, 5). Of these species, the most severe infections are caused by *P. falciparum* parasites, which are most prevalent in sub-Saharan Africa (2). It is the most severe because of *P. falciparum* parasites' ability to adhere to microvasculature and evade the host's immune system (6).

1.2 The life cycle of *P. falciparum* parasites

P. falciparum parasites are transmitted by the female *Anopheles* mosquito, which is mostly attributed to the *Anopheles gambiae* complex in Africa (7). The parasites' life cycle spans various developmental stages within the mosquito vector and the human host (Figure 1) (8).

When an infected mosquito bites a human, sporozoites are injected into the host and enter the bloodstream (9). The extraerythrocytic development cycle (EDC), lasting 2-10 days, starts when the sporozoites reach the liver. The sporozoites traverse the liver sinusoids and invade hepatocytes, allowing the parasites to evade the host's immune system (8, 10, 11). The parasites establish a parasitophorous vacuole (PV) and undergo their first round of asexual replication (8). In other species of *Plasmodium*, i.e., *P. ovale* and *P. vivax*, non-replicating hypnozoites can be formed in place of schizonts, leading to a period of dormancy. This enables the parasites' lasting survival and can cause disease relapse (8). During the EDC, hepatic schizont maturation occurs, forming multinucleated exo-

© University of Pretoria

erythrocytic hepatic schizonts (also known as meronts) that lead to the production of thousands of hepatic merozoites (1N).

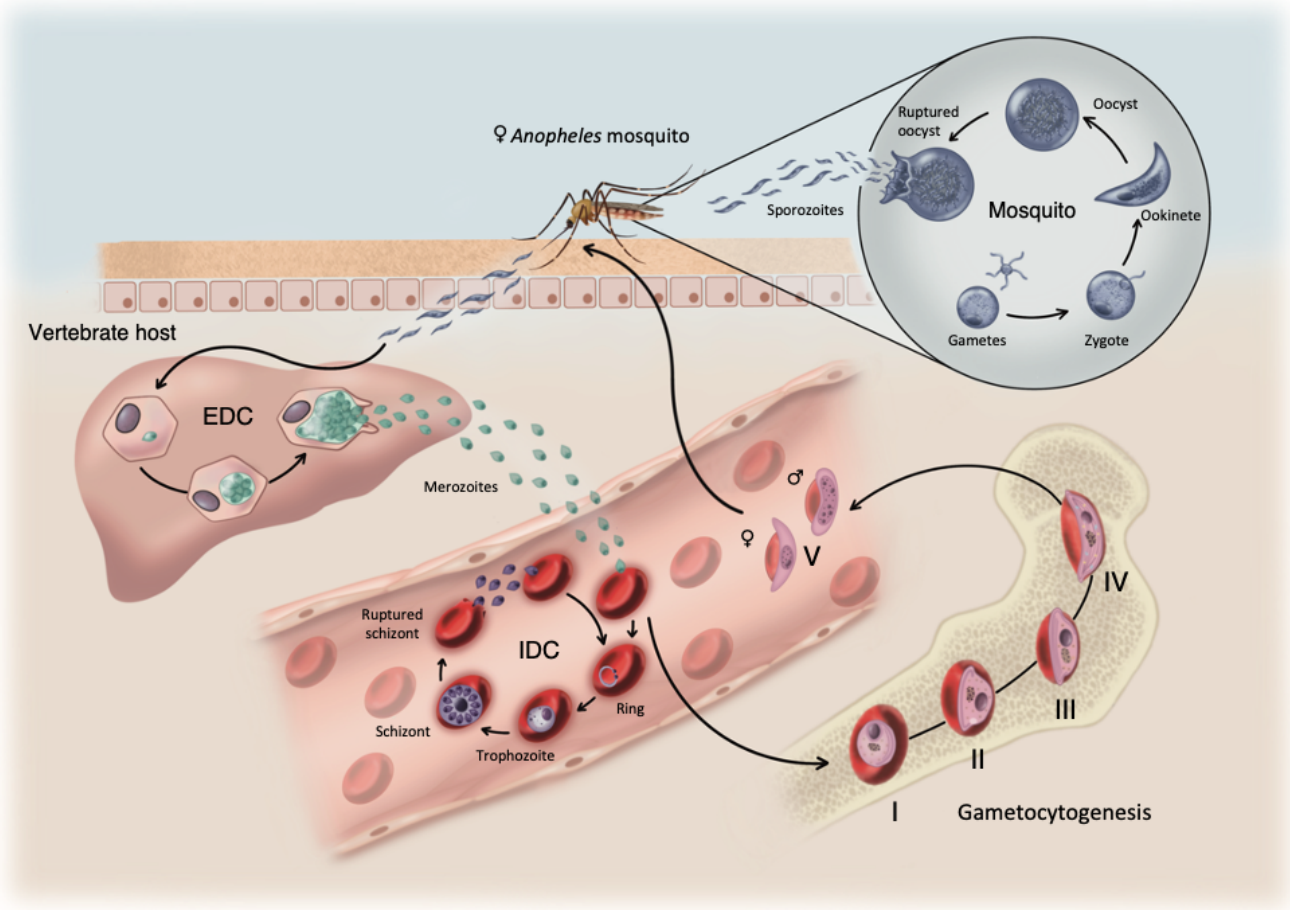


Figure 1. The life cycle of *P. falciparum* parasites. Shown is the progression of the parasite from the injection of sporozoites into the vertebrate host's bloodstream through each developmental stage until the mature gametocytes are again consumed by the mosquito for further development in the vector. This figure was created in Procreate (https://procreate.art) and modified from Connacher *et al.* (2022) (10). EDC = extraerythrocytic development cycle and IDC = intraerythrocytic development cycle.

When the hepatic schizonts burst, thousands of hepatic merozoites are released into the bloodstream, where they invade the host's erythrocytes (8, 10). This is the start of the 48 h intraerythrocytic development cycle (IDC), where malaria symptoms arise due to the rupture of erythrocytes at the end of each cycle (10, 12). *P. falciparum* parasites infect a large number of the host's erythrocytes, creating a high parasite burden and contributing to the severity of the disease (8). During the IDC, parasites undergo development in a PV that forms during the invasion of erythrocytes and supports a second round of asexual proliferation (13). Merozoites develop into ring-stage parasites that import haemoglobin via a cytostome. This is followed by the development of the parasites into trophozoites, wherein haemoglobin digestion occurs, which leads to the formation of a crystallised by-product called haemozoin. Haemoglobin digestion takes place in the parasite's digestive vacuole, primarily serving as a source of amino acids for protein biosynthesis in the parasite (14). Additionally, this creates space in the erythrocyte for the parasite to develop and grow. In mature trophozoites, DNA synthesis commences (1N to 2N) and the trophozoites develop into schizont-stage parasites (> 2N) by undergoing

© University of Pretoria

nuclear division (4, 13, 15, 16). During schizogony, multiple rounds of asynchronous DNA replication and nuclear division without cytokinesis take place, which form multinucleated schizonts. The last round of DNA replication and segregation, however, are synchronous, leading to the formation of individual daughter merozoites (1N) encapsulated in parasite plasma membranes and inner membrane complexes (17, 18). The mature schizonts then rupture and up to 32 merozoites are released into the bloodstream by a single schizont, where each can invade other erythrocytes, initiating the next cycle (8).

A small proportion of asexual parasites (< 10 %) commit stochastically to sexual differentiation during each cycle (10). *P. falciparum* parasites remain in this development stage for approximately 8-10 days while differentiating through five distinct stages (9). Similarly to asexual parasites, early gametocyte stages (I-III) digest haemoglobin and form haemozoin deposits. However, gametocytes do not undergo DNA replication and nuclear division. During stage I-IV, the gametocytes are not circulated in peripheral blood but rather instead sequester in the bone marrow and spleen of the host (10, 19, 20). Stage I gametocytes morphologically resemble trophozoites, however, as the parasites progress through gametocytogenesis, the distinctive crescent shape of *P. falciparum* parasites emerges. Stage II gametocytes become lemon-shaped, while stage III gametocytes start to elongate and become D-shaped. Stage IV gametocytes are elongated and thin, with male and female gametocytes becoming distinguishable. Female gametocytes usually have dense haemozoin distribution, while male gametocytes have dispersed haemozoin distribution (19, 21). After maturation, the stage V male and female gametocytes are released back into the bloodstream, where they can remain for several days to increase their likelihood of being transmitted to the mosquito upon its next blood meal (8, 22, 23).

Gametogenesis takes place in the mosquito midgut after ingestion of mature stage V gametocytes and is triggered by a pH change, temperature drop, and exposure to xanthurenic acid that is present in the mosquito midgut. The male gametocytes undergo three swift rounds of DNA replication that result in an octoploid nucleus. Subsequently, eight flagella are assembled, which results in eight motile male microgametes that are released by exflagellation (6, 24). In contrast, female gametocytes develop into a single immotile macrogamete (20). Fertilisation of the macrogamete (1N) by the microgamete (1N) creates a diploid zygote, which matures into a motile ookinete (22). Ookinetes can traverse the mosquito midgut and form oocysts, where the parasites undergo a third round of asexual replication to form sporozoites. The sporozoites undergo maturation and are released into the haemolymph, from where they can migrate to the salivary glands of the mosquito (8). The mosquito's next blood meal allows the transmission of the sporozoites to a new host (8, 25).

1.3 Prevention and treatment

The complex life cycle of *P. falciparum* parasites, which spans many distinct developmental stages, provides many opportunities for intervention to prevent and treat malaria.

© University of Pretoria

1.3.1 Prevention of malaria

The first form of defence against malaria is through vector control, which targets the *Anopheles* mosquito. *An. gambiae* favours human blood to complete its egg production cycle and mainly bites and rests indoors (26). Bite prevention can be an effective means of disease control, and this is primarily achieved by using insecticide-treated nets, long-lasting insecticide-treated nets, and indoor residual spraying (27). The primary chemical component that makes these measures lethal is the mosquito-repellent pyrethroid, which is not harmful to mammals and stays in the environment (26, 28). The lingering effect of the pyrethroid in the environment, however, results in mosquitoes' behavioural changes and resistance to insecticides (29, 30).

Chemoprophylaxis is another useful measure to prevent infection, as these drugs are typically ingested before exposure to malaria (31). Chemoprophylaxis in malaria targets the parasites during the EDC and IDC stages (32). These drugs have shown efficacy in reducing morbidity and mortality and are often used by travellers, pregnant women, and children (33, 34). Common examples of chemoprophylactic drugs include mefloquine, primaquine, and sulfadoxine-pyrimethamine. However, these drugs are often accompanied by unpleasant side effects and resistance has been shown to develop (35). Chemoprophylaxis has not been widely used for individuals living in moderate/high transmission areas. This is due to the high cost of acquiring and delivering chemoprophylactic drugs to large populations, requiring multiple doses, potential accelerated drug resistance, and potentially compromised development of naturally acquired immunity (36).

Developing a vaccine is considered to be of great importance in the malaria community and will provide a means of long-term protection that is not possible using chemical intervention (28). The current frontrunners are the RTS,S/AS01 subunit vaccine and the newly introduced R21/Matrix-M vaccine. The RTS,S/AS01 vaccine was approved by the WHO in 2021 for use in children who have a high risk of exposure, with the rollout of this vaccine approved in 18 countries (1). The R21/Matrix-M vaccine became the second WHO-recommended vaccine for use in children in high-risk areas in 2023. This vaccine is meant to assist in the shortage of RTS,S/AS01 vaccine supply, as it has great potential for large-scale and low-cost manufacturing (1, 37). The R21/Matrix-M vaccine displays greater efficacy (77 % at 12 months after vaccination) compared to RTS,S/AS01 (56 % at 12 months after vaccination and 33 % at 18 months after vaccination) (37). Although the efficacy of these vaccines seems encouraging, malaria elimination still requires a multifaceted approach (38).

1.3.2 Treatment of malaria

1.3.2.1 Severe Malaria

Malaria chemotherapy targets the parasites to alleviate symptoms and prevent death. For severe cases of malaria, it is recommended to administer artesunate parenterally. Severe malaria is caused by the

© University of Pretoria

sequestration of mature parasites in erythrocytes in the vasculature, causing reduced blood flow to vital organs (39). Artesunate shows activity against circulating ring-stage parasites and has been proven to reduce the incidence of death by approximately a third in clinical trials (40). Parenteral administration of artesunate allows the rapid removal of parasites from the blood (50 % parasite clearance after 6-9 h) and can be administered even while the patient is in a coma, as with cerebral malaria (3, 34, 41). This pharmacodynamic advantage of artesunate kills the parasites before sequestration (39).

1.3.2.2 Uncomplicated malaria

Uncomplicated malaria cases caused by *P. falciparum* parasites are currently treated by artemisinin combination therapies (ACTs). ACTs comprise a derivative of artemisinin (dihydroartemisinin (DHA), artemether, and artesunate) and an accompanying drug, where the artemisinin component is responsible for the rapid clearance of parasites while the accompanying drug is responsible for clearing the remaining parasites. Artemether-lumefantrine is most commonly used and can even be administered during pregnancy (42). Artemisinin inhibits the parasite's metabolic processes (glycolysis, nucleic acid, and protein synthesis) and requires the endoperoxide bridge of the drug to be activated by haem derived from haemoglobin digestion (43, 44).

ACTs were considered to have high efficacy, however, resistance has emerged recently (34). Other antimalarials are available and the ones currently in use target diverse processes of the parasites across multiple life cycle stages, most commonly the IDC of the parasite that is responsible for malaria symptoms. These antimalarials include antifolates (e.g. pyrimethamine), atovaquone, antibiotics, chloroquine derivatives, and quinolone derivatives (e.g. mefloquine) (28). However, resistance has also threatened their efficacy as antimalarials.

1.3.2.3 Drug resistance

Antimalarial drug resistance has been a great impediment in the fight against malaria with many antimalarials used currently and historically eventually experiencing resistance development (45). Chloroquine, which inhibits haem detoxification, was the standard treatment for malaria as it displayed rapid action, good tolerability, and simple dosing. However, resistance to chloroquine developed soon after its implementation (~10 years) due to a single nucleotide polymorphism in the chloroquine resistance transporter gene (*pfcrt*) (46). Chloroquine was also prescribed for pregnant women in malaria-endemic areas as a chemoprophylactic treatment against *P. falciparum* malaria, however, with the spread of chloroquine resistance, it was replaced with sulphadoxine-pyrimethamine in some African countries (39). Pyrimethamine is an antifolate that inhibits dihydrofolate reductase (DHFR) and is partnered with sulphadoxine, which inhibits dihydropteroate synthase (DHPS), affecting the parasite's folate biosynthesis pathway. A mutation of the *pfdhfr* gene conferred widespread pyrimethamine resistance, while a mutation of the *pfdhps* gene conferred resistance to sulphadoxine (39, 46, 47).

© University of Pretoria

The decreased efficacy of chloroquine and sulphadoxine-pyrimethamine led to their replacement by ACTs (39). However, the development of resistance against artemisinin and its derivatives, largely due to the non-compliance of patients, is a growing threat to the progress made against malaria (48). Mutations in the *kelch13* gene (encoding the PfK13 protein) confer partial resistance to artemisinin by permitting a portion of the treated early rings to survive, as the mutation leads to impaired haemoglobin uptake and, therefore, reduced artemisinin activation (49, 50). Clinically, partial artemisinin resistance is observed by delayed parasite clearance (parasites present after 72 h) (51). This delayed parasite clearance places greater pressure on the combination's partner drug, as a larger population of parasites that remain after the artemisinin component has been cleared leads to a greater potential for resistance mutant selection (39). Temporary measures to curb the effect of drug resistance include a prolonged course of treatment and evaluation of triple combination therapies (32). Triple combination therapies entail two slow-acting partner drugs to protect against resistance development and have proven to be safe and effective (39).

Despite effective antimalarials currently available to combat malaria, the potential for widespread resistance development, specifically against ACTs, remains a great threat. New antimalarials are under development, however, they will only become readily available within a few years.

1.4 Drug Discovery

The need for novel antimalarials drives antimalarial drug discovery to produce new drugs that can combat the emerging resistance while also considering the drug's safety (51). The Medicines for Malaria Venture (MMV) is a non-profit organisation that actively coordinates and drives the discovery and development of new antimalarials by partnering with experts in drug discovery and parasitology (https://www.mmv.org). Potential therapies are designed based on specific Target Product Profiles (TPPs) (32, 51).

1.4.1 Target Product Profiles and Target Candidate Profiles

A TPP refers to the final product that can contain two or more active compounds. TPPs fall into two categories: TPP1's intended use is case management, while TTP2's intended use is chemoprotection. Both consider different target candidate profiles (TCPs) (Figure 2). A TCP outlines the candidate's (a compound's) qualities (52). Some compounds can achieve multiple TCPs. For example, the compound can target hypnozoites (TCP3) and have transmission-blocking activity (TCP5), as with Primaquine (51).

© University of Pretoria

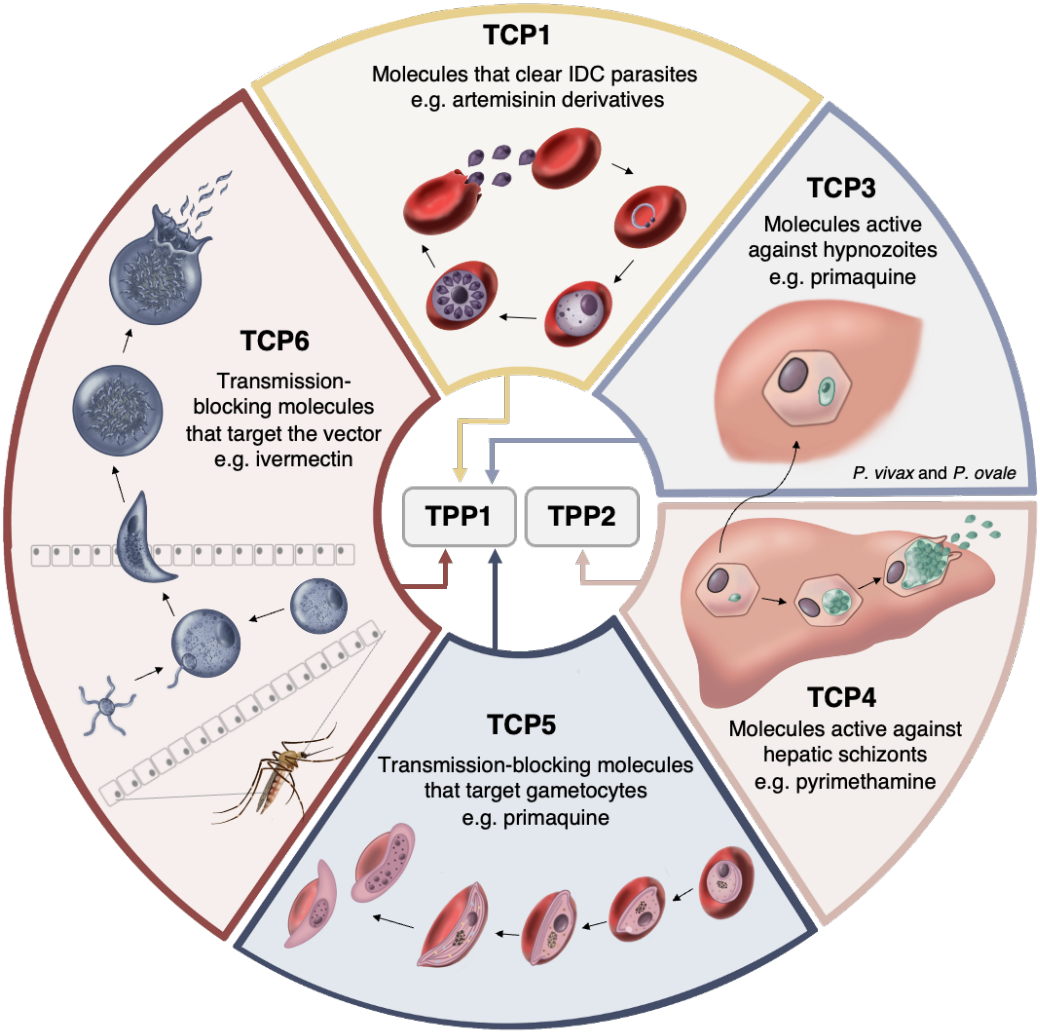


Figure 2. An outline of TPPs and TCPs. Two target product profiles (TPPs) that consider five different target candidate profiles (TCPs) are outlined in the figure. The information was adapted from the MMV (https://www.mmv.org) and Burrows *et al.* (2013) (52), and the figure was created using Procreate (https://procreate.art). IDC = intraerythrocytic development cycle.

TPP1, which concerns case management, usually considers compounds classified as TCP1, TCP3, and TCP5 (32). TCP1 would represent the basis of malaria treatment and indicate compounds that are responsible for the clearance of parasites in the IDC to alleviate malaria symptoms and prevent death (45, 51, 53). TCP2 is a retired profile that was used to consider antimalarials with sustained antiparasitic activity (45). Compounds classified as TCP3 would prevent relapse by targeting dormant liver-stage parasites (hypnozoites) in *P. vivax* and *P. ovale* infections (45, 51, 52). Compounds classified as TCP5 would prevent the development of gametocytes and block transmission. This goes hand in hand with compounds classified as TCP6 that would target the mosquito vector to decrease transmission and defend the uninfected population (53). Choosing which compounds to combine is dependent on their safety, resistance potential, capacity to be co-formulated, and efficacy (compounds could work synergistically or even antagonistically).

TPP2 concerns chemoprotection and is used by travellers entering malaria-endemic areas and occasionally during epidemic outbreaks. It mostly considers compounds classified as TCP4 and is at

© University of Pretoria

times accompanied by compounds classified as TCP1 (32, 51). Chemoprotection is accomplished by targeting sporozoites, hepatic schizonts, or hepatic merozoites. TCP4 is the key profile in TPP2 and has presented some challenges as different antimalarials must be used for chemoprotection and treatment of active malaria cases in the same region to avoid resistance development. The administration of these antimalarials also needs to be taken into account and should preferably only be administered once in an outbreak (52).

Identifying clinical candidates that fulfil these TPPs and TCPs requires continued investigative efforts. Once identified, these compounds move through the drug discovery pipeline to hopefully advance to preclinical and clinical development stages.

1.4.1.1 Preclinical and clinical candidates for malaria treatment

There are several antimalarial candidates currently under investigation that fit the clear criteria for development, including 1) their efficacy against multidrug-resistant parasites, 2) novel mode of action (MoA), and 3) having the ability to target multiple life cycle stages of the parasite. Particularly interesting is the number of antimalarial candidates with a broad range of target proteins as discussed below.

The aminopyridine compounds MMV390048 (MMV048) and its derivative UCT943 (both targeting phosphatidylinositol 4-kinase (PI4K) in the parasite) were promising candidates that were investigated for their antimalarial potential. PI4K is an attractive, novel drug target as these kinases play a key role in intracellular signalling by catalysing the conversion of phosphatidylinositol to phosphatidylinositol-4-phosphate (54). MMV048 was the first PI4K inhibitor to undergo clinical development and has been investigated in phase II clinical trials, where it showed multi-stage activity against the parasite (except against the hypnozoite stages). Parasites treated with MMV048 did, however, show recrudescence in a subset of participants (45, 55). MMV048 showed a good safety profile in initial investigations, however, embryo-foetal development studies in rats indicated teratogenicity (development of diaphragmatic hernias and cardiovascular malformations) (56).

M5717, a quinoline diamine, is another promising clinical candidate that showed multi-stage activity against EDC and IDC parasites, as well as transmission-blocking activity and activity against resistant strains (45). M5717 targets *P. falciparum* translation elongation factor 2 (EF2), therefore affecting protein synthesis as it is responsible for the translocation of mRNA and tRNA (57). In preclinical studies, M5717 showed good pharmacokinetic properties such as long plasma half-life, good oral availability, and minimal side effects, promoting its further investigation in phase II clinical trials (45).

DSM265, a triazolopyrimidine, was evaluated in phase II clinical trials. DSM265 shows multi-stage activity (EDC and IDC parasites) by inhibiting dihydroorotate dehydrogenase (DHODH) that catalyses pyrimidine nucleotide formation for nucleic acid synthesis. DSM265 is well tolerated, shows good safety in humans, and has a long half-life (58). However, DSM265's propensity for inducing resistance

© University of Pretoria

development made it a less favourable candidate as it showed the development of multiple mutations capable of conferring resistance (59, 60).

MMV030084 (MMV084), a trisubstituted imidazole that targets cGMP-dependent protein kinases (PKG) in *P. falciparum* parasites, is a promising potential candidate. MMV084 showed potent activity against the IDC of the parasite by inhibiting merozoite egress and the EDC by inhibiting hepatocyte invasion, and showed promising results for resistance selection, where it was denoted as resistance-refractory (61, 62).

To investigate parasite biological and, in turn, identify new druggable targets, tool compounds can be used as biological probes. These compounds are often effective in cancer or other parasitic diseases and are investigated for their potential activity against *P. falciparum* parasites. One such compound was difluoromethylornithine (DFMO), which was used as a lead compound along with other inhibitors of ornithine decarboxylase (ODC) to investigate polyamine metabolism as a drug target after the success of DFMO in cancer and African sleeping sickness management (63-65). DFMO did have activity against asexual *P. falciparum* parasites, however, it was found to be cytostatic and merely prevented the parasites from entering the schizont stage during the IDC. This inhibition was also reversible upon the addition of polyamines to the culture (63). Another compound that showed good efficacy in cancer is suberoylanilide hydroxamic acid (SAHA), which is a histone deacetylase (HDAC) inhibitor that has a key role in gene expression regulation (66). SAHA does show activity against *P. falciparum* parasites below the range that is considered clinically relevant for this compound and moderate selectivity towards the parasite (67). These compounds have not undergone clinical testing to be implemented as antimalarials, but they have provided insight into druggable targets in the parasite.

The discovery of new drugs with novel targets and MoAs is of utmost importance. These drugs must have novel targets to reduce the likelihood of pre-existing resistance affecting their efficacy (51). This holds particular importance for *P. falciparum* parasites, which have, to various extents, developed resistance against all currently implemented antimalarials and will undoubtedly do so again (28).

1.4.2 Antimalarial drug discovery

Diverse approaches to antimalarial drug discovery can encompass the phenotypic screening of compound libraries as well as the design of inhibitors for established druggable targets (target-based discovery). Phenotypic screening, however, has led to the discovery of most clinical antimalarial candidates currently under development (45) since it overcomes the challenge of compound uptake experienced because *P. falciparum* parasites are obligate intracellular parasites enclosed in three different membrane systems (68). Although target-based screening has identified strong protein inhibitors, the majority of these did not show any activity against the parasite simply because they were not able to enter the parasite to engage their target.

© University of Pretoria

Phenotypic screening takes place in biological systems that mimic physiological conditions and are unperturbed as far as possible, and often exploits a disease-related characteristic to screen large compound libraries for hit compound identification (69). It, therefore, provides an objective approach to identifying compounds that have a desired effect or phenotype and has the additional advantage of considering cell permeability (51, 70, 71). The *in vitro* activity of compounds is assessed using whole-cell parasites, cultured to the desired life cycle stage, and often leads to identifying compounds that can target various proteins and signalling pathways (51). Examples of antimalarials currently in the drug discovery pipeline that were discovered using phenotypic screening include cipargamin (KAE609) and M5717, the latter of which retains activity against parasites resistant to current antimalarials (45).

Although highly successful in identifying compounds with clear activity against *Plasmodium* parasites during the hit identification phase, phenotypic screening is challenged since hits identified this way cannot progress as preclinical candidates in the drug discovery pipeline unless their target is known (53, 70). This is for several reasons: 1) it would be almost impossible to motivate the progress of a compound with an unknown MoA in the possibility of pre-existing resistance in parasites if the compound acts on known drug targets, 2) different compounds targeting different proteins/MoAs will be combined to prevent resistance development, and 3) not knowing the target or the MoA of the compounds makes rational improvement of the potency and safety of the compound challenging as it is unknown which part of the compound interacts with the target protein (51, 72). It is, therefore, critical that a compound's MoA or target be elucidated as early on as at the hit-to-lead phase of discovery. If successful, the compounds progress to toxicology evaluations and, eventually, clinical trials (Figure 3).

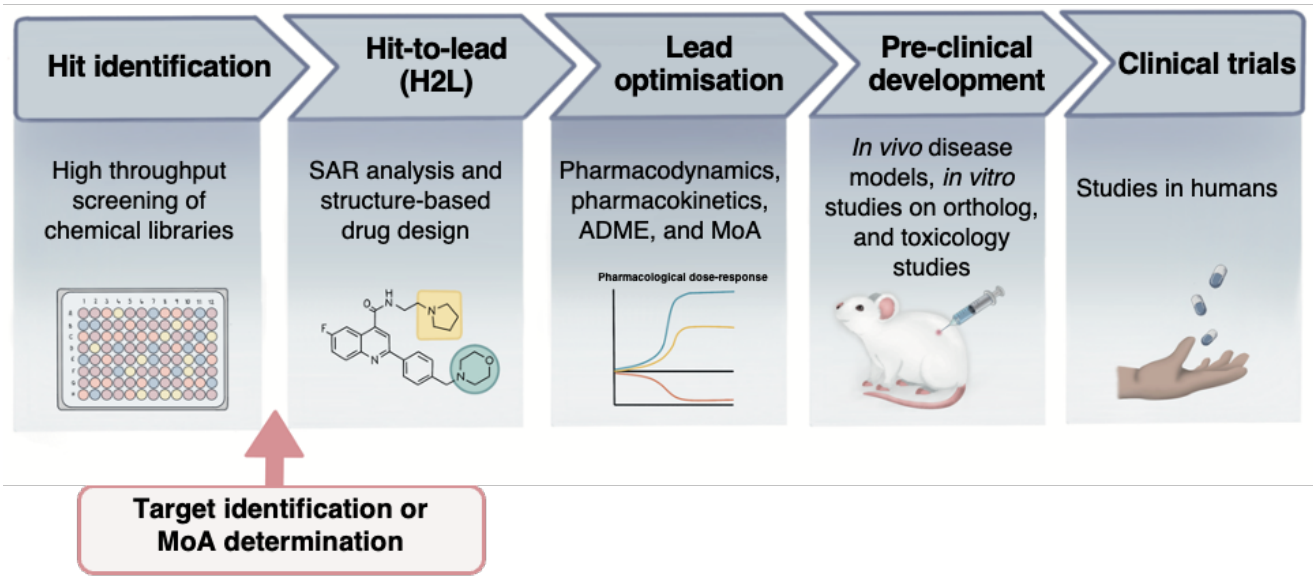


Figure 3. The drug discovery pipeline. The phenotypic drug discovery pipelines that shows when target identification (pink) takes place. This figure was adapted from Zheng *et al.* (2013) **(73)** and created using Procreate (https://procreate.art). SAR = structure-activity relationship, ADME = Absorption, distribution, metabolism, and excretion and MoA = mode of action.

1.5 Target identification and MoA determination

© University of Pretoria

Target identification and MoA determination of antimalarial compounds are complex and time-consuming processes and a wide variety of approaches have been developed by international consortia like the Malaria Drug Accelerator (MalDA, https://www.malariada.org) (72).

1.5.1 Target identification strategies

1.5.1.1 A genomics-based strategy

In vitro evolution and whole genome analysis (IVIEWGA) is a frequently applied forward-genomics approach (enhanced evolution) that allows for the evaluation of potential resistance mechanisms, and if resistance-conveying mutations are within the target protein, this allows for target identification. It involves exposing asexual *P. falciparum* parasites to sublethal levels of the compounds repeatedly until recrudescent, resistant parasites emerge (72, 74). The resistant parasite clones can then undergo whole genome sequencing to allow the identification of mutations such as single nucleotide variants (SNVs) or copy number variants (CNVs) that bestow resistance compared to the parent clone. The mutations often directly relate to measures taken by the parasite to overcome the pressure conferred by compound treatment or to compensate for the effect of compound treatment, which can result in a mutation in the gene of the compound's target protein (75, 76).

This method has led to the successful identification of numerous targets, however, despite this success, some compounds do not easily develop resistance or develop resistance through non-specific mutations such as mutations in the *P. falciparum* multi-drug resistance 1 transporter (51, 72). When the compound does not lead to resistance development, it is denoted as 'irresistible', an attractive quality during drug development (72). Additionally, the development of resistance could potentially only indicate the resistance mechanism instead of the target.

The identified mutations must be validated by reintroducing the mutation in wild-type parasites or lengthy mechanistic studies to verify that the identified SNVs/CNVs are what confers the resistant phenotype (76, 77). One way to validate the mutations identified during IVIEWGA is by introducing the mutant allele into a wild-type parent parasite clone. Clustered regularly interspaced short palindromic repeats/Cas9 (CRISPR/Cas9) allow the introduction of double-stranded breaks at a specified sequence directed by gRNAs and the donor-repair sequence can then incorporate the identified SNV followed by monitoring the parasites to discover if the mutation had the desired effect (72). When resistance cannot be induced, alternative target identification approaches can be applied.

1.5.1.2 Proteomics-based strategies

Chemical proteomics can provide evidence of direct interaction between the investigative compound and its target and can be applied to asexual and gametocyte-stage parasites. One approach is to use parasite cell extracts to identify protein interactions with the compound covalently immobilised on magnetic beads. The proteins identified in this manner undergo quantification via mass spectrometry

© University of Pretoria

and are considered putative drug targets (72). This approach assisted in confirming the target of MMV048 in *P. falciparum* parasites as PI4K (78). This method is simplistic and can identify targets that bind with medium and high affinity to the compound with minimal uncertainty. However, the disadvantages of this method include not being able to immobilise certain compounds on a matrix or the compounds losing their activity due to the modification, not being able to release the target due to the compound covalently modifying the target and the added difficulty of working with protein complexes and low abundance proteins (79).

When a compound cannot be immobilised or the immobilisation compromises the target-binding ability of the compound, a cellular thermal shift assay (CETSA) can be applied to investigate a compound's target. The assay is built on the principle of the thermodynamic stabilisation acquired by the protein upon binding of a ligand. CETSA can be performed using intact cells (*P. falciparum* parasite-infected erythrocytes) or a lysate of the soluble protein fraction, the distinction being the environment of the drug-target interaction (80). Lysate CETSA only monitors the binding of the ligand to the target protein, while intact-cell CETSA provides greater insight into the complexity of the intracellular environment of the interaction (80).

In addition to the two sample preparation methods above, there are two assay versions. Isothermal dose response (ITDR) CETSA tests ten different concentrations of the compounds at three set temperatures, while in melt-curve CETSA a single compound concentration is tested across a thermal gradient of ten temperatures. The temperatures for the ITDR-CETSA heat challenge are dependent on the protein's thermal stability. As a starting point, 51 °C and 57 °C are recommended as 51 °C is the average melting temperature of *P. falciparum* parasites' proteome, whereas 57 °C allows the detection of thermostable proteins. A control temperature (37 °C) where denaturation does not occur is also included (80). For melt-curve CETSA a high drug concentration to ensure ligand-induced protein stabilisation and a range of 37-73 °C is suggested for *Plasmodium* proteins. CETSA shows limitations with low abundance proteins and membrane-bound proteins and has to be accompanied by downstream efficacy and functionality studies (81).

The Drug Affinity Responsive Target Stability (DARTS) assay overcomes the challenge of working with membrane proteins faced by CETSA. It detects protein-ligand interactions based on the principle that some ligands protect their target protein from proteolysis (82). A sample of the cell lysate is treated with the compound under investigation together with either the vehicle control or an inactive analogue, followed by subsequent protein-limited digestion using proteases. The samples are then separated by sodium dodecyl-sulphate polyacrylamide gel electrophoresis (SDS-PAGE), where a non-degraded band that shows protection from proteolysis is the target protein, the protein is then identified using mass spectrometry (83). DARTS is advantageous as it does not require the modification or immobilisation of the compound (83). It is, however, limited by the binding affinity of the compound for its target and how susceptible the protein is to proteolysis (84).

© University of Pretoria

The genomic and proteomic approaches mentioned are expensive, labour-intensive, and sometimes the target is not discovered. When the identification of a specific target is not possible, the MoA can be determined to ascertain to some extent what the compound-induced effect is to enable the successful progression of the compound through the drug discovery pipeline.

1.5.2 MoA determination strategies

1.5.2.1 A metabolomics-based strategy

Inhibition of metabolic-associated enzymes can lead to the diminution of their products and accumulation of the substrates used by the enzyme. These changes can then be used to probe the compound's target by comparing the untreated and treated metabolome (85). Biochemical pathway analysis after compound treatment using metabolomics has proven effective in identifying the MoA of several antimalarials found in the MMV Malaria Box. Metabolic effects of the compounds were evaluated using ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS) and the data were represented as metabolic fingerprints (86). Metabolome coverage, however, tends to be much lower than that of genomes, transcriptomes, and proteomes (86, 87). This is owed to metabolites' fluctuating complexity, large dynamic *in vivo* concentration ranges, and the sheer number of metabolites in existence (87).

1.5.2.2 Transcriptomics-based strategies

Compound-induced fluctuations in transcriptional response provide an objective approach to probing the compound's MoA (88). Transcriptomics are useful to determine the rapid and long-lasting effects of compound treatment because the transcriptome provides dynamic insight into the changes in the cell's requirements after compound treatment (85). The rationale behind this method is that compounds that have overlapping biological effects or targets will show similar transcriptional responses. Transcriptomics has proven to be a useful tool for MoA deconvolution in other fields of clinical research, such as oncology and tuberculosis, and will be discussed in more detail below.

1.6 Transcriptomics for drug MoA studies

In oncology research, the Connectivity Map (CMap) is an example of a database of genome-wide expression profiles obtained from human cancer cell lines after compound treatments. The CMap database can be applied to drug repositioning, lead identification, and as an indication of compound-induced transcriptional effects (89). This approach allowed the identification of a breast cancer stem cell-targeting compound, lovastatin, to combat tumour growth and recurrence (90).

Baillif *et al.* (2020) investigated large publicly available compound-induced gene expression datasets to predict a compound's activity by building machine-learning models with the data (91). The authors showed that compounds elicited similar transcriptional responses in one or multiple different cell lines

© University of Pretoria

independently of their chemical structures. Iorio *et al*. (2010) also developed a method that uses similarities in compound-induced gene expression fluctuations to predict MoA (92). They used a network theory to stratify compounds into groups of interconnected nodes/communities, which are enriched for compounds impacting the same biological pathways or that share MoAs. Using this method, Iorio *et al.* accurately predicted the MoA of nine anticancer compounds (92). Similarly, Mashima *et al.* (2015) showed that transcriptional responses could be used to reflect the anticancer compounds' MoAs (93). The authors treated human lung cancer cell lines with various commonly used anticancer compounds and promising clinical candidates, followed by gene expression analysis using microarrays. Hierarchical clustering of the expression profiles revealed that compounds that target the same or related pathways cluster together, such as HDAC inhibitors and oncogenic kinase inhibitors (93).

In the tuberculosis research field, Boshoff *et al*. (2004) generated a set of 430 microarray profiles of the transcriptional response of *Mycobacterium tuberculosis* to compounds. They successfully grouped the microarray profiles of the compound treatments with similar MoAs using unbiased grouping and predicted the MoA of multiple compounds that were previously uncategorised (94). Similarly, Murima *et al*. (2013) developed a microfluidics-based miniaturised gene expression assay to assist in MoA deconvolution. They successfully identified a minimum gene set to function as biomarkers for MoA deconvolution by analysing the transcriptional responses of the whole Mycobacterial genome to antibacterial compounds. They used unsupervised hierarchical clustering to group compounds that have similar MoAs (95).

The success of transcriptomics-based methods for MoA identification in other disease models, such as oncology and tuberculosis, made it an attractive avenue to investigate the MoA of antimalarials in the drug discovery pipeline similarly.

1.6.1 Transcription and transcriptional regulation in the IDC of *P. falciparum* parasites

P. falciparum parasites developed an exceptionally specialised method of transcriptional regulation that creates a continuous cascade of gene expression (Figure 4) (96). This periodic gene expression during the IDC follows a "just-in-time" expression profile wherein the parasite only expresses a gene during a development stage when it is specifically required.

The different developmental stages of IDC parasites have tightly regulated gene expression that is distinctive to each stage. During the ring stages, there is low transcriptional activity due to the compact nature of the DNA and the greater presence of nucleosomes (96, 97). The trophozoite stages are accompanied by erythrocyte remodelling and many morphological changes. Trophozoites have substantially increased genome volume and euchromatin conformation, correlating to the active transcription seen in trophozoites (96). In schizont-stage parasites, there is reduced transcription as nucleosomes are repacked, leading to chromatin compaction, which is required to form daughter

© University of Pretoria

merozoites for the next IDC (97). However, the merozoites show highly active transcription for genes that contribute to invasion despite the compact nature of the DNA (97). To achieve this continuous cascade of gene expression, which changes based on parasite development and environmental conditions, tightly coordinated transcriptional control is crucial during every stage. Transcriptional control is facilitated by the interplay between chromatin organisation, epigenetic marks and transcription factors (TF).

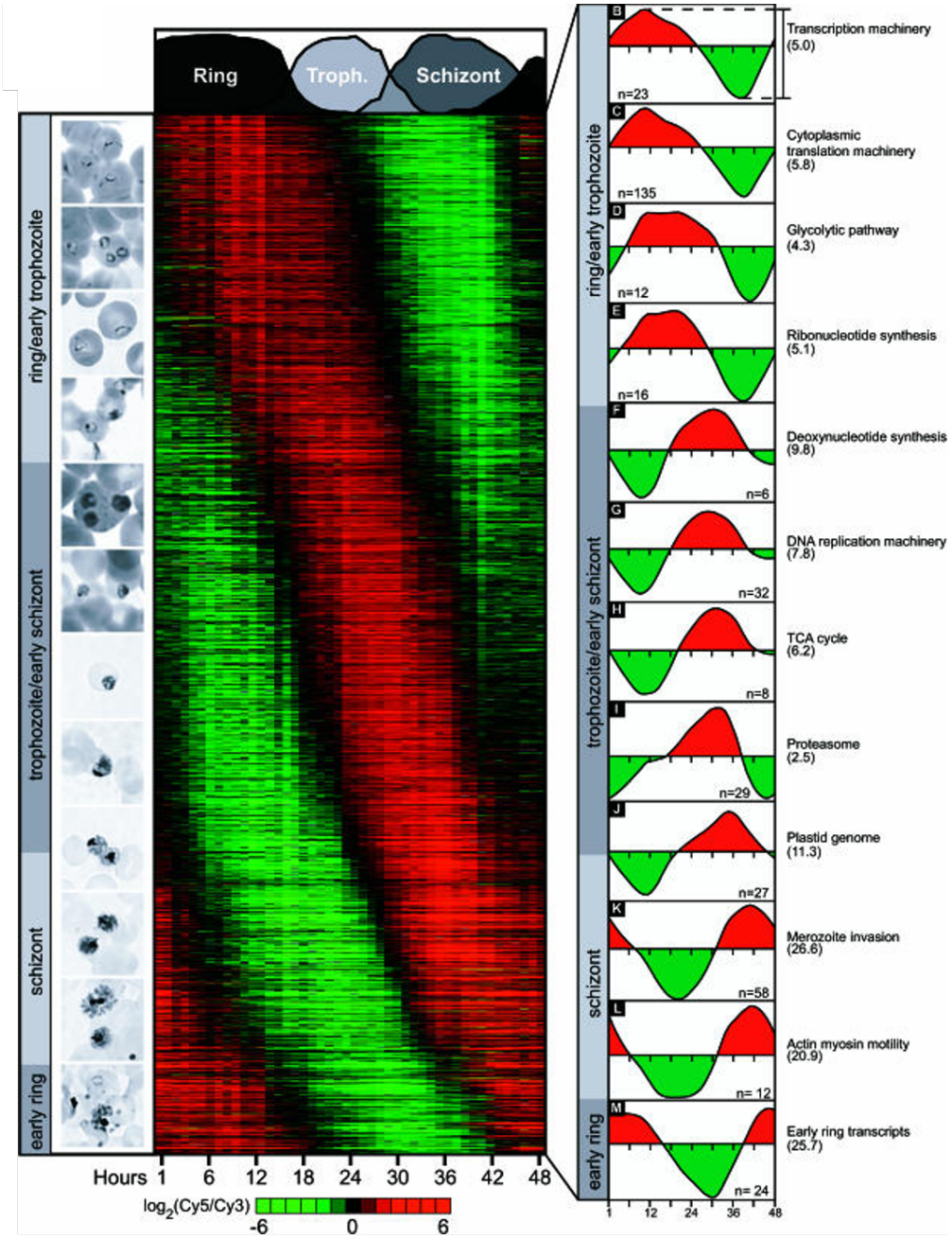


Figure 4. The IDC transcriptome of *P. falciparum* parasites. The representative morphology of the parasites throughout the intraerythrocytic development cycle (IDC) is shown on the left and on the right, 2712 genes were ordered by their phase of expression on the y-axis to create a phaseogram showing the tightly regulated transcription of *P. falciparum* parasites during the IDC. For the $\log_2$(Cy5/Cy3) scale green indicates downregulated gene expression while red indicates upregulated gene expression. Figure obtained from Bozdech *et al.* (2003) (96) (Copyright © 2003 Bozdech *et al.*).

P. falciparum parasites possess relatively few sequence-specific TFs, with the largest classified family being the Apicomplexan APETALA2 (ApiAP2) family, which contains less than 30 members (98). The

© University of Pretoria

ApiAP2 family are master regulators responsible for the stage-specific transcription of hundreds of genes (97). The AP2 TFs regulate specific subsections of genes such as AP2-I, which is important for the expression of invasion-related genes, while AP2-G regulates gametocytogenesis and AP2-O regulates expression in ookinetes (99). Another important TF family in *P. falciparum* parasites are the Myb (from myeloblastosis) proteins, which regulate growth and differentiation (100). TFs have DNA-binding domains that can bind to promoter regions or enhancers to recruit chromatin remodelling complexes and the pre-initiation complex.

Genomic DNA is wrapped around nucleosomes in eukaryotes, including *P. falciparum* parasites, which allows for compacted chromatin and plays a regulatory role in gene expression. Nucleosome-depleted regions of active promoters are generally present upstream of the transcription start site for binding of the pre-initiation complex. In contrast, silenced genes have higher nucleosome occupancy to obstruct pre-initiation complex interaction (97, 101). Additionally, the histones that make up nucleosomes have protruding N-terminal tails, which can undergo post-translational modifications (PTMs) to assist in gene expression regulation (97). PTMs contribute to the interaction of nucleosomes with DNA and, therefore, chromatin conformation and, ultimately, gene expression (102).

1.6.2 Transcriptomics in antimalarial MoA deconvolution

Hu *et al*. (2010) measured the independent response of *P. falciparum* parasites' transcriptome after exposing the parasites to 20 active compounds, followed by transcriptional analysis using microarrays (Figure 5) (103). Three compound classes were identified, and compounds were sorted into classes based on their transcriptional responses and the number of genes in the parasites that were affected (minimum 3-fold difference) by treatment with the compound.

Compounds in the first class induced < 50 genes with an overall transcriptional impact on < 250 genes. Parasite populations treated with compounds in the first class, like quinine and chloroquine, showed reproducible low levels of transcription (overall showing Log_2 fold change (FC) in transcription close to 0) in response to compound treatment, possibly due to their high level of toxicity. Chloroquine was used to treat at three different concentrations and showed a dose-dependent response. Compounds in the second class induced > 50 genes and had an overall impact on 250-500 genes and typically inhibited the development of schizont-stage parasites, which included ML-7 and W-7, which inhibit calcium/calmodulin-dependent protein kinases (CDPKs) and cyclosporine A, which inhibits the calcineurin pathway. There was minimal overlap in transcriptional responses between CDPK and calcineurin pathway inhibitors, which indicated the signalling pathways have non-overlapping functions, but both ultimately influenced the regulation of transcription. Compounds in the third class induced transcription in > 250 genes, had an overall impact on > 500 genes and arrested IDC development, which indicated that they interfered with life cycle regulations and included staurosporine and trichostatin A (TSA), which both inhibit HDACs (103).

© University of Pretoria

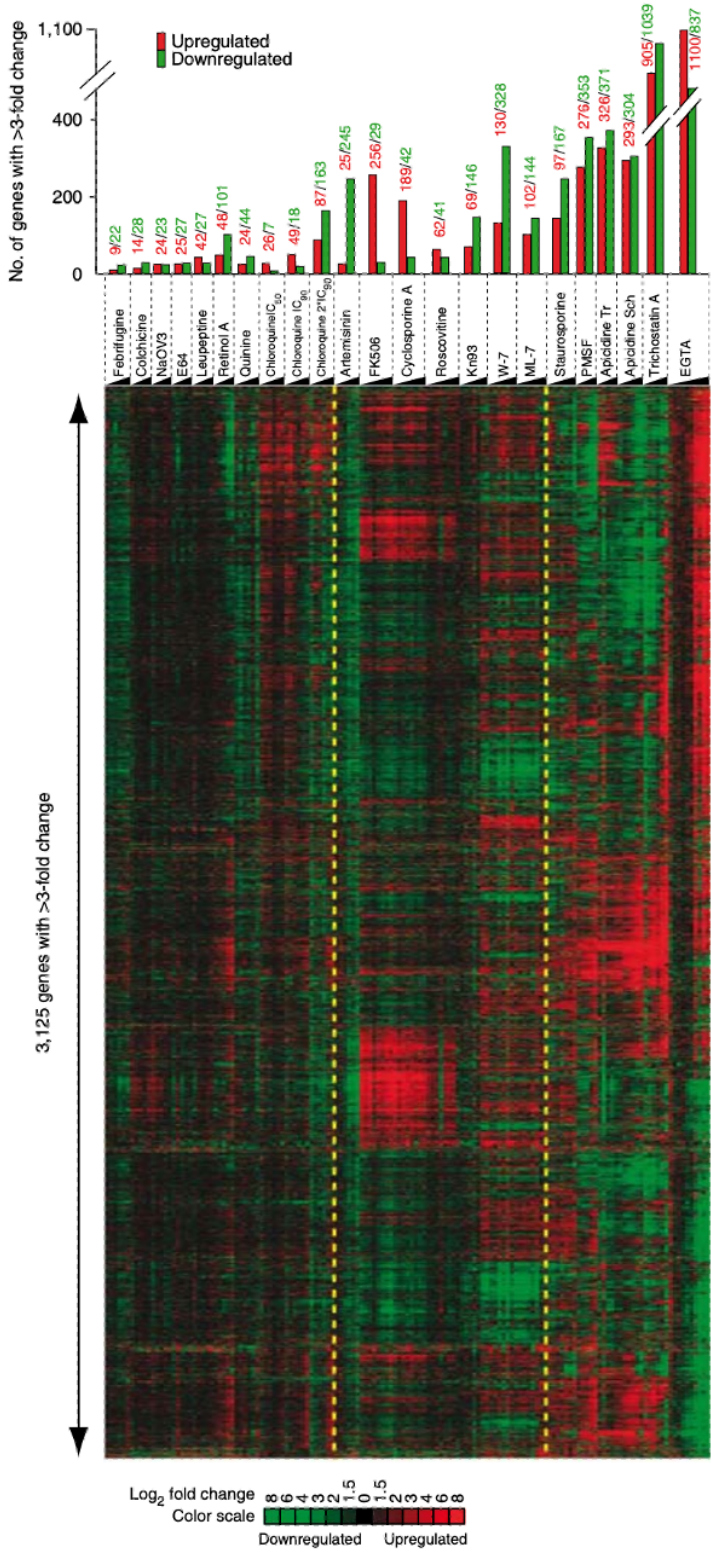


Figure 5. *P. falciparum* parasites' transcriptional responses due to compound treatments. The top bar graph indicates the number of genes showing a minimum of 3-fold change in gene expression in response to treatment with each of the compounds. The bottom heatmap shows the transcriptional response of 3125 genes elicited by each of the compounds in their set, which displays distinctive profiles for each compound. The yellow dashed lines indicate the three different compound classes. For the $\log_2$ fold change scale, green indicates downregulated gene expression, while red indicates upregulated gene expression. Figure obtained from Hu *et al.* (2010) (103) (Copyright © 2010 Nature America, Inc.).

Hu *et al.* (2010) then investigated co-transcriptional properties of functionally related genes through the Pearson correlation coefficient of a subsection of the genes investigated between the compound-induced transcription profiles. This led them to conclude that there was a high level of transcriptional

© University of Pretoria

co-regulation between genes that had related functions. This meant that this approach could provide insights into *P. falciparum* parasites' functional genome that was not possible when simply looking at gene expression throughout the IDC (103). They additionally observed that treatment with the compounds showed distinguishable expression profiles, allowing them to identify differentially expressed genes. DNA microarray-based profiling of compound-induced effects in *P. falciparum* parasites can, therefore, be applied to generate a high-resolution transcriptional dataset that shows functional relationships between genes. Additionally, the compound-induced effect is reproducible and dose-dependent (103).

Another study by Siwo *et al*. (2015) also showed that exposure of *P. falciparum* parasites to different compounds results in a transcriptional response specific to the MoA and chemical structure of the compound (104). They used microarrays to estimate genome-wide transcript abundance in response to treatment with the individual compounds to create genome-wide response indices (a ratio between the average transcript levels of the gene of interest (GOI) after compound treatment and the average of the transcript levels of the GOI across all the compound treatments). These indices were then hierarchically clustered, which led to the discovery that similarities in chemical structure and more often, similarities in MoA between compounds produced similar transcriptional responses (Figure 6). This provided a framework for relating compound MoA to transcriptional responses in *P. falciparum* parasites.

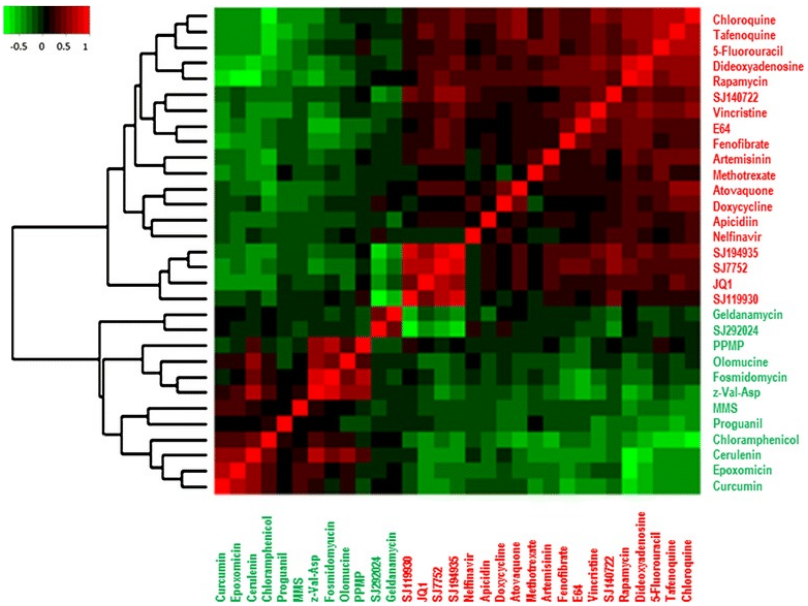


Figure 6. *P. falciparum* parasites' transcriptional responses to compound treatment show a relation to chemical structure and MoA. The compounds were hierarchically clustered based on their genome-wide response indices. For the colour scale indicating the correlation of each compound pair using the genome-wide response index scale, green indicates positive correlation while red indicates negative correlation. The red and green colour of the compound name was used to indicate that compounds of the same colour show a positive correlation while different colours show a negative correlation. Figure obtained from Siwo *et al*. (2015) (104) (Copyright © Siwo *et al*. 2015. (http://creativecommons.org/licenses/by/4.0/)). MoA = mode of action.

This study proved that there are two main contributors to similarities in transcriptional responses, similarities in chemical structures and similarities in targeted proteins or pathways (Figure 7). Notably,

© University of Pretoria

not all compounds that share chemical structures induce similar transcriptional responses, while compounds that have overlapping biological effects do have similar transcriptional responses. With a principle component analysis, Siwo *et al.* (2015) determined that 65 % of the variation in correlation was explained by the first variation dimension (Dim 1), where the compounds were divided into Class I and II, which corresponded to the clusters observed in the heatmap (Figure 7) (104).

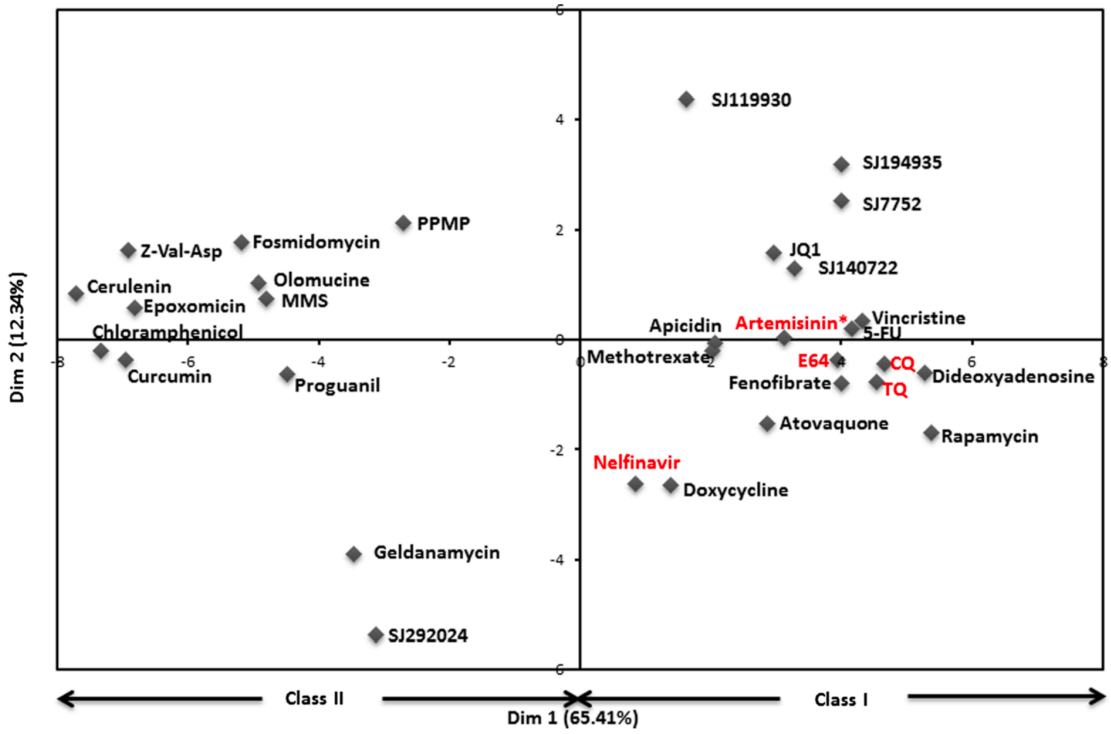


Figure 7. The principle component analysis of compound-induced transcriptional profiles. The compounds are divided into two main clusters by the first component of variation (Dim 1), which indicates compounds in class I and class II (based on chemical substructure differences). Dimension 2 (dim 2) indicates small molecule transcriptional response. The plot indicates that chemical similarities and converging modes of action lead to similar transcriptional responses. Figure obtained from Siwo *et al.* (2015) (104) (Copyright © Siwo *et al.*, 2015. (http://creativecommons.org/licenses/by/4.0/)).

Together, the Hu *et al.* (2010) and Siwo *et al.* (2015) studies provided proof of concept that compound-induced transcriptional profiling was compound-specific and dose-dependent, providing insight into the biological effects of a compound, which will allow the MoA to be elucidated. Such insights have been applied in other studies to interrogate the on-target MoA of antimalarial candidates, including kinase-specific inhibitors (105, 106). This revealed the MoA of new chemical compounds in a series to still target the same drug target as the parent compound since they had similar transcriptome fingerprints. Whilst these studies are highly informative and support the hypothesis that transcriptional profiling is useful to describe the MoA of antimalarial candidates, similar to what is seen in tuberculosis and oncology research fields, these data were all based on whole transcriptome analyses by e.g. DNA microarrays. However, such transcriptome-associated approaches would not be conducive to routine drug discovery because they are time- and labour-intensive and possess inherent complexities.

This underscores the potential of using these transcriptional profiles to extract a limited set of transcripts to function as biomarkers for MoA stratification. van Heerden *et al.* (2021) used the basis of these

© University of Pretoria

studies to explore if machine learning could be applied to extract biomarkers for MoA stratification from pre-existing compound-induced transcriptional response data. Machine learning to classify compounds based on their MoA has not been established in the antimalarial drug discovery process until van Heerden *et al*. (2021) recently developed a model to do so (107).

1.7 A chemo-transcriptomic MoA classification model in *P. falciparum* parasites

van Heerden *et al*. (2021) developed a rational gene selection approach that could identify a limited set of predictive biomarkers for MoA classification to accelerate and inform antimalarial drug discovery. They aimed for their machine learning model and biomarkers to provide a medium-throughput approach to MoA deconvolution. The predictive features could then be used to generate and train machine-learning models to accurately classify compound MoA (107).

Meta-analysis of published gene expression profiles of IDC *P. falciparum* parasites after compound treatment was performed (103, 108-111). From these gene expression profiles, they could identify differentially expressed (upper or lower 5th percentile expression) transcripts that would be useful for compound MoA classification to build machine-learning models based on the feature selection criteria set out. Feature selection for the rational selection models included: 1) removal of non-informative transcripts that did not show significant differential expression, 2) removal of differentially expressed genes that show sporadic profiles throughout compound treatment and time points, 3) genes differentially expressed at all time points of treatment were selected, and 4) genes that are shared between compound treatments were not considered further as they potentially show general drug stress and are not specific to a MoA (107). The transcripts were ranked by MoA stratification variable importance, indicating how reliant the machine learning model is on a specific transcript for accurate classification. The same number of transcripts as was identified using the rational selection approach were obtained using a machine-learning-inferred approach to evaluate which set of transcripts performed better.

van Heerden *et al*. (2021) used 80 % of the compounds to build the models and the remaining 20 % to test the models. They compared the two different gene selection approaches, machine-learning-inferred and rationally-selected, using a sliding gene-scale method. This entails building smaller models (mini models) using sequentially fewer transcripts to ascertain the minimum number of transcripts that allows for accurate MoA stratification. Most of their models showed acceptable accuracy (~70 %) (Figure 8A). However, some of the models overfit the data, leading to too many differences in accuracy when the training sets were compared to the test sets, while other models had undesirable variability. When choosing the best-performing model, they considered the model's performance with the test set of compounds, the accuracy and variability of the model, and the minimum number of transcripts used to train the model. They identified the rational selection 50 transcript model as the optimum minimum

© University of Pretoria

transcripts that were still capable of stratifying compounds having similar MoAs with an accuracy of 76.6 ± 6.4 % (red block Figure 8A).

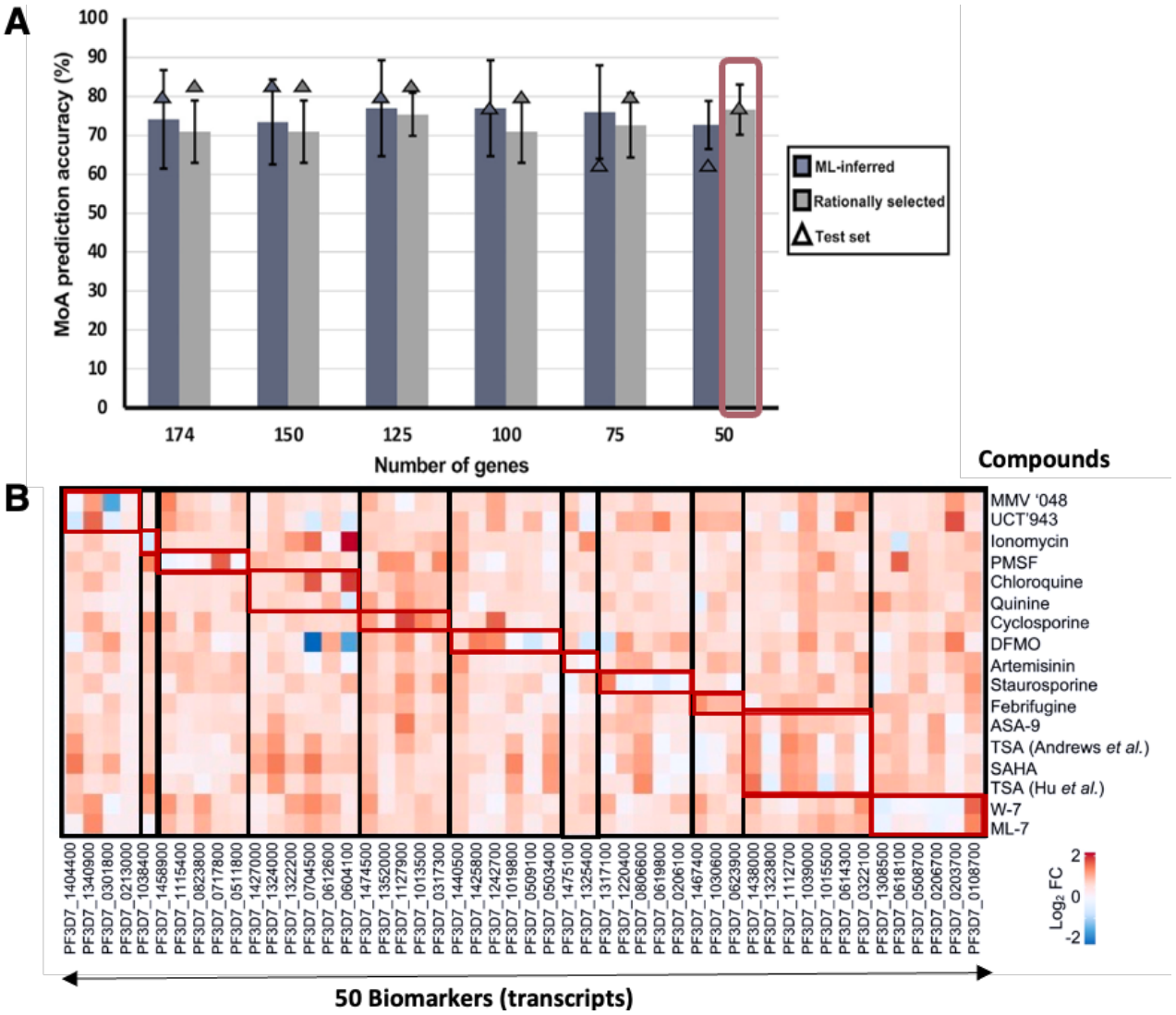


Figure 8. Identification of the best model, the rational selection 50 transcript model. (A) Investigating the minimal number of transcripts needed for compound MoA stratification by machine learning models. In the red block is the model investigated in this project. **(B)** Evaluating the correlation between different compounds using the top 50 transcripts (biomarkers). Figure obtained from van Heerden *et al.* (2021) (107) (Copyright © 2021 van Heerden, van Wyk and Birkholtz).

The 50 identified, potential biomarkers from the rational selection model were evaluated for each compound from the data set to ascertain the compound-specific transcriptional profiles. From this, it was clear that each compound treatment had distinguishable transcriptional responses for the different compounds and that each of the 50 biomarkers identified in the rational selection model was unique (Figure 8B). The 50 transcripts they selected as biomarkers represent specific compound classes and, therefore, MoAs (Table 1).

© University of Pretoria

Table 1. The top 50 transcripts identified as potential biomarkers for the rational selection 50 transcript model. The compounds from which the biomarkers were obtained are also shown, as well as the MoA associated with them.

Biomarker (PlasmoDB Gene ID)	Compound	Reference	Mode of Action
1. PF3D7_0108700 2. PF3D7_0203700 3. PF3D7_0206700 4. PF3D7_0508700 5. PF3D7_0618100 6. PF3D7_1308500	ML-7 & W7	(112)	Calcium/calmodulin-dependent protein kinase (CDPK) inhibitors
7. PF3D7_0322100 8. PF3D7_0614300 9. PF3D7_1015500 10. PF3D7_1039000 11. PF3D7_1112700 12. PF3D7_1323800 13. PF3D7_1438000	Trichostatin A (TSA), suberoylanilide hydroxamic acid (SAHA) and 2-aminosuberic acid derivative (ASA-9)	(108, 112)	Histone deacetylase (HDAC) inhibitors that perturb the transcriptome
14. PF3D7_0623900 15. PF3D7_1030600 16. PF3D7_1467400	Febrifugine	(112)	Targets *P. falciparum* prolyl-tRNA synthetase activity
17. PF3D7_0206100 18. PF3D7_0619800 19. PF3D7_0806600 20. PF3D7_1220400 21. PF3D7_1317100	Staurosporine A	(112)	Serine/threonine kinase inhibitors, reduces merozoite invasion
22. PF3D7_1325400 23. PF3D7_1475100	Artemisinin	(112)	Involved in the production of carbon-centered free radicals that alkylate haem and proteins (hypothesised)
24. PF3D7_0503400 25. PF3D7_0509100 26. PF3D7_1019800 27. PF3D7_1242700 28. PF3D7_1425800 29. PF3D7_1440500	Difluoromethylornithine (DFMO)	(109)	Inhibits ornithine decarboxylase (ODC) causing parasite arrest
30. PF3D7_0317300 31. PF3D7_1013500 32. PF3D7_1127900 33. PF3D7_1352000 34. PF3D7_1474500	Cyclosporine A	(112)	Inhibits merozoite invasion by binding sphingomyelin
35. PF3D7_0612600 36. PF3D7_0704500 37. PF3D7_1324000	Chloroquine	(112)	Inhibits haem polymerase enzyme
38. PF3D7_0604100 39. PF3D7_1322200 40. PF3D7_1427000	Quinine	(112)	Accumulates in the digestive vacuole of parasites (possible inhibition of haem detoxification)
41. PF3D7_0511800 42. PF3D7_0717800 43. PF3D7_0823800 44. PF3D7_1115400 45. PF3D7_1458900	Phenylmethylsulfonyl fluoride (PMSF)	(112)	Serine protease inhibitor
46. PF3D7_1038400	Ionomycin	(110)	Calcium-binding ionophore
47. PF3D7_0213000 48. PF3D7_0301800 49. PF3D7_1340900 50. PF3D7_1404400	MMV048 & UCT943	(111)	Inhibits phosphatidylinositol 4-kinase (PI4K)

© University of Pretoria

Self-organising maps (SOMs) were used to cluster and condense the expression patterns of the biomarkers and allow their visualisation as two-dimensional (2D) suprahexagonal chemo-transcriptomic fingerprints (further referred to as chemo-transcriptomic fingerprints) (Figure 9A). SOMs are a form of artificial neural network and function based on unsupervised learning (113). Within the fingerprints, each small hexagon represents a cluster of genes with either the same or very similar expression levels, as there are fewer hexagons than biomarkers.

Biomarkers with the most influence are located at the edge of the fingerprint, while biomarkers with small or no changes in gene expression levels will be located at the centre. Where the hexagons are located in the fingerprint shows how similar the hexagon is to the neighbouring hexagons. These chemo-transcriptomic fingerprints are unique to each compound, and compounds with similar targets or MoAs will cluster together. The unique fluctuation in transcriptional responses between the different compounds could differentiate between the compounds based on their MoA, which allowed separation between CDPK inhibitors and PI4K inhibitors, for example (Figure 9B).

This model provides a new means to indicate compound MoA quickly and specifically using their chemo-transcriptomic fingerprints and would, therefore, help to accelerate antimalarial drug discovery by highlighting the most viable drug candidates with preferential MoAs to be investigated. These potential biomarkers still need to be validated *in vitro* before they can be used for drug discovery purposes.

© University of Pretoria

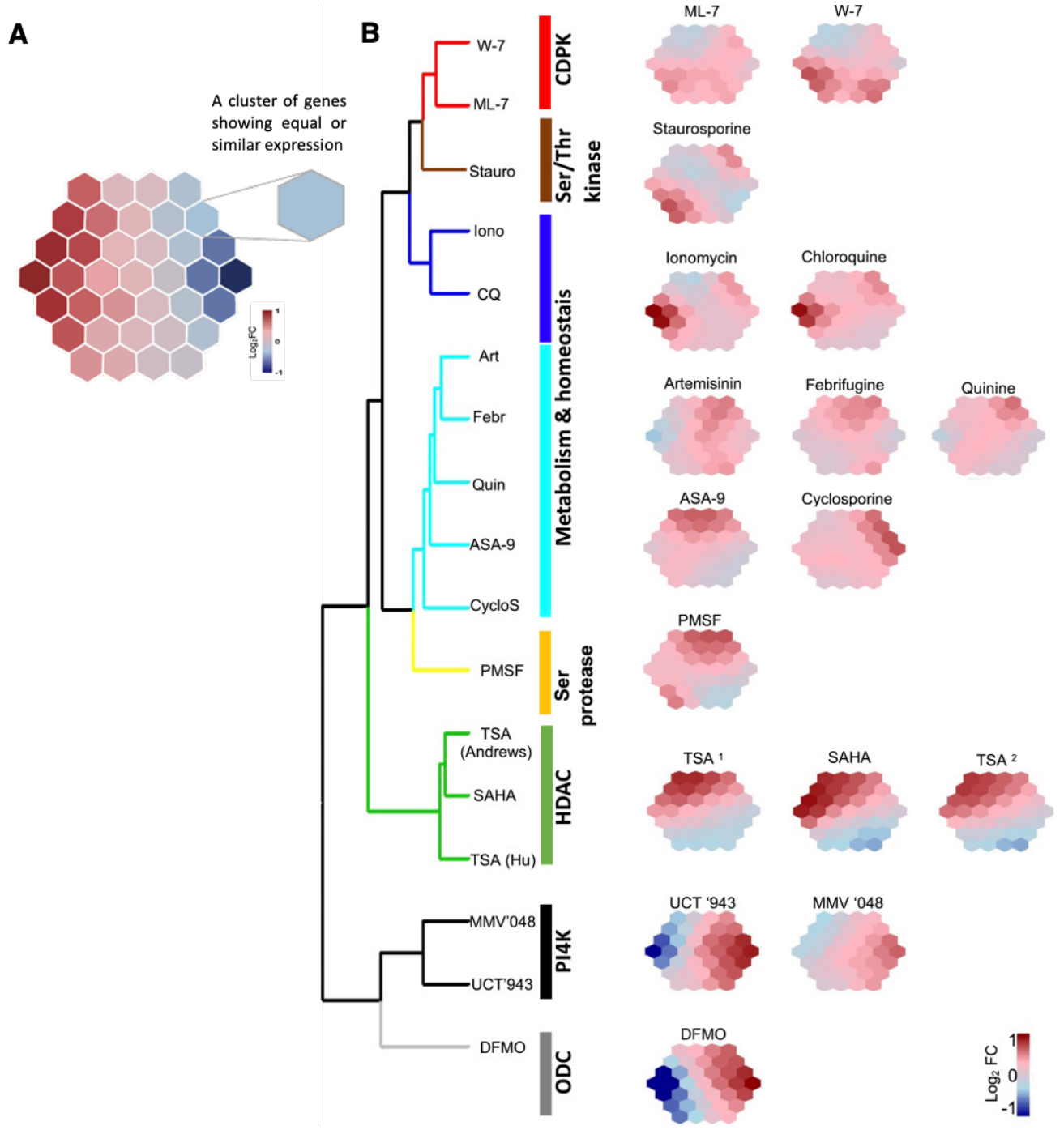


Figure 9. Clustering of the suprahexagonal chemo-transcriptomic fingerprints of different compounds according to their MoAs. (A) Self-organising maps (SOMs) depicted as a suprahexagonal chemo-transcriptomic fingerprint where a single hexagon within the chemo-transcriptomic fingerprint represents a cluster of biomarkers having the same/similar expression patterns. The colour of the individual hexagons represented the $\log_2$ fold change in gene expression, values > 0 represented increased transcriptional abundance, while values < 0 represented decreased transcriptional abundance. This figure was created using Procreate (https://procreate.art). **(B)** SOMs were used to cluster and condense the data followed by hierarchical clustering of the compound treatments using Ward linkage on Euclidian distance of expression. The colour of the individual hexagons represents the $\log_2$ fold change in gene expression. This figure was obtained from van Heerden *et al.* (2021) (107) (Copyright © 2021 van Heerden, van Wyk and Birkholtz). CDPK = calcium/calmodulin-dependent protein kinase, Ser/Thr kinase = serine/threonine kinase, Ser protease = serine protease, HDAC = histone deacetylase, PI4K = phosphatidylinositol 4-kinase, ODC = ornithine decarboxylase, Stauro = staurosporine, Iono = ionomycin, CQ = chloroquine, Art = artemisinin, Febr = febrifugine, Quin = quinine, ASA-9 = 2-aminosuberic acid derivative, CylcoS = cyclosporine A, PMSF = phenylmethylsulfonyl fluoride, TSA = trichostatin A, SAHA = suberoylanilide hydroxamic acid, DFMO = difluoromethylornithine, TSA[1] = data from Hu *et al.* (2009) (112), TSA[2], data from Andrews *et al.* (2012) (108).

© University of Pretoria

1.8 Aim, hypothesis, and objectives

1.8.1 Aim

Design and optimise a real-time qPCR-based platform to validate chemo-transcriptomic profiles of antimalarial compounds generated using machine learning as a means of predicting a compound's mode of action.

1.8.2 Hypothesis

The machine learning predicted biomarkers can separate compounds based on transcriptional profiling into specific mode of action classes.

1.8.3 Objectives:

1. To generate a real-time qPCR platform for chemo-transcriptomic fingerprints from the predicted 50 biomarkers obtained from the machine learning model.

- Build and optimise a qPCR-based platform to evaluate the biological relevance of the 50 predicted biomarkers.
- Introduction of control compounds from the van Heerden *et al.* (2021) dataset to evaluate the platform's capabilities.

2. Introduction of clinical candidates not previously associated with the model.

3. Introduction of compounds with shared MoAs to evaluate the clustering capacity of the platform.

1.9 Research outputs

Natanya Venter, Ashleigh van Heerden, and Lyn-Marié Birkholtz. *In vitro* validation of chemo-transcriptomic profiles for drug mode of action classification. 28th South African Society for Biochemistry and Molecular Biology Congress 2024. Poster presentation, Polokwane, July 2024.

Natanya Venter, Ashleigh van Heerden, and Lyn-Marié Birkholtz. *In vitro* validation of chemo-transcriptomic profiles for drug mode of action classification. 9th Southern Africa Malaria Research Conference. Poster presentation, Pretoria, July 2024.

© University of Pretoria

Chapter 2: Materials and Methods

2.1 Ethical clearance statement

This project holds approval from the University of Pretoria research ethics committee of the Faculty of Health Sciences and the Faculty of Natural and Agricultural Sciences for *in vitro* cultivation of intraerythrocytic *P. falciparum* parasites (NAS040/2023). Human A^+ blood for the cultivation of intraerythrocytic *P. falciparum* parasites was obtained from the South African National Blood Service (SANBS). All research using human blood was performed in a biosafety level 2 (BSL2) certified culturing facility (registration number 39.2/ University of Pretoria -23/160). Additionally, all project-specific work was performed with approval from the University of Pretoria research ethics committees of the Faculty of Health Sciences and the Faculty of Natural and Agricultural Sciences (both under NAS062/2023).

2.2 Culturing of IDC *P. falciparum* parasites

Drug-sensitive *P. falciparum* NF54 parasites (certified mycoplasma free) were obtained in a cryopreserved state from The Malaria Research and Reference Reagent Resource Center ((MR4) https://www.beiresources.org/About/MR4Home.aspx), thawed, and maintained *in vitro* as previously established (114, 115). Parasites were maintained under shaking conditions (60 rpm) in a hypoxic environment (5 % O_2, 5 % CO_2, and 90 % N_2, Afrox, RSA) at 37 °C to replicate the physiological environment. Parasites were cultured in A^+ human erythrocytes at 5 % haematocrit in complete culture medium (RPMI 1640 culture medium (Gibco or Sigma-Aldrich, USA), supplemented with 5 g/L AlbuMax II (lipid-rich bovine serum albumin that served as a substitute for human serum) (Gibco, USA), 25 mM HEPES at pH 7.5 (Sigma-Aldrich, USA), 24 μg/mL gentamicin (Fresenius Kabi, Germany), 200 μM hypoxanthine (Sigma-Aldrich, USA), 11 mM D-glucose (Sigma-Aldrich, USA), and 0.2 % (w/v) sodium bicarbonate (Glentham Life Sciences, UK)). The parasites' progression through the IDC was examined using light microscopy and a Rapi-Diff staining kit (composed of methanol fixative, eosin, and methylene blue) (Clinical Sciences Diagnostics, RSA). The Rapi-Diff staining kit was used to fix a dry, thin blood smear to the microscope slide with methanol, followed by the eosin stain to stain the DNA of the parasite pink and methylene blue stain to stain the cytoplasm blue. All microscopy pictures were taken using a Nikon Eclipse 50i light microscope (Nikon, Japan) with the 100× oil immersion objective and a Nikon D5-Fi1 camera (Nikon, Japan) accompanied by the NIS-elements F software. Parasite cultures were maintained at a parasitaemia (the proportion of infected erythrocytes compared to all erythrocytes counted) between 2-5 %.

2.3 Sorbitol synchronisation of IDC *P. falciparum* parasites

P. falciparum parasites possess the ability to synchronise their development *in vivo,* this allows the parasites to delay their development or enter dormancy due to drug treatment or nutrient starvation (116). Parasites cultivated *in vitro* lose their natural *in vivo* synchronicity after multiple asexual life cycles, this leads to multiple developmental stages being present in a culture. Therefore, when

© University of Pretoria

synchronous parasites were required for experiments, intraerythrocytic asexual parasites were treated twice with 5 % (w/v) D-sorbitol (Sigma-Aldrich, USA) 8 h apart, 48 h before an experiment to yield a > 90 % ring-stage parasite population (117). Sorbitol causes trophozoite and schizont parasite stages to lyse because of their greater erythrocyte membrane permeability caused by the formation of new permeation pathways as the parasites mature, leaving only intact ring-stage parasites (118, 119). The parasites were incubated in sorbitol for 15 min at 37 °C. Afterwards, the culture was centrifuged at 3200 ×*g* for 3 min to pellet the culture and the supernatant aspirated. The remaining pelleted culture was washed twice with culture medium (without AlbuMax II) and then resuspended in fresh complete culture medium to a 5 % haematocrit.

2.4 *In vitro* inhibition of asexual proliferation of parasites

A SYBR Green I asexual proliferation assay was used to determine the *in vitro* activity of the compounds (120). SYBR Green I is a DNA intercalating dye that allowed the quantification of double-stranded DNA (dsDNA) and, subsequently, parasite proliferation since erythrocytes lack nuclei (121). The following compounds were tested at the indicated starting concentrations: 200 nM chloroquine (Sigma-Aldrich, USA), 160 nM artemisinin (Sigma-Aldrich, USA), 1 mM DFMO, 12 μM SAHA (Cayman, Estonia), 400 nM MMV048 (UCT, RSA), 600 nM UCT943 (UCT, RSA), 75 nM DSM265 (Malaria Box, MMV, Switzerland (https://www.mmv.org/mmv-open/malaria-box/about-malaria-box)), 250 nM pyrimethamine (Malaria Box, MMV, Switzerland), 25 nM M5717 (Malaria Box, MMV, Switzerland), and 1.75 μM MMV084 (MMV, Switzerland). The test compounds underwent a two-fold serial dilution in complete culture medium (three-fold for UCT943 and chloroquine) for nine dilutions total in a 96-well plate. A background control (500 nM chloroquine) and a positive for proliferation control (untreated parasites in complete culture media) were included in the experiments. A > 95 % synchronised ring-stage parasite culture suspension (1% parasitaemia, 1% final haematocrit) was prepared for the assays and incubated with the test compounds (prepared as stated above) at 37 °C for 96 h under hypoxic conditions (5 % O_2, 5 % CO_2, and 90 % N_2, Afrox, RSA) to allow the progression of the parasites through two IDCs.

After the incubation period, plates were frozen to assist with cell lysis. An equal volume of SYBR Green I lysis buffer (containing 20 mM Tris-HCl (pH 7.5) (Sigma-Aldrich, USA), 5 mM EDTA (Merck, USA), 0.008 % (w/v) Saponin (Sigma-Aldrich, USA), 0.08 % Triton X-100 (Sigma-Aldrich, USA), and 0.2 μL/mL 10 000× SYBR Green I dye (Thermo Fisher Scientific, USA)) was added to the incubated parasite suspension. An additional 15 min of incubation in the dark at room temperature was done before measuring the fluorescence of the SYBR Green I dye on a Fluoroskan Ascent FL microplate fluorometer (Thermo Labsystems, USA). An excitation wavelength of 485 nm and an emission wavelength of 538 nm were used.

The background fluorescence was accounted for by subtracting the average of the chloroquine control from all data. These adjusted values were then normalised using the average of the corrected untreated

© University of Pretoria

(positive for proliferation) control and the data was expressed as a percentage of parasite proliferation compared to the untreated control. The data was further analysed using GraphPad Prism 10 to generate a four-parameter dose-response curve from which the half maximal inhibitory concentration (IC_{50}), as well as 90 % of the maximum inhibition concentration (IC_{90}) (using the Find EC anything function on GraphPad Prism 10 (F = 10 % proliferation)), was determined.

2.5 RNA isolation and cDNA synthesis

2.5.1 NucleoZOL RNA isolation

A 10 mL asexual NF54 culture (5 % haematocrit, 10 % parasitaemia for method optimisation (~33-37 hours post-invasion (hpi) parasites) and 4-10 % parasitaemia for compound treatments, depending on the incubation period (see section 2.7)) was sedimented. Thereafter, the culture medium was aspirated, leaving only the parasite pellet. Mid to late trophozoite-stage parasites were used for optimisation and comparison to the RNeasy Mini kit (Qiagen, Germany), as this is where *P. falciparum* parasites are the most metabolically and transcriptionally active (122). The fresh parasite pellet was homogenised by lysing the erythrocyte pellet with the addition of 500 µL per 5 × 10^6 cells NucleoZOL one-phase RNA isolation reagent (Macherey-Nagel, Germany) and pipetting. UltraPure™ DNase/RNase-free distilled water (Thermo Fisher Scientific, USA) (200 µL per 500 µL NucleoZOL) was added to the lysate, shaken vigorously, and incubated at room temperature for 5 min. The samples were centrifuged using an Eppendorf Minispin Plus (Eppendorf, Germany) for 15 min at 12 000 ×*g* to form a pellet containing DNA, polysaccharides, and proteins. The supernatant was removed and combined with an equal volume of 100 % isopropanol (Glentham Life Sciences, UK) to precipitate the RNA, followed by an incubation at room temperature for 10 min. The samples were then centrifuged for 10 min at 12 000 ×*g* and the supernatants were discarded. The RNA pellet was washed twice with 75 % ethanol (500 µL per 1000 µL supernatant in the previous step) and the supernatant was removed after centrifugation at 8 000 ×*g* for 3 min. The RNA pellet was reconstituted in 35 µL UltraPure™ DNase/RNase-free distilled water (Thermo Fisher Scientific, USA) and vortexed for 3 min at room temperature to dissolve the RNA pellet.

2.5.2 Qiagen RNeasy Mini kit RNA isolation

A 20 mL asexual NF54 culture (5 % haematocrit, ~10 % parasitaemia (~33-37 hpi parasites)) was sedimented, the pellet was washed three times with 1× PBS (3 mM KCl (MINEMA Chemicals, RSA), 0.137 M NaCl (Glentham Life Sciences, UK), 8.1 mM Na_2HPO_4, 1.9 mM NaH_2PO_4 (Merck, Germany)) and frozen at -80 °C overnight to allow better cell lysis when using a RNeasy Mini kit (Qiagen, Germany). The RNA isolation was performed as per the manufacturer's guidelines. To 500 µL of the still frozen pellet, 700 µL buffer RLT (a lysis buffer with guanidium isothiocyanate and β-mercaptoethanol) was added and the sample was then transferred to a Qiashredder column (Qiagen, Germany) and centrifuged for 90 s at 14 000 ×*g*. The full pellet did not fit in a single column, therefore

© University of Pretoria

the flowthrough of each column was transferred to a single clean nuclease-free tube and 600 µL TRIzol™ LS Reagent (Thermo Fisher Scientific, USA) was added, followed by vigorous vortexing for ~20 s. TRIzol™ LS Reagent is a monophasic solution containing phenol and guanidium isothiocyanate to allow the isolation of RNA, DNA, and proteins in separate fractions. The samples were then incubated at room temperature for 5 min, after which 400 µL chloroform (Associated Chemical Enterprises, RSA) was added and the samples were vortexed vigorously again for ~20 s. The samples were incubated for 10 min at room temperature and centrifuged for 15 min at 14 000 ×*g*. The addition of chloroform and the subsequent centrifugation allowed the separation into an aqueous phase and an organic phase. The upper aqueous phase that contains the RNA was transferred to a new nuclease-free tube; care was taken to leave the lower phases completely undisturbed. An equal volume of 70 % (v/v) ethanol was added to the aqueous phase and mixed using inversion.

The sample was loaded onto an RNeasy column and centrifuged for 15 s at 14 000 ×*g*. Buffer RW1 (700 µL) was loaded to wash the column and samples were centrifuged again for 15 s at 14 000 ×*g*. This was followed by the addition of a mixture of 70 µL buffer RDD and 10 µL RNase-free DNase (to allow on-column digestion of DNA and to ensure the RNA stays column bound) and a 15 min room temperature incubation. Buffer RW1 (700 µL) was loaded onto the column again, followed by centrifugation for 15 s at 14 000 ×*g*. Buffer RPE (500 µL) was loaded to remove contaminating salts from previous buffers, and samples were centrifuged for 2 min at 14 000 ×*g*. The column was then transferred to a new nuclease-free tube and 35 µL RNase-free water from the kit was pipetted directly onto the column, followed by a 3 min incubation. Finally, RNA was eluted by centrifugation at 14 000 ×*g* for 90 s.

2.5.3 RNA integrity evaluation

The RNA concentration as well as the A_{260}/A_{230} and A_{260}/A_{280} ratios were measured spectrophotometrically using a Nanodrop One^{C} (Thermo Fisher Scientific). The integrity of the RNA was evaluated by non-denaturing agarose gel electrophoresis using a 1.5 % (w/v) agarose (Thermo Fisher Scientific, USA) gel dissolved in 1× TAE (40 mM tris-acetate pH 8.0 (Sigma-Aldrich, USA), 1 mM EDTA (Merck, USA)), and pre-stained with 0.75 µg/mL ethidium bromide (EtBr) (Invitrogen, USA) before proceeding with downstream experiments. The RNA samples (300-500 ng) were heated to 70 °C (Accublock digital dry bath) for 5 min to denature the RNA before loading into the solidified gel along with 6× TriTrack DNA loading dye (10 mM Tris-HCl (pH 7.6), 0.03 % (w/v) bromophenol blue, 0.03 % (w/v) xylene cyanol FF, 0.15 % (w/v) orange G, 60 % (v/v) glycerol, and 60 mM EDTA) (Thermo Fisher Scientific, USA)) at a final concentration of 1× TriTrack DNA loading dye. A GeneRuler 1 kb DNA ladder (500 ng) (Thermo Fisher Scientific, USA) was included for size reference. The gel was then visualised using a Gel Doc XR+ imaging system (Bio-Rad, USA). The RNA samples were either directly converted to cDNA or stored at -80 °C until cDNA synthesis could be performed.

© University of Pretoria

2.5.4 cDNA synthesis

For cDNA synthesis, 1 μg of RNA was used in each sample and 5 pmol/μL random N9 primers (Promega, USA) and 5 pmol/μL oligo (dT) primers (Integrated DNA Technologies, USA) were added in the same tube and mixed. The samples were incubated at 70 °C for 5 min to allow annealing of the primers, followed by an incubation at 4 °C for 5 min. ImProm-II 1× reaction buffer (50 mM Tris-HCl pH 8.3, 75 mM KCl, and 10 mM DTT) (Promega, USA), 0.5 mM dNTP mix (KAPA Biosystems, USA), 3 mM $MgCl_2$ (Promega, USA), and 15 U/μL ImProm-II reverse transcriptase (Promega, USA)) was added to the samples. The samples were placed in the Thermal Cycler (Applied Biosystems, USA) under the following conditions: 25 °C for 5 min to allow primer annealing, first strand extension at 42 °C for 60 min, 70 °C for 15 min to allow inactivation of the reverse transcriptase and cooling to 4 °C. RNA-DNA hybrids are formed during reverse transcription; therefore, RNA hydrolysis was performed by adding 6 μL NaOH (1M) (Sigma-Aldrich, USA) and 6 μL EDTA (0.5 M) (Merck, USA) to each 20 μL cDNA reaction followed by a 15 min incubation at 65 °C in the Thermal Cycler (Applied Biosystems, USA).

cDNA clean-up was performed using the NucleoSpin Gel and PCR clean-up kit (Macherey-Nagel, Germany) according to the manufacturer's specifications. To the cDNA samples, 300 μL buffer NT was added to assist with binding to the column. The solution was transferred to a NucleoSpin Extract II column followed by a 5-min incubation at room temperature. The samples were then centrifuged at 13 000 ×*g* for 1 min, and the flow-through was discarded. The column-bound cDNA was washed thrice with wash buffer NT3 (500 μL) and centrifuged for 2 min at 13 000 ×*g* to dry the membrane. The cDNA was eluted using 30 μL of pre-warmed (70 °C) elution buffer to increase cDNA yields. The cDNA concentration as well as the A_{260}/A_{230} and A_{260}/A_{280} ratios were measured spectrophotometrically using a Nanodrop One^C (Thermo Fisher Scientific), and samples were stored at -20 °C until downstream use.

2.6 Primer design and validation

2.6.1 Primer design

Specific primers for the 50 biomarkers were required to perform qPCR. The PlasmoDB (available at: https://plasmodb.org/plasmo/app) accession codes for the genes were used to obtain the mRNA Sequences (Introns spliced out; UTRs highlighted) of the 50 potential biomarkers and seryl-tRNA synthetase (PF3D7_0717700). Seryl-tRNA synthetase is used as it is a housekeeping gene that is constitutively expressed throughout the life cycle and is further referred to as the endogenous control gene (ECG) (123, 124). These sequences were then imported into Benchling (available at: https://benchling.com/editor), a cloud-based platform that was used for primer design using its primer Wizard application, based on Primer3. Primers were designed to have a predicted melting temperature (T_m) (as calculated by Benchling) as close to 60 °C as possible (range of 57.5 °C to 60.5 °C). Melting

© University of Pretoria

temperatures between primer pairs should preferably not differ more than 1-3 °C. *P. falciparum* parasites have an A-T rich genome, which meant that a wider range of temperatures needed to be considered. Primers were designed to ideally be between 15-30 base pairs (bp). The 3' end of the primers was adjusted to end on a Guanine/Cytosine (GC clamp) to allow for better binding of the primer to the template. When the primers predicted by the Wizard function could not be adjusted to meet the requirements, primers were manually designed according to the parameters. The primers were evaluated for the formation of secondary structures using the minimum Gibbs free energy (△G) of the Homodimer and Monomer functions and a cut-off ΔG < -37.7 kJ/mol at 50 °C was used (125).

2.6.2 End-point PCR and agarose gel electrophoresis for primer validation

To assess the primers T_a and their ability to produce a singular band of the expected size, a PCR was performed using cDNA obtained as per section 2.5.4. Each PCR sample contained 30 ng of cDNA template, 0.5 pmol/μL of the forward primer and 0.5 pmol/μL reverse primer (Inqaba Biotec, RSA) specific to each transcript, 1× standard Taq reaction buffer (New England Biolabs, USA), 200 μM dNTPs (KAPA Biosystems, USA), and 1.25 U/50 μL Taq DNA polymerase (New England Biolabs, USA). UltraPure™ DNase/RNase-free distilled water (Thermo Fisher Scientific, USA) was added to achieve reaction volumes of 10 μL. The reactions were subjected to 30 cycles of PCR under the following conditions: 95 °C initial denaturation for 5 min, denaturation at 95 °C for 30 s, annealing at 58 °C for 30 s, extension at 68 °C for 30 s, and a final extension at 68 °C for 5 min. The samples were stored at 4 °C until they could undergo electrophoretic analysis.

To visualise the PCR products and to establish whether products of expected band sizes were obtained, agarose gel electrophoresis was used. A 2 % (w/v) agarose gel was prepared by dissolving agarose (Thermo Fisher Scientific, USA) in 1× TAE (40 mM tris-acetate pH 8.0 (Sigma-Aldrich, USA) and 1 mM EDTA (Merck, USA)) assisted by heating of the solution in a microwave (LG). The samples were loaded into the solidified gel with 6× TriTrack DNA loading dye (10 mM Tris-HCl (pH 7.6), 0.03 % (w/v) bromophenol blue, 0.03 % (w/v) xylene cyanol FF, 0.15 % (w/v) orange G, 60 % (v/v) glycerol and 60 mM EDTA)) (Thermo Fisher Scientific, USA) at a final concentration of 1× TriTrack DNA loading dye. A GeneRuler 100 bp DNA ladder (500 ng) (Thermo Fisher Scientific, USA) was included for size reference. The gel was pre-stained using 0.75 μg/mL EtBr (Invitrogen, USA) and electrophoresis was performed at 100 V for approximately 30 min. The gel was then visualised using a Gel Doc XR+ imaging system (Bio-Rad, USA). Reactions that yielded multiple bands or no bands at all were eliminated from downstream experiments.

2.7 Compound treatment

Before compound treatment, cultures were synchronised twice (8 h apart) as per section 2.3 to ensure a tightly synchronised parasite population. The culture was left to recover for a life cycle (48 h) before compound treatment.

© University of Pretoria

The *in vitro* parasite reduction ratio (PRR) assay data and stage-specific activity were used to inform the treatment periods for each compound. PRR is a measure of parasite viability in response to drug treatment and provides insight into pharmacodynamic parameters, including the PRR value, lag phase (the period before the compound notably affects the parasite), and parasite clearance time (PCT) (clearance of 99.9 % of parasites) (126). Artemisinin and chloroquine are fast-acting with PRR ≥ 4, DSM265, pyrimethamine, MMV084, and M5717 have intermediate rate of action with 3 ≤ PRR < 4 and MMV048 and UCT943 are slow-acting with PRR < 3 or lag of > 48 h.

A ~6-9 hpi ring culture was treated at IC_{90} (obtained as per section 2.4) with the respective compounds and incubated for a time that fits their stage of activity and PRR (Table 2). For DFMO and SAHA that do not have PRR data, the time was customised to fit their peak activity stage (trophozoites). The starting parasitaemia was chosen to not induce stress due to proliferation and to ensure sufficient parasite starting material for RNA isolation. Treatments lasting 12 h and 24 h had a starting parasitaemia of ~7 %, while 30 h and 48 h treatments were done at 3-5 % parasitaemia. After incubation, the cultures underwent RNA isolation and cDNA synthesis.

Table 2. Activity profiles for each of the compounds in this study.

Compound	Stage of activity	Ref.	Parasite reduction ratio (PRR)	Ref.	Treatment period
Chloroquine	Trophozoites and early schizonts are most susceptible	(123)	PRR: 4.6 – 4.9 PCT: 29.4 – 31.1 h	(122)	12 h
Artemisinin	Rings, trophozoites and early schizonts	(124)	PRR: 8.0 – 8.7 PCT: 18.8 – 18 h	(122)	12 h
DFMO	Trophozoites	(112)	Not available	N/A	24 h
SAHA	Trophozoites	(104)	Not available	N/A	24 h
MMV048	Schizonts	(125)	PRR: 2.7	(125)	48 h
DSM265	Late trophozoites	(55)	PRR: 3.1 PCT: 85 h	(126)	30 h
Pyrimethamine	Early schizonts	(127)	PRR: 3.4 – 4.2 PCT: 51.3 – 59.4 h	(122)	30 h
MMV084	Schizonts	(56)	PKG inhibitors have PRRs similar to pyrimethamine	(56)	30 h
M5717	Treated rings only develop into abnormal trophozoites	(128)	Lag of 24-48 h, rapid killing after 48+ h of exposure	(128)	30 h
UCT943	Schizonts	(125)	PRR: 2.5	(125)	48 h

PRR = parasite reduction ratio, Ref. = reference, PCT = parasite clearance time, DFMO = difluoromethylornithine, and SAHA = suberoylanilide hydroxamic acid.

The parasite populations' stage distribution was determined by counting a minimum of three microscope fields, 1000 erythrocytes and 30 parasites per sample. Two-tailed unpaired Student's t-tests were applied to calculate statistical significance between samples, considering each microscope field separately.

© University of Pretoria

2.8 Real-time amplification of the genes of interest

cDNA (1.5 ng per reaction) was used with a final reaction volume of 10 μL to a 384-well plate. qPCR was performed using a QuantStudio 5 Dx real-time PCR system (Applied Biosystems, USA) using 1× PowerUp SYBR Green master mix (SYBR Green dye, Dual-Lock Taq DNA Polymerase, dNTPs with dUTP/dTTP blend, heat-labile uracil-DNA Glycosylase (UDG), ROX passive reference dye) (Thermo Fisher Scientific, USA)) with 0.5 pmol/μL forward primer and 0.5 pmol/μL reverse primer (Inqaba Biotec, RSA), respectively, specific to each transcript. Heat-labile UDG prevents carryover contamination, while the Dual-Lock Taq DNA Polymerase is a combination of two hot-start mechanisms to prevent premature activity. Plates were centrifuged to collect the contents at the bottom of the plate before insertion into the QuantStudio 5 Dx real-time PCR system (Applied Biosystems, USA). The cycling conditions were: UDG activation at 50 °C for 2 min, Dual-lock Taq DNA polymerase at 95 °C for 2 min, followed by 40 cycles of denaturation at 95 °C for 15 s, annealing at 58 °C for 2 min and extension at 72 °C for 1 min. Reactions were run in technical duplicate and a single biological repeat was done for each experiment.

Standard curves for each biomarker and the ECG were generated by plotting the mean quantitative cycle (Cq) values against the log (amount of cDNA (ng) present in the reaction) over a range of 5 cDNA quantities starting at 1.5 ng to 0.1 ng using a 2-fold dilution series. The slope of each curve was determined using a linear regression line of best fit. The individual primer efficiencies were calculated using the formula $E = 10^{(-1/slope)}$ from the curves. The formula to express efficiency as a percentage was: % Efficiency = (E – 1) x 100.

Two methods of relative quantification of expression were applied. The data were first analysed using the well-known $2^{-\Delta\Delta Cq}$ method (127), which assumes primer efficiencies close to 100 %. In this method, ΔC_q for the treated and untreated samples were calculated using Equation 1. This was followed by normalising the treated samples to the untreated control using Equation 2. The $\Delta\Delta C_q$ could then be converted to the fold change in gene expression using Equation 3. The data were finally represented as the $\log_2$FC in gene expression.

Eq. 1: $\Delta C_q = C_q(GOI) - C_q(ECG)$

Eq. 2: $\Delta\Delta C_q = \Delta C_q(\text{treated sample}) - \Delta C_q(\text{untreated control})$

Eq. 3: $\text{Fold change} = 2^{-\Delta\Delta Cq}$

The Pfaffl method of relative quantification was additionally included to evaluate the effect of differences in primer efficiencies on the qPCR-based platform as it considers the primer efficiency of a specific primer set and represents the change in gene expression as a ratio, defined in Equation 4 (128). This method was only applied to the set of five control compounds to investigate the effect of primer efficiency on relative quantification. The efficiency was calculated using the equation $E = 10^{(-1/slope)}$, this

© University of Pretoria

value was converted to a percentage when discussing efficiency but left as is when incorporated into the Pfaffl equation. The data were finally represented as the log_2FC (Log_2ratio) in gene expression.

Eq. 4: $$\text{Fold change ratio} = \frac{(E_{GOI})^{\Delta Cq_{GOI}(\text{control} - \text{sample})}}{(E_{ECG})^{\Delta Cq_{ECG}(\text{control} - \text{sample})}}$$

2.8.1 Data visualisation

The log_2FC values were visualised using SOMs created in R with the R package SupraHex (version 1.28.1), which uses a self-organising learning algorithm and only the input features to identify the underlying structure and patterns of the data (129). Clustering was performed on a distance matrix using median distance and average cluster linkage (average between all items). These maps were used to compress and cluster the data, reducing the dimensionality to make it easier to observe trends within the data that were obscured when using heatmaps. The data were capped at 1 and -1 Log_2FC and the SOMs were visualised as suprahexagonal chemo-transcriptomic fingerprints.

© University of Pretoria

Chapter 3: Results

3.1 *In vitro* culturing of asexual *P. falciparum* parasites

P. falciparum parasites (drug-sensitive NF54 strain) were cultured and used for downstream applications, which included proliferation inhibition studies, RNA isolation, and subsequent PCR and qPCR. The development of the parasites through the IDC was monitored using Rapi-Diff-stained thin blood smears to establish that the parasites progressed through the life cycle as expected (Figure 10). The acidic eosin dye component of the Rapi-Diff staining kit stains the parasite's chromatin pink, while the cytoplasm is stained blue by the basic methylene blue dye component (130).

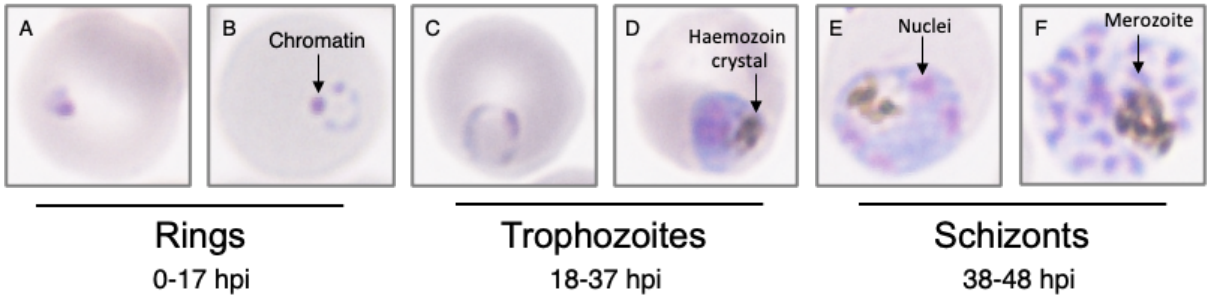


Figure 10. Development of IDC *P. falciparum* parasites over 48 h. Representative pictures of the progression of asexual IDC parasites through the life cycle. The stained thin blood smears were viewed with a light microscope using the 100× oil immersion objective. hpi = hours post-invasion.

A single IDC of the parasites lasts ~48 h, wherein the parasites progress through three distinct morphological stages. The cycle starts when a merozoite invades an erythrocyte (0 hpi) and develops into a ring-stage parasite (Figure 10A and B), with the older ring-stage parasite showing two distinctive dark pink dots (chromatin) common to *P. falciparum* parasites and a clear vacuole in the middle (Figure 10B). This is followed by the trophozoite-stage parasite (~18 hpi), early-stage trophozoites also show a clear vacuole and the definite fine ring shape of the ring stages starts to thicken and becomes more irregular as the parasite matures into trophozoites (Figure 10C). During the later trophozoite stage of development (30-37 hpi), the vacuole becomes less noticeable, and the parasites show the formation of haemozoin crystals (brown), a by-product of active haemoglobin digestion, as is evident in the second trophozoite picture (Figure 10D). Schizont-stage parasites (~38 – 48 hpi) are characterised by multiple pink dots (> 2N) as nuclear division commences (116). This is evident in the young schizont, where four nuclei (pink dots) are visible (Figure 10E). The mature schizont has already segregated into individually packaged merozoites, which each have a pink dot with blue cytoplasm around them (Figure 10F).

3.2 Establishing the *in vitro* activity of the test compounds

To allow downstream assessment of the compound-induced effect on gene expression, it was necessary to generate dose-response curves for each compound investigated using the SYBR Green I asexual proliferation assay. After 96 h of incubation with the compounds, which cover two IDCs, the IC_{50} and IC_{90} were determined. Only a single biological repeat was performed for each compound as

© University of Pretoria

the compounds used here have well-established activities against *P. falciparum* parasites and we only confirmed pre-existing activity data to ensure the correct concentration to be used in downstream experiments. Five compounds (artemisinin, chloroquine, DFMO, SAHA, and MMV048) were chosen from the van Heerden *et al*. (2021) paper as control compounds to test the platform as they are well-characterised compounds from different clusters in the dendrogram (Figure 11).

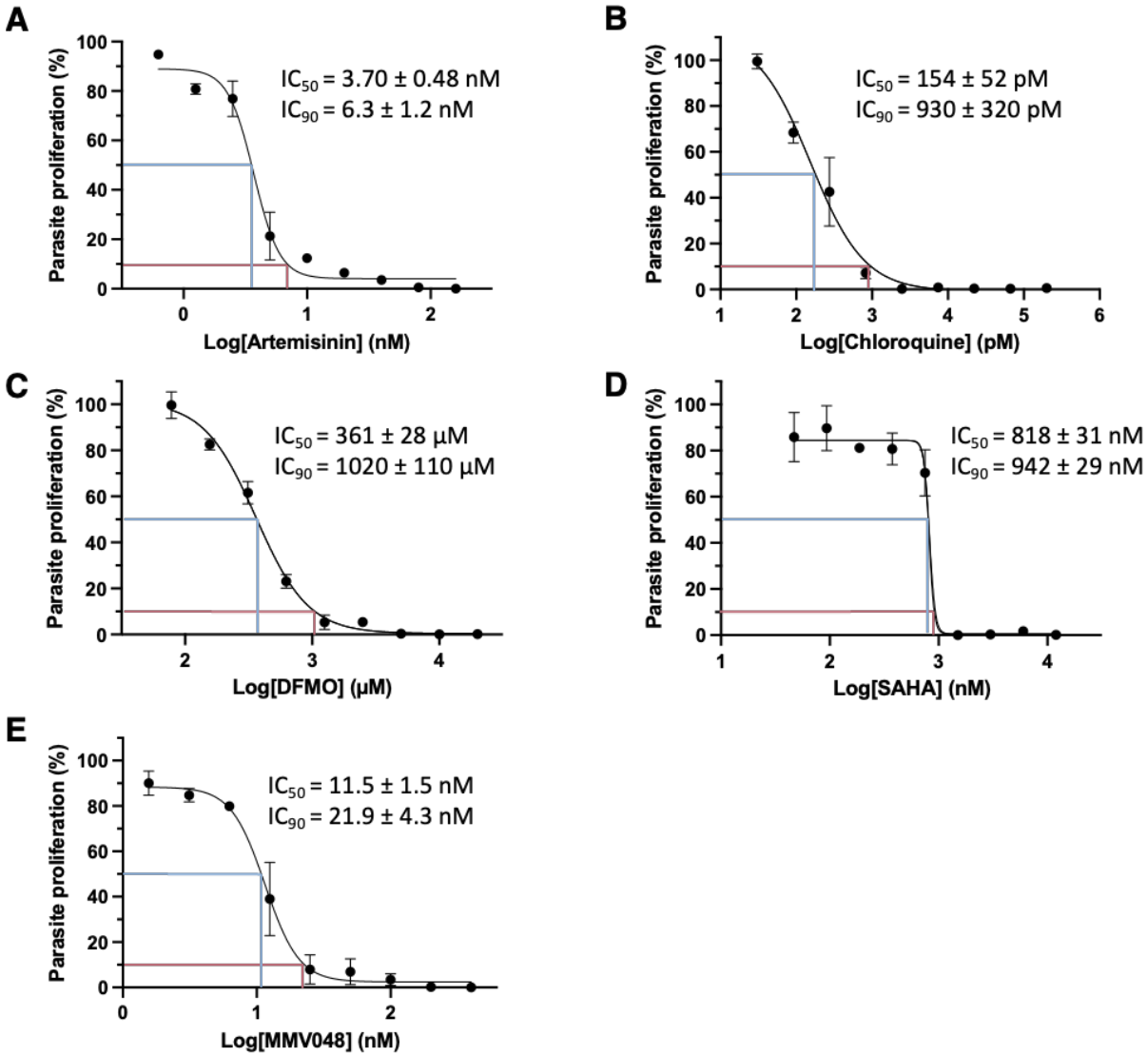


Figure 11. The activity of control compounds included in the rational selection 50 transcript model. The dose-response curves of **(A)** artemisinin, **(B)** chloroquine, **(C)** DFMO, **(D)** SAHA, and **(E)** MMV048 against intraerythrocytic asexual parasites. The red lines indicate the IC_{90} and the blue lines indicate the IC_{50}. The dose-response curves, IC_{50}, and IC_{90} values are provided by GraphPad Prism 10. Data are from one biological repeat (n = 1) and are represented as the mean ± SD. Where error bars are not visible, they fall within the symbol. DFMO = difluoromethylornithine, SAHA = suberoylanilide hydroxamic acid, IC_{50} = half maximal inhibitory concentration, and IC_{90} = 90 % of the maximum inhibition concentration.

The IC_{50} values for the control compounds featured in van Heerden *et al*. (2021) were compared to literature values (Table 3) and were found to be similar. The dose-response curves are, therefore, an accurate representation of the activity of the compounds *in vitro*.

© University of Pretoria

Table 3. Literature IC_{50} values for the control compounds.

Compound	Obtained IC_{50}	Literature IC_{50}	Reference
Artemisinin	3.70 ± 0.48 nM	2.2 nM	(131)
Chloroquine	154 ± 52 pM	10 ± 3 nM	(132)
DFMO	361 ± 28 μM	~560 μM	(116)
SAHA	818 ± 31 nM	~100 nM	(108)
MMV048	11.5 ± 1.5 nM	28 nM	(78)

DFMO = difluoromethylornithine and SAHA = suberoylanilide hydroxamic acid.

In addition to the control antimalarial compounds evaluated above, four antimalarial clinical candidates with different MoAs, DSM265 (targets DHODH) (133), pyrimethamine (targets DHFR) (134), MMV084 (targets PKG) (62), and M5717 (targets EF2) (62), were selected to further evaluate the performance of the platform. Additionally, UCT943 was also included to evaluate how the platform deals with compounds that share a target/MoA with another tested compound (MMV048). Their inhibitory concentrations were again verified (Figure 12).

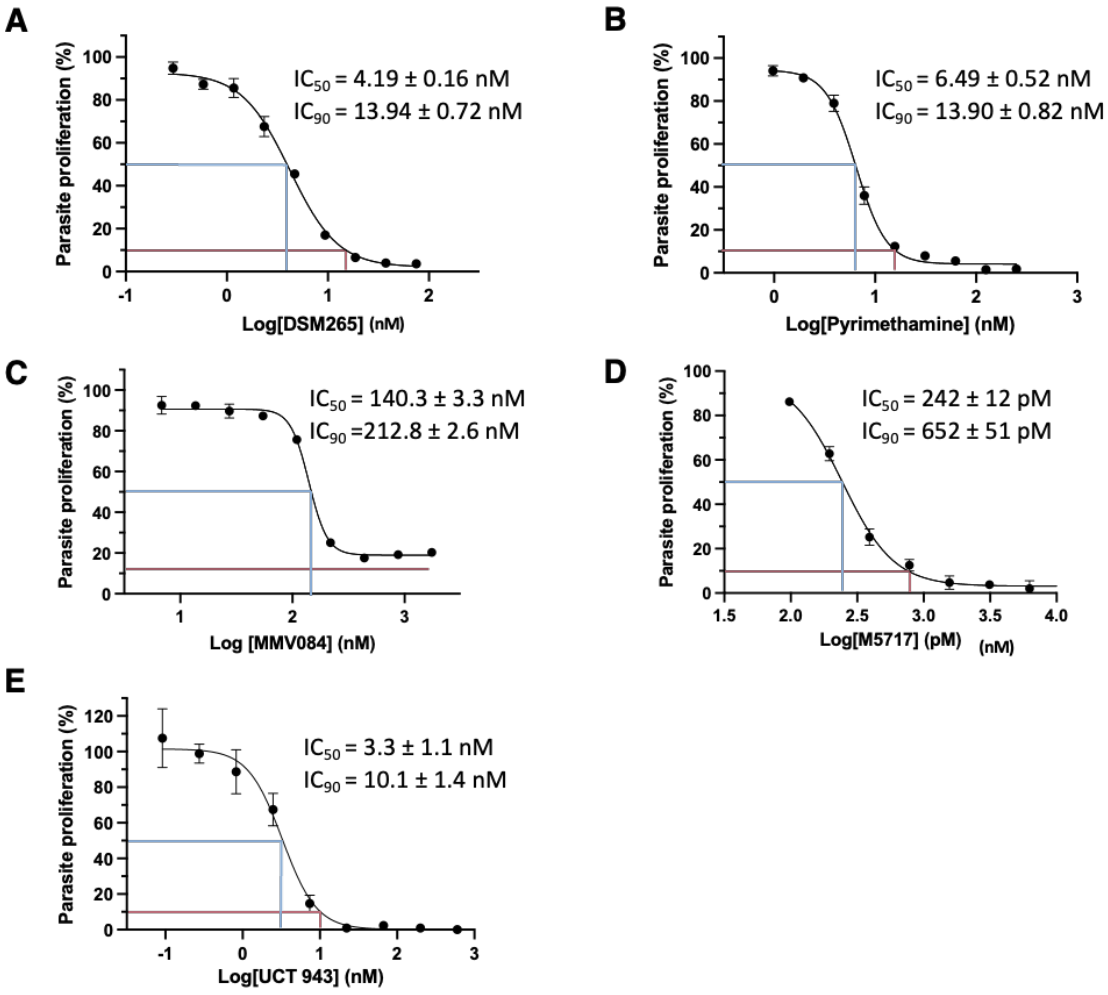


Figure 12. The activity of clinical candidates with known MoAs that are unfamiliar to the model. The dose-response curves of **(A)** DSM265, **(B)** pyrimethamine, **(C)** MMV084, **(D)** M5717, and **(E)** UCT943 against intraerythrocytic asexual parasites. The red lines indicate IC_{90} and the blue lines indicate IC_{50}. The dose-response curves, IC_{50}, and IC_{90} values are provided by GraphPad Prism 10. Data are from one biological repeat (n = 1) and

© University of Pretoria

are represented as mean ± SD. Where error bars are not visible, they fall within the symbol. IC_{50} = half maximal inhibitory concentration and IC_{90} = 90 % of the maximum inhibition concentration.

The IC_{50} values obtained for the clinical candidates with distinct MoAs are comparable to those obtained in the literature and could, therefore, be used to inform the concentrations required downstream (Table 4). The IC_{50} of UCT943 is 3.26 nM, which also closely resembles the activity in the literature of 5.4 nM against asexual *P. falciparum* parasites.

Table 4. Literature IC_{50} values for the clinical candidates.

Compound	Obtained IC_{50}	Literature IC_{50}	Reference
DSM265	4.19 ± 0.16 nM	8.9 nM (0.01 μg/mL)	(135)
Pyrimethamine	6.49 ± 0.52 nM	< 10 nM	(136)
MMV084	140.3 ± 3.3 nM	109 nM	(62)
M5717	242 ± 12 pM	~810 pM	(62)
UCT943	3.26 ± 1.1 nM	5.4 nM	(137)

3.3 NucleoZOL allows for efficient RNA isolation and cDNA synthesis

A critical first step when building the qPCR-based biomarker transcription analysis platform was to optimise the isolation of RNA from small amounts of parasite starting material to obtain a large enough yield for downstream qPCR. NucleoZOL uses a no-chloroform and no-phase separation approach to simplify RNA isolation and has successfully extracted RNA from a wide variety of sources, including bacteria, yeasts, plants, cultured cells, and human tissue. The singular phase RNA isolation provided a rapid and scalable procedure, minimising the effect of technical variability between samples introduced during RNA isolation. To evaluate the use of NucleoZOL for RNA isolation from *P. falciparum* parasites, the yield, purity, and integrity of the RNA were compared to an established method that uses a Qiagen RNeasy mini kit (Table 5).

Table 5. Comparison between the NucleoZOL and RNeasy methods of RNA isolation.

Sample	Concentration	Total RNA yield	A_{260}/A_{280} ratio	A_{260}/A_{230} ratio
RNeasy Mini kit	1.28 μg/μL	37 μg	2.27	2.29
NucleoZOL	2.22 μg/μL	64 μg	2.19	2.69

RNA was isolated from 10 and 20 mL IDC *P. falciparum* parasite cultures (NucleoZOL and RNeasy, respectively) of ~10 % parasitaemia. The cultures were majority trophozoites at ~33-37 hpi, which are the most transcriptionally active and provided the greatest chance for high yields of RNA. The NucleoZOL sample had an RNA concentration 1.7-fold greater than the RNeasy sample (1.28 and

© University of Pretoria

2.22 μg/μL, respectively). Therefore, the total yield from the NucleoZOL method was greater than that of the RNeasy method (64 μg *vs.* 37 μg, respectively).

The purity of the samples obtained for both methods was assessed using the A_{260}/A_{280} and A_{260}/A_{230} ratio. The A_{260}/A_{280} ratio assesses the amount of protein contamination, measured at 280 nm, in the samples and should ideally be above 1.9. The A_{260}/A_{230} assesses the presence of contaminants like guanidine, phenol, or other organic substances at 230 nm, and this value should be above 1.8. The NucleoZOL sample had a lower A_{260}/A_{280} ratio than the RNeasy sample (2.19 and 2.27, respectively) but a higher A_{260}/A_{230} ratio (2.69 and 2.29, respectively). The samples from the NucleoZOL and RNeasy methods, however, both had purity ratios well above the required values, indicating both samples are pure (Table 5).

The integrity of the isolated RNA was subsequently determined using agarose gel electrophoresis. The RNA was denatured (heated to 70 °C) and ran on an agarose gel (1.5 % (w/v) agarose) (Figure 13). Intact RNA shows two distinct bands on an agarose gel, which correspond to the ribosomal RNA (rRNA) subunits (28S rRNA and 18S rRNA) present in eukaryotes.

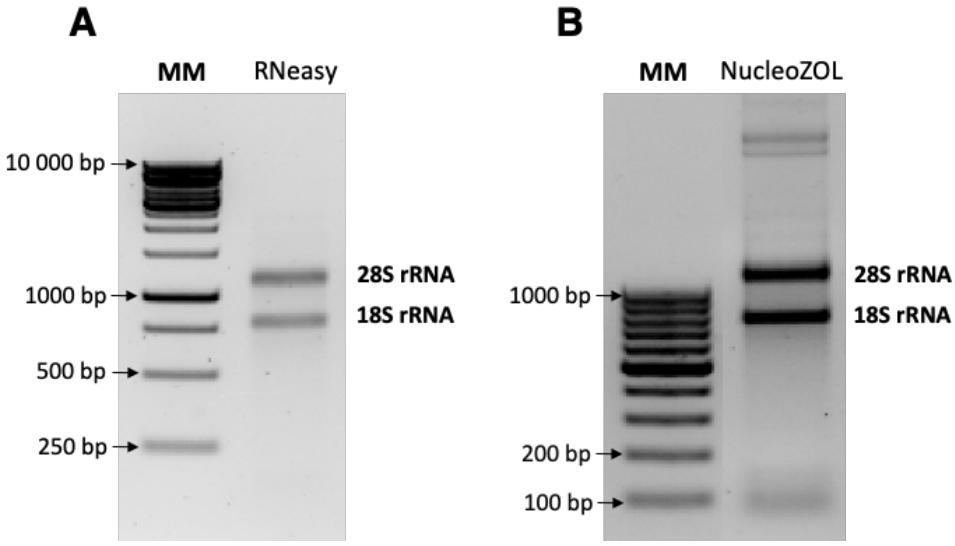


Figure 13. Evaluation of the established RNeasy method *vs.* the NucleoZOL method of RNA isolation. (A) The RNeasy mini kit RNA samples with the 28S and 18S rRNA indicated. **(B)** The NucleoZOL RNA samples with the 28S and 18S rRNA indicated. The RNA and a 1 kb ladder (A) and 100 bp ladder (B) (MM) (Thermo Fisher Scientific, USA) were separated by a 1.5 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). bp = base pairs.

Both the NucleoZOL and RNeasy samples show distinct bands that correspond to 28S and 18S rRNA (Figure 13A and B), indicating both samples are intact. The faint smearing observed for the NucleoZOL sample does not indicate degradation and is most likely due to using non-denaturing agarose gels and a large amount of RNA (Figure 13B). There were fainter bands present at the top of the NucleoZOL RNA sample, and these are potentially long non-coding RNA found in *P. falciparum* parasites seemingly characteristic of RNA isolation using NucleoZOL (138). It is not likely to be genomic DNA contamination

© University of Pretoria

as that is typically observed as a singular band. The additional faint band at the bottom of the NucleoZOL sample represents a low concentration of low molecular weight RNAs (small RNAs) (139).

3.3.1 RNA isolated using NucleoZOL can be converted to functional cDNA

The ability of the RNA to be converted to cDNA was evaluated by end-point PCR using the primers for the ECG (seryl-tRNA synthetase) as it is considered constitutively expressed throughout the *P. falciparum* parasites' IDC (140).

The RNeasy and the NucleoZOL samples produced bands of the expected sizes (~160 bp), indicating that the cDNA is functional (Figure 14). The band in the no-template control (NTC) lane (< 100 bp) corresponds to primer dimers formed in the absence of cDNA template, a band of the same size is faintly present for the RNeasy samples. This means that NucleoZOL provides a rapid and scalable approach for RNA isolation from small amounts of *P. falciparum* parasites. Therefore, this method of RNA isolation was applied to all further samples.

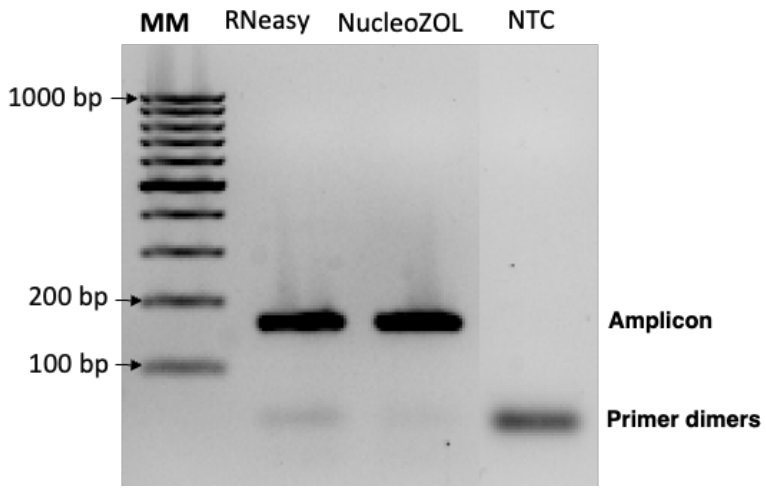


Figure 14. Functionality evaluation of cDNA synthesised from samples from both methods of RNA isolation. The cDNA and a 100 bp ladder (MM) (Thermo Fisher Scientific, USA) were separated by a 1.5 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). The endogenous control gene was used to evaluate the RNA's ability to produce functional cDNA. cDNA was obtained using RNA from the RNeasy mini kit and NucleoZOL, respectively. NTC = no-template control and bp = base pairs.

3.4 The RNA is of sufficient quality for downstream experiments

IDC *P. falciparum* parasite cultures (6-9 hpi) were treated with the respective compounds and RNA was isolated at the time points indicated in Figure 15. For each incubation period (12, 24, 30, or 48 hours post-treatment (hpt)) an untreated control is included as it is used for normalisation purposes in qPCR to ascertain a compound-induced effect compared to an untreated population at the same developmental stage. Different time points were chosen based on the activity profiles of the compounds (Table 2, section 2.7), a single time point could not be chosen because the compounds act on different developmental stages in the parasites' IDC.

© University of Pretoria

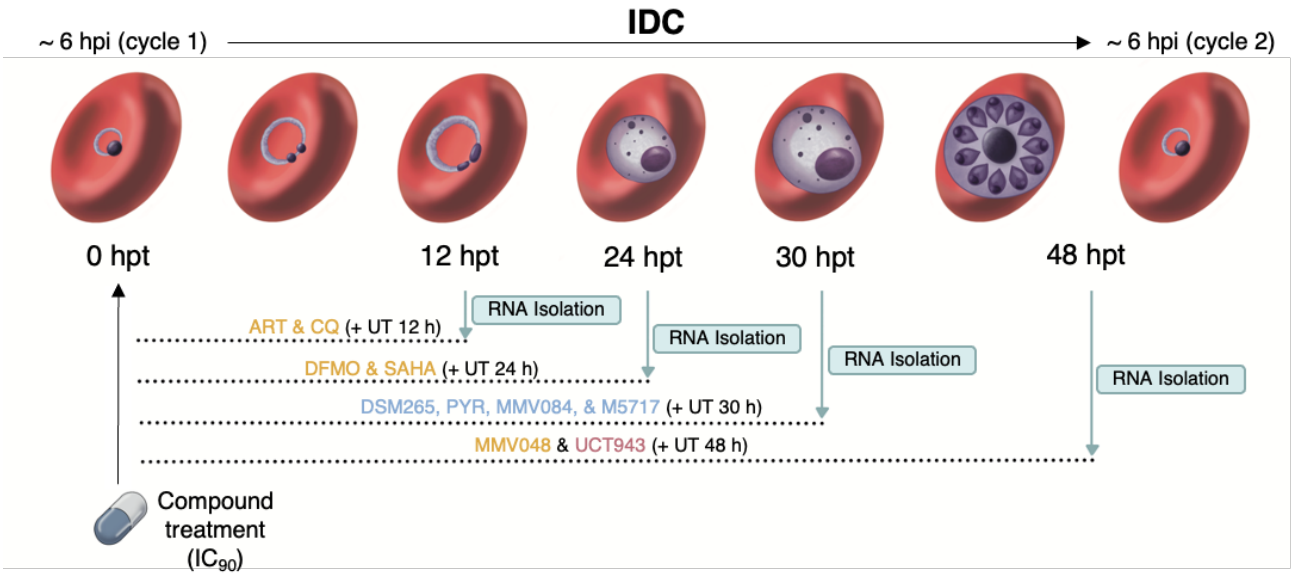


Figure 15. Schematic overview of compound treatment and subsequent RNA isolation time points. Compounds in yellow are control compounds, compounds in blue are clinical candidates included to extend the platform and UCT943, indicated in red, shares a target with MMV048 and was added to evaluate how the platform deals with shared MoAs. This figure was created using Procreate (https://procreate.art). IDC = intraerythrocytic development cycle, hpi = hours post-invasion, hpt = hours post-treatment, UT = untreated control, IC_{90} = 90 % of the maximum inhibition concentration, and MoA = mode of action.

3.4.1 Control compounds

All RNA samples for the control compounds were spectrophotometrically evaluated (as per section 2.5.3) to obtain their concentrations and purity ratios (A_{260}/A_{280} and A_{260}/A_{230}) to ensure the RNA is fit for downstream experiments (Table 6).

Table 6. Control compound and untreated control samples' RNA concentrations and purity values.

Sample	Concentration (ng/μL)	A_{260}/A_{280}	A_{260}/A_{230}
Untreated 12 h	24	1.34	0.22
Artemisinin	15	1.35	0.37
Chloroquine	49	1.69	0.47
Untreated 24 h	1420	2.18	2.45
DFMO	134	2.03	0.37
SAHA	491	2.16	2.71
Untreated 48 h	1280	2.14	2.65
MMV048	1547	2.17	2.46
Untreated 48 h (for ~6× IC_{90})	1061.9	2.12	2.88
MMV048 (~6x IC_{90})	532.0	2.08	2.60

DFMO = difluoromethylornithine and SAHA = suberoylanilide hydroxamic acid.

The RNA samples, except for the untreated control sample at 12 h and the artemisinin and chloroquine treatment samples, showed purity ratios greater than 1.9 and 1.8, respectively, and high RNA concentrations (Table 6). In the cases of treatment with artemisinin and chloroquine (and their untreated control), RNA isolation was performed on early-stage trophozoites that are less transcriptionally active,

© University of Pretoria

contributing to lower RNA yields. It has been shown that lower concentrations of nucleic acids do cause decreased purity ratios (141). DFMO also had a lower RNA concentration (134 ng/μL) and showed a similar trend for the A_{260}/A_{230} ratio. The integrity of the RNA samples was maintained, as was evident by the two distinctive bands representing 28S and 18S rRNA for each sample (Figure 16).

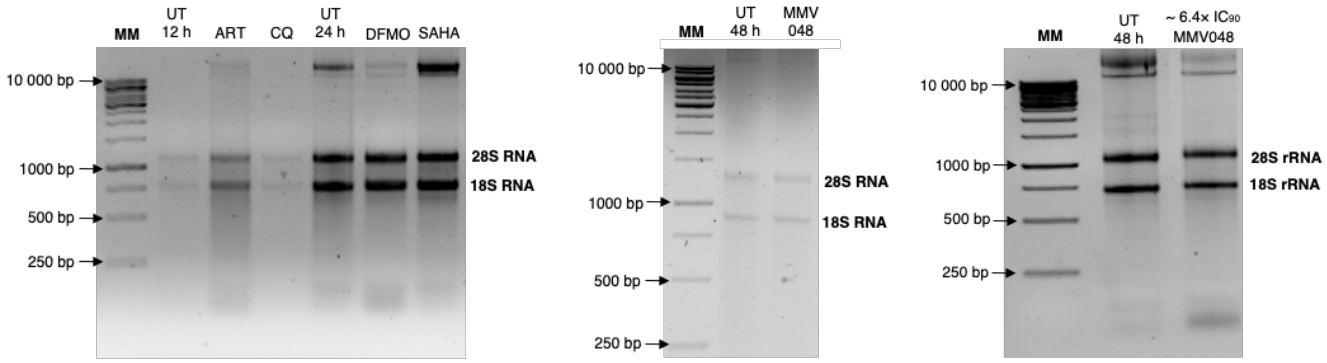


Figure 16. RNA integrity evaluation of the control compound samples. The RNA and 1 kb ladder (MM) (Thermo Fisher Scientific, USA) were separated by a 1.5 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). The 28S and 18S rRNA bands are indicated. UT = untreated and bp = base pairs.

3.4.2 Clinical candidates

The RNA samples for the clinical candidate treatments were also spectrophotometrically evaluated (as per section 2.5.3) in the same manner as for the control compound treatments (Table 7).

Table 7. Clinical candidate and untreated control samples' RNA concentrations and purity values.

Sample	**Concentration** (ng/μL)	**A_{260}/A_{280}**	**A_{260}/A_{230}**
Untreated 30 h	1244	2.18	2.80
DSM265	635	2.19	2.78
Pyrimethamine	1231	2.10	2.70
MMV084	1377	2.10	2.53
M5717	577	2.18	2.74
Untreated 48 h	1280	2.14	2.65
UCT943	1382	2.15	2.63

The A_{260}/A_{280} and A_{260}/A_{230} ratios of the untreated control (30 h) parasites RNA sample and the compound-treated parasites RNA samples were acceptable and deemed free of contaminants. The integrity of the samples was evaluated using agarose gel electrophoresis (Figure 17). All RNA samples showed two distinctive bands representative of the 18S and 28S rRNA, and the RNA was, therefore, intact and could be used for downstream qPCR evaluation.

© University of Pretoria

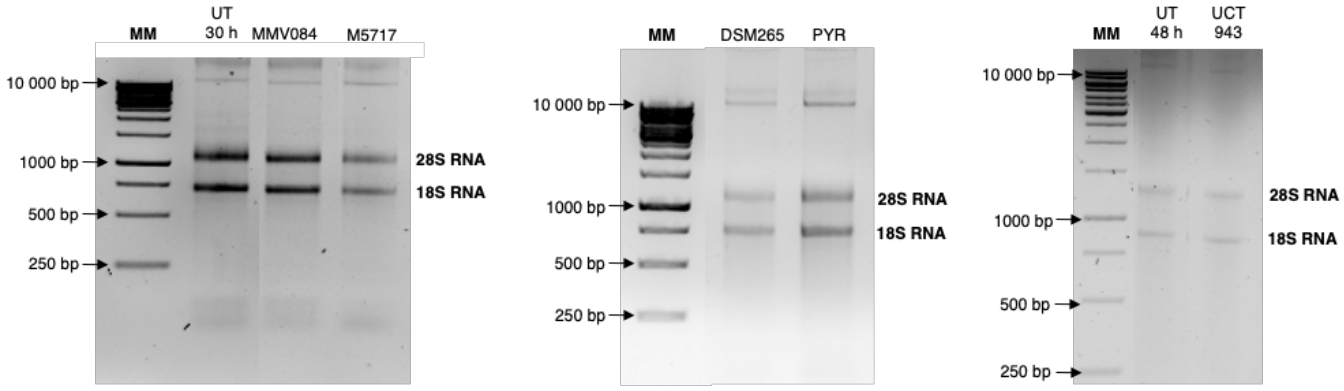


Figure 17. RNA integrity evaluation of the clinical candidate treatment samples. The RNA and 1 kb ladder (MM) (Thermo Fisher Scientific, USA) were separated by a 1.5 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). The 28S and 18S rRNA bands are indicated. UT = untreated, bp = base pairs, and PYR = pyrimethamine.

3.5 Primers for the predicted biomarkers were designed and validated for qPCR analysis

3.5.1 Primer design

To use primers for qPCR in parallel, under uniform conditions, it was important to design them to fall within set parameters. The primers were all designed to have amplicons of roughly the same size, 80-200 bps, as large variations in size can lead to variations in the efficiency of amplification in qPCR. Another important parameter was the T_m of the primers. However, *P. falciparum* parasites have an A-T rich genome that presented some challenges for T_m. The primers were designed to have a 1-3 °C difference in T_m between all primer pairs and a < 1 °C difference in T_m between primers of the same set (Figure 18). Minimising differences in T_m allowed for uniform conditions between genes when doing qPCR, which should decrease variability in amplification. A full list of primer sequences and other primer-specific information is available in Supplementary Table 1.

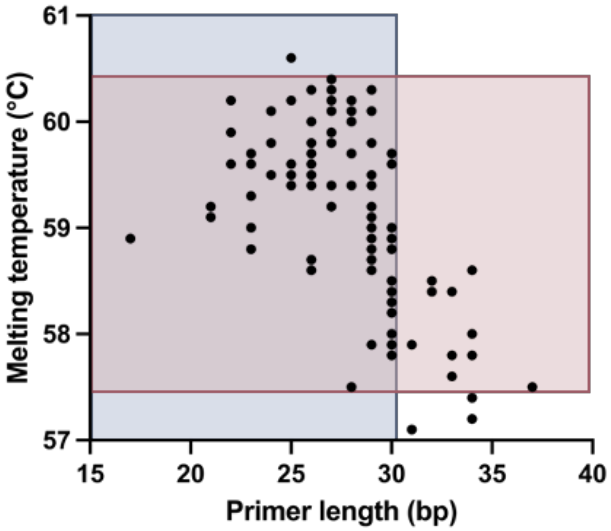


Figure 18. Key primer design factors. Indicated using the pink block is the desired range of T_m (57.5 – 60.5 °C) and indicated using the blue block is the desired range of primer length (15-30 bp). The overlap between the blocks shows the ideal range of both parameters.

© University of Pretoria

The optimal primer length of 15-30 bp (indicated in blue) was in some cases exceeded in favour of the primer having a T_m that falls within the specified range (indicated in pink) as T_m has a greater effect on the efficiency of amplification compared to primer length (Figure 18). This was the case with nine primers outlined in Table 8. Three primers fell outside both parameters (transcript 20 forward primer, transcript 42 forward primer, and transcript 47 reverse primer), resulting from the primer length already exceeding the maximum and the T_m not increasing closer to the minimum required T_m, even with further lengthening (Table 8).

Table 8. Primers that do not fall within the parameters of T_m and primer length.
Indicated in red is the parameter not within the specified range of 57.5 – 60.5 °C or 15-30 bp.

Primer	T_m	Length
Biomarker 3 reverse	58.5	32
Biomarker 18 forward	57.8	34
Biomarker 18 reverse	58	34
Biomarker 20 forward	57.4	34
Biomarker 26 forward	58.6	34
Biomarker 26 reverse	57.6	33
Biomarker 37 forward	57.9	31
Biomarker 40 reverse	58.4	33
Biomarker 42 forward	60.6	25
Biomarker 43 reverse	57.8	33
Biomarker 45 forward	57.1	31
Biomarker 47 forward	57.5	37
Biomarker 47 reverse	57.2	34
Biomarker 49 reverse	58.4	32

The $\triangle G$ of the homodimer is a measure of the probability that a primer will form a dimer between the two identical primers, while the $\triangle G$ of the monomer measures the probability the primer will fold on itself and have secondary structure formation at a given temperature. Each of the primers was evaluated for their probability of forming secondary structures that could impede the amplification process. Out of the 102 primers, 99 primers fell within the set parameter for the $\triangle G$ of the homodimer and the monomer (Figure 19).

© University of Pretoria

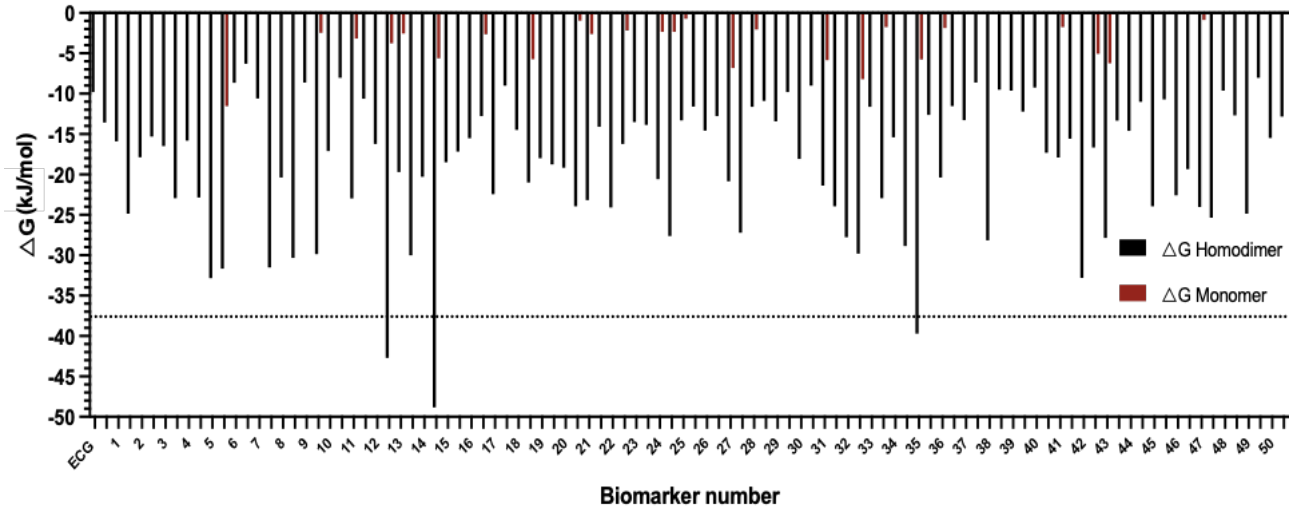


Figure 19. Gibbs free energy (△G) of the primers. The △G of the homodimer is indicated in black and the △G of the monomer in red for each primer. The dotted line indicates the cut-off (-37.7 kJ/mol).

The △G of the monomer for the primers, on average, were in the range of 0 to -11.5 kJ/mol, indicating the primers are not likely to fold on themselves. Primer dimer formation is more likely, as is seen by the △G of the homodimer for the primers, which were in the range of -6.3 to -48.8 kJ/mol. Some primers (biomarker 12 reverse primer, biomarker 14 reverse primer, and biomarker 35 forward primer) fell outside the design parameter for the △G of the homodimer (< -37.7 kJ/mol) suggested for qPCR primers as these were the only suitable primers for those biomarkers. All the primers were evaluated using PCR, followed by qPCR, as it was deemed more important to keep as many biomarkers included in the set of 50 predicted biomarkers as possible.

3.5.2 Primer validation

The primers' specificities were evaluated using end-point PCR to determine whether the primer sets could function under uniform conditions. The primers were examined for their ability to produce a single amplicon of the expected size. The T_a of each primer was calculated using T_a = 69.3 °C + 0.41(GC content) – 650/primer length) as it considers the GC content and the number of nucleotides in the primer (142). The primer sets that had a T_a of < 64 °C (for the forward primer) were initially tested at a T_a of 58 °C while primer sets that had a T_a of > 64 °C (for the forward primer) were initially tested at a T_a of 60 °C (Figure 20). Primer sets for biomarkers 14 and 44 were tested at 58 °C and 60 °C as they were close to 64 °C.

© University of Pretoria

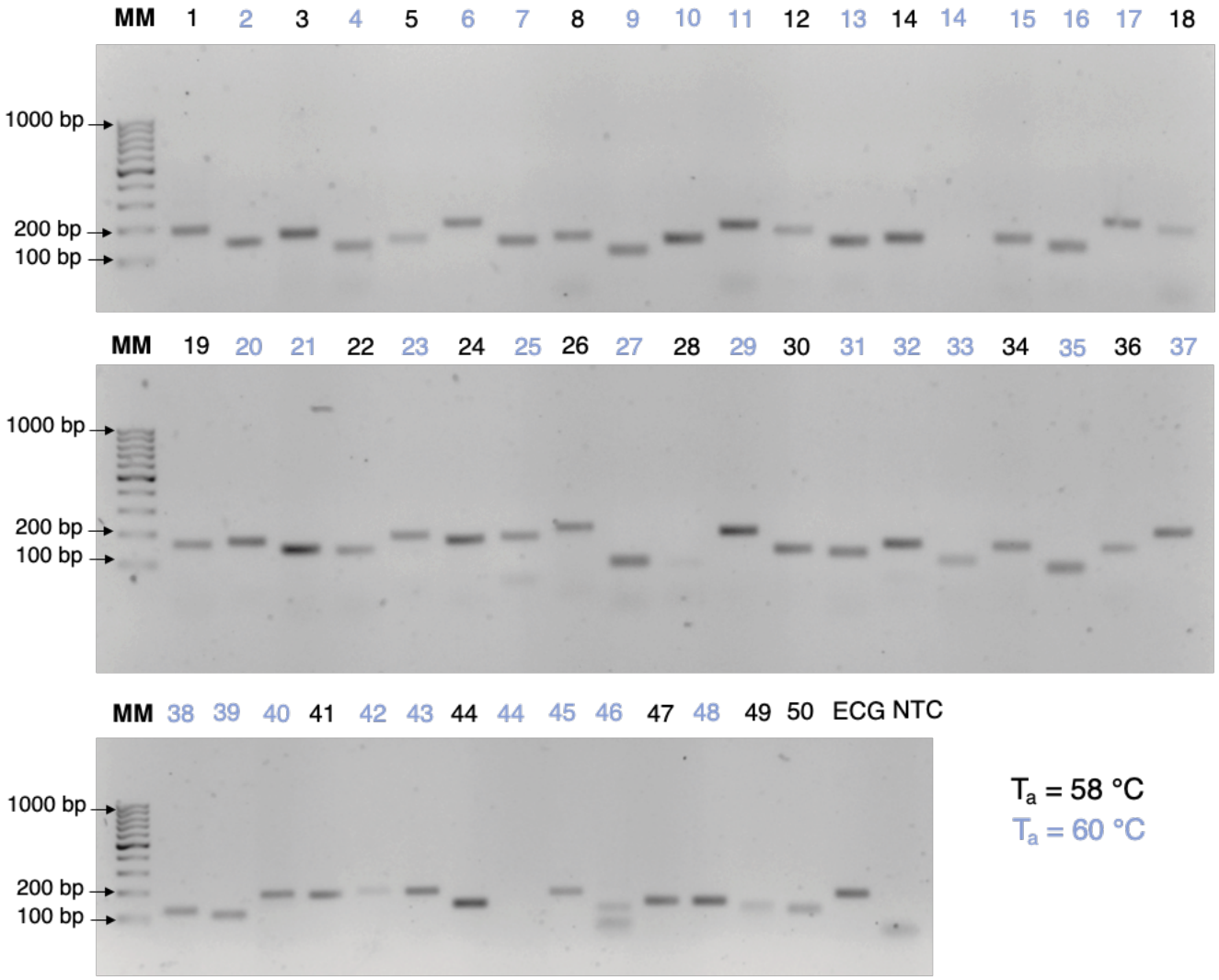


Figure 20. Primer validation at a T_a of 58 °C and 60°C, respectively. The DNA amplicons and 100 bp ladder (MM) (Thermo Fisher Scientific, USA) were separated by a 2 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). In black is primer sets tested at a T_a of 58 °C, while blue indicates primer sets tested at a T_a of 60 °C. ECG = endogenous control gene, NTC = no-template control for ECG, T_a = annealing temperature, and bp = base pairs.

Most of the primer sets produce single bands at their respective T_a. However, at a T_a of 60 °C (blue biomarker numbers), some primer sets did not produce a single band (biomarker 46) or any amplicon at all (biomarker 14 and 44 at 60 °C). A single T_a would have allowed all reactions to be run under uniform qPCR conditions to decrease variability between reactions. For this reason, all primers were re-evaluated at a T_a of 58 °C (Figure 21).

© University of Pretoria

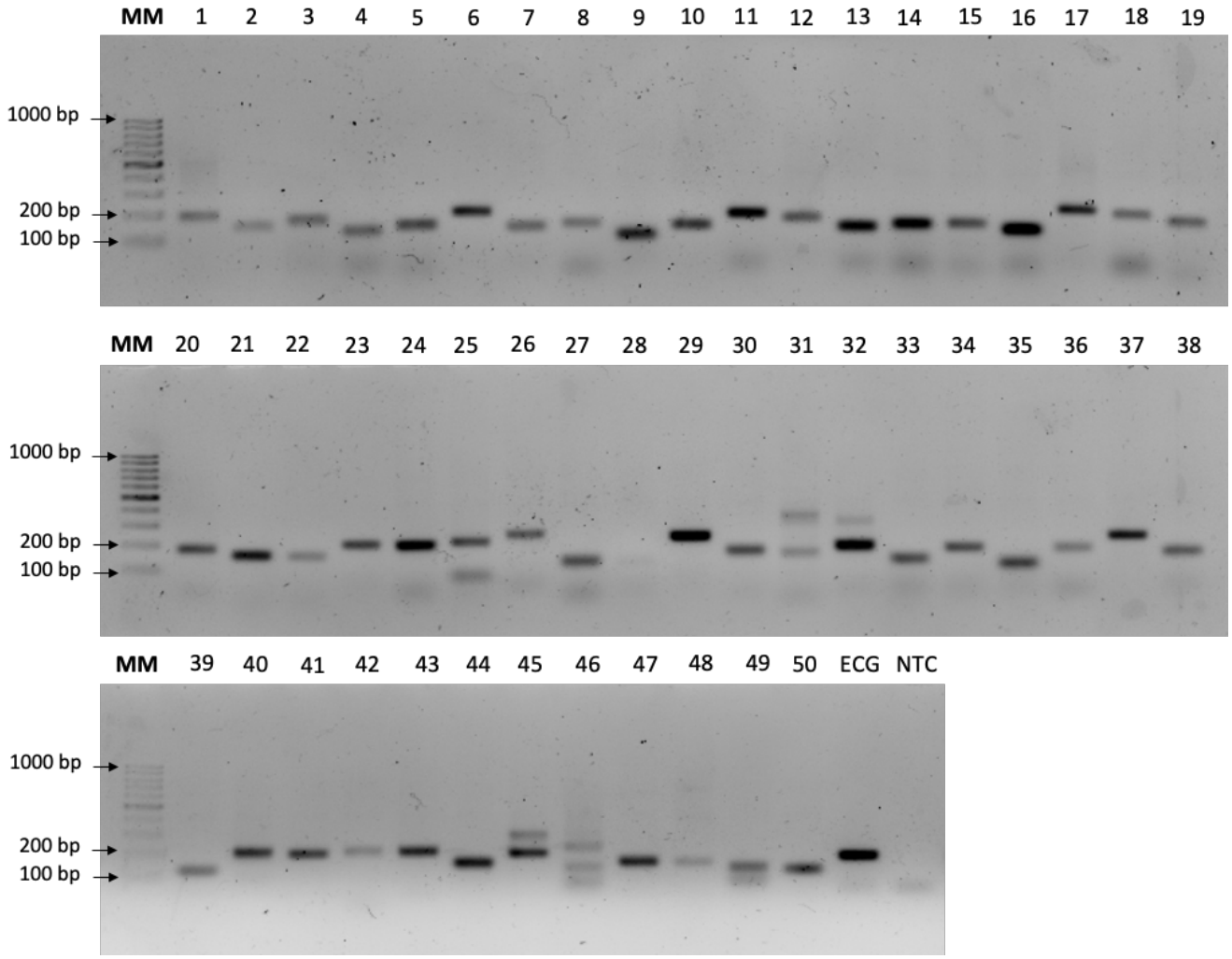


Figure 21. Primer validation at a uniform T_a of 58 °C. The DNA amplicons and 100 bp ladder (MM) (Thermo Fisher Scientific, USA) were separated by a 2 % (w/v) agarose gel and visualised using EtBr (0.75 μg/mL). ECG = endogenous control gene, NTC = no-template control for ECG, and bp = base pairs.

At a T_a of 58 °C, primers for 45 of the biomarker sets and the ECG showed a single amplification band, indicating that only one product of the expected size is amplified. Primer sets for biomarkers 31, 32, 45, and 46 showed multiple bands and, therefore, multiple amplicons. This led to these biomarkers being excluded from further investigation. The primer set for biomarker 28 produced no amplicon and was also excluded from further investigation. Since the primers will be investigated further using qPCR, it was not concerning that some amplicon bands were fainter than others. Only primers that produced multiple bands or no bands were not examined further.

3.6 Primer sets efficiency and specificity

In addition to evaluating the primer sets' specificity using end-point PCR, a melt curve analysis was also performed using qPCR. Melt curves provide a means of detecting any non-specific amplification. The ECG and the remaining primer sets produced single melt curve peaks, which indicated a single amplicon is being amplified (Figure 22 and Supplementary Table 3). This meant that no additional primer sets needed to be excluded based on the melt curve data.

© University of Pretoria

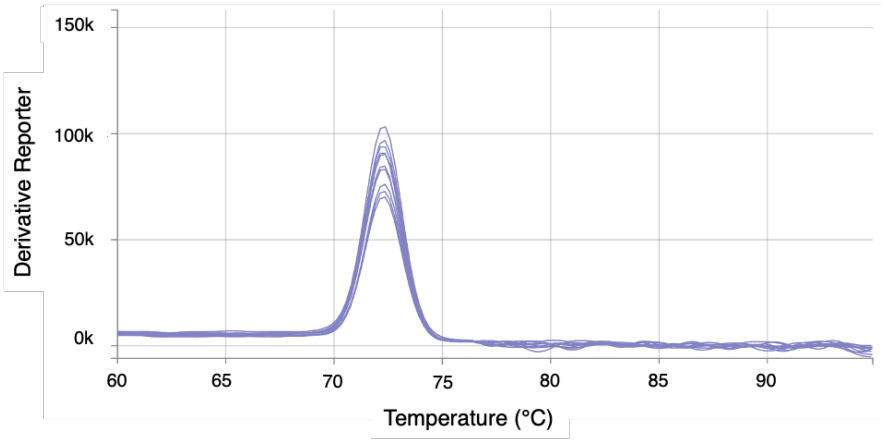


Figure 22. Melt curve analysis of the ECG primers using qPCR. As an example, the ECG was used as a representative of the melt curves to indicate specificity towards the intended target.

Primer efficiency is an important factor that provides insight into the quantification accuracy, quality, and reliability of the qPCR as it influences the replication efficiency. In qPCR, the template should ideally double in quantity every cycle, which indicates an efficiency of 100 %. The $2^{-\Delta\Delta Cq}$ method assumes that both primer sets have an efficiency equal to 100 % and primers should, therefore, be as close to 100 % efficient as possible. Primer efficiency is determined using standard curves with the slope of the standard curve converted into a percentage efficiency for the individual primer sets (Figure 23).

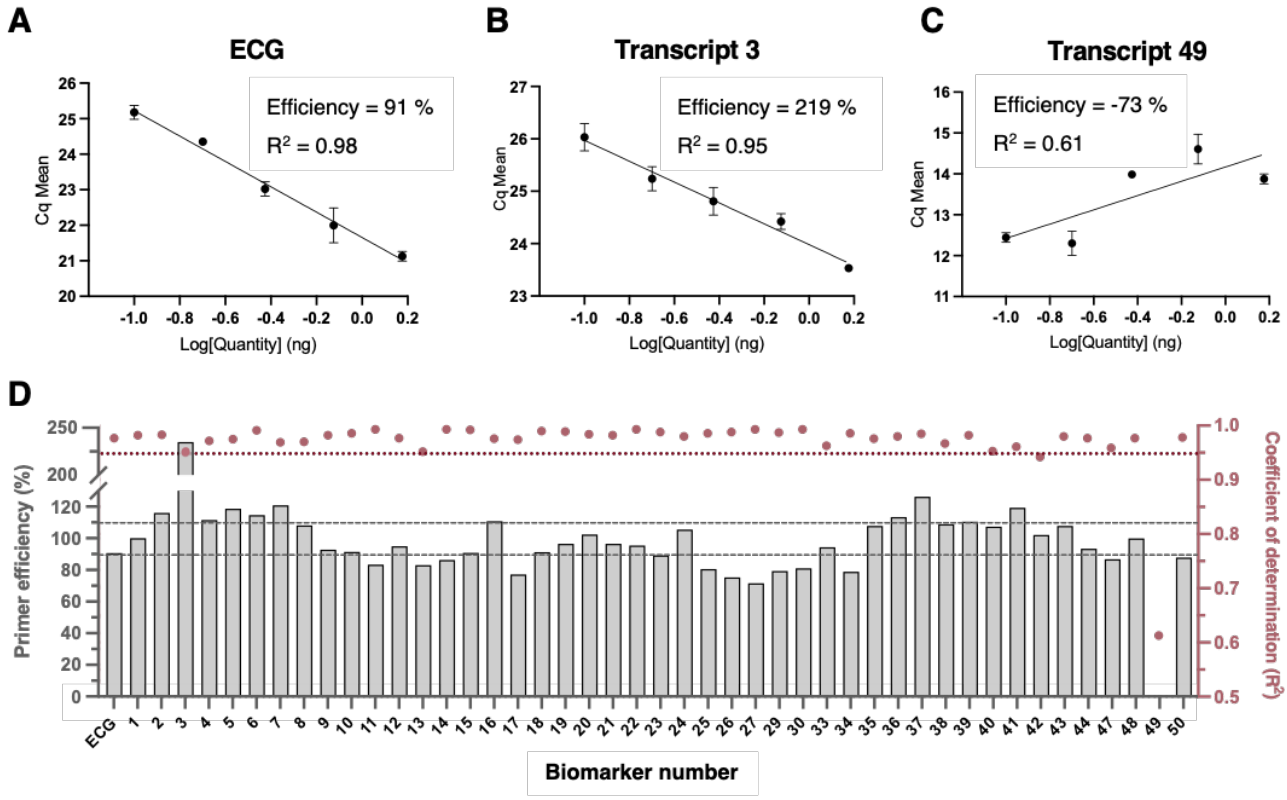


Figure 23. Evaluation of primer efficiencies using standard curves. The standard curves of the primers for the ECG **(A)** with preferable efficiency, biomarker 3 **(B)** with too high efficiency, and biomarker 49 **(C)** with too low efficiency. **(D)** The primer efficiencies for the full set of biomarkers are on the left y-axis (the grey dotted lines on the graph indicate the literature-suggested range of primer efficiencies) and the coefficient of determination (R^2) for each standard curve is on the right y-axis (indicated by the red line is 0.95). Data are from one biological repeat (n = 1) and A, B and C are represented as the mean ± SD. Where error bars are not visible, they fall within the symbol. ECG = endogenous control gene.

© University of Pretoria

The curve for the ECG indicated a preferable primer efficiency of 91 % and an R^2 of 0.98 (Figure 23A). By contrast, biomarker 3 had a primer efficiency almost double that of any other biomarker at 219 % efficiency, this seemed to be independent of technical variability (R^2 = 0.95) and it was also not due to non-specific amplification (Figure 23B). A BLAST search using the primers for biomarker 3 showed that the sequence of biomarker 3 was located within the untranslated region of another gene (PF3D7_0206800), potentially explaining the abnormally high primer efficiency. Biomarker 49 had a negative efficiency, possibly because of a poor linear fit (R^2 = 0.61) (Figure 23C). The remaining standard curves all showed good linearity with coefficients of determination (R^2) > 0.95 (indicated by the red dotted line), except for primer set 42 (R^2 = 0.94) (Figure 23D). The remainder of the standard curves are available in Supplementary Table 3. The grey dotted lines on this graph (Figure 23D) indicate the literature-suggested primer efficiency range of 90-110 %. However, only primer sets that were very far out of the range, biomarker 3 and biomarker 49, were excluded from further analysis and optimisation.

3.7 Evaluation of the chemo-transcriptomic fingerprints of the selected biomarkers

van Heerden *et al.* (2021) defined 50 predictive biomarkers to allow for chemo-transcriptomic fingerprinting of antimalarial compounds. To evaluate if the removal of the seven biomarkers that were identified as performing poorly in the qPCR would affect the chemo-transcriptomic fingerprints, the microarray data presented by van Heerden *et al.* (2021) were re-evaluated. SOMs were again used to generate chemo-transcriptomic fingerprints for the original set of 50 predicted biomarkers (Figure 24A), and this was compared to fingerprints generated with a set of 43 biomarkers that remained after performing the quality checks described above. (Figure 24B). In each case, the resultant chemo-transcriptomic fingerprints were hierarchically clustered to identify groupings of compounds with similar/related MoAs.

The grouping of the nodes (smaller hexagons) within the larger chemo-transcriptomic fingerprints of the compound-treated parasites are not evidently changed between Figure 24A and B and showed barely any visual distinction between the original set of 50 biomarkers and the set of the remaining 43 biomarkers. This was expected as a single node can contain multiple biomarkers with similar or the same expression, and the removal of some biomarkers should not detrimentally affect the chemo-transcriptomic fingerprint as the remaining biomarkers should compensate for the removed biomarkers. This showed that the exclusion of the seven biomarkers does not have a detrimental effect on the chemo-transcriptomic fingerprints' ability to distinguish between different compound treatments and it will, therefore, not affect the platform's ability to differentiate between compounds and their respective MoAs using chemo-transcriptomic fingerprints.

© University of Pretoria

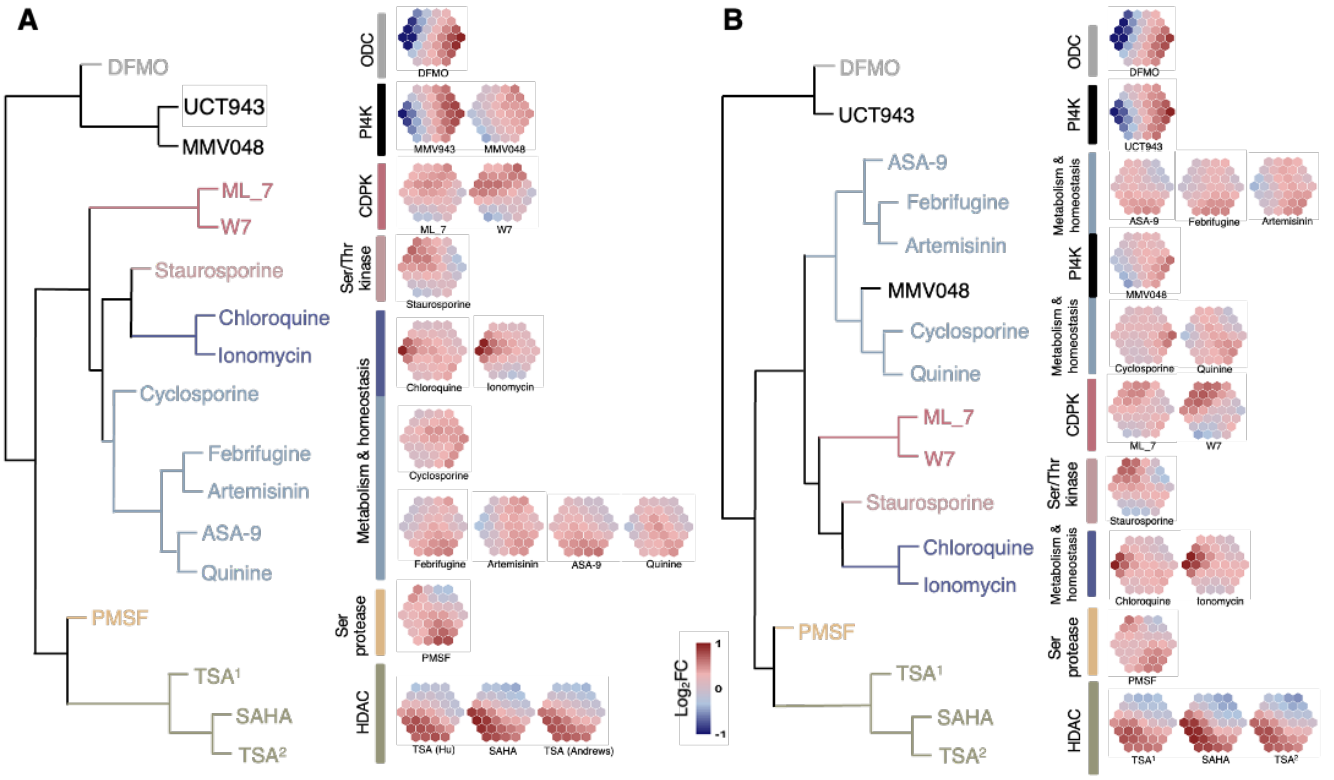


Figure 24. Evaluating the effect caused by the exclusion of biomarkers on the ability of chemo-transcriptomic fingerprints to cluster compound treatments based on their MoAs. (A) The original set of 50 predicted biomarkers (107) and **(B)** the reduced 43 predicted biomarkers. The chemo-transcriptomic fingerprints are created using self-organising maps (SOMs). The colour of the individual hexagons represents the $\log_2$ fold change in gene expression. Hierarchical clustering of the compound treatments was performed using Ward linkage on Euclidian distance of expression. DFMO = difluoromethylornithine, ODC = ornithine decarboxylase, PI4K = phosphatidylinositol 4-kinase, ASA-9 = 2-aminosuberic acid derivative, CDPK = calcium/calmodulin-dependent protein kinase, Ser/Thr kinase = serine/threonine kinase, ser protease = serine protease, PMSF = phenylmethylsulfonyl fluoride, TSA = trichostatin A, TSA¹ = TSA data from Hu *et al.* (2009) **(112)**, SAHA = suberoylanilide hydroxamic acid, TSA² = TSA data from Andrews *et al.* (2012) **(108)**, and HDAC = histone deacetylase.

However, it shows a change in where the MMV048 treatment clusters within the dendrogram. Biomarker 49, excluded during primer validation, is a biomarker identified from the MMV048 and UCT943 treatment datasets. It, along with three other predicted biomarkers (biomarkers 47, 48, and 50), are associated with PI4K inhibition as a MoA. Removing this biomarker may affect the resultant clustering of the chemo-transcriptomic fingerprints of parasites treated with compounds that target PI4K. UCT943 (that also targets PI4K) is more potent than MMV048 and the difference in potency between the compounds is reflected in the degree of increased or decreased transcript abundance, possibly also contributing to these two compound treatments' chemo-transcriptomic fingerprints clustering separately when using 43 predicted biomarkers. The remaining MoA clusters (grouped by colour in Figure 24) remained intact, but there was some change in where the individual clusters were located after removing the seven biomarkers. This was seen with the homeostasis and metabolism cluster (indicated in light blue in Figure 24) that switched with the branches containing the CDPK and serine/threonine kinase inhibitors clusters. This could be attributed to the specific biomarkers that were removed, for example, the removal of biomarker 46, which is associated specifically with treatment with

© University of Pretoria

ionomycin and its MoA, most likely contributed to the relocation of the associated cluster (indicated in dark blue in Figure 24). However, provided that the chemo-transcriptomic fingerprints resulting from treatment with compounds that share MoAs remain clustered together, it should be sufficient to predict new compounds' MoAs downstream.

3.8 Evaluation of the influence of the parasite's life cycle progression on the biomarker detection

The biomarkers' ability to distinguish between parasite populations treated with compounds with different MoAs and, therefore, their potential use as biomarkers was first evaluated by using compounds that were already included in the rational selection model developed by van Heerden *et al.* (2021). These compounds (artemisinin, chloroquine, DFMO, SAHA, and MMV048), therefore, have MoAs that are recognised and distinguished by the model. The transcription of *P. falciparum* parasites is tightly regulated, and genes are only expressed in the stage that they are needed by the parasite. Since compound treatment was performed using multiple incubation periods to ensure that a compound-induced transcriptional response could be observed, the stage-specific expression profile of transcripts had to be considered. We, therefore, evaluated the parasite populations' stage distribution and morphology after treatment with the control compounds to evaluate if stage distribution was severely affected (Figure 25).

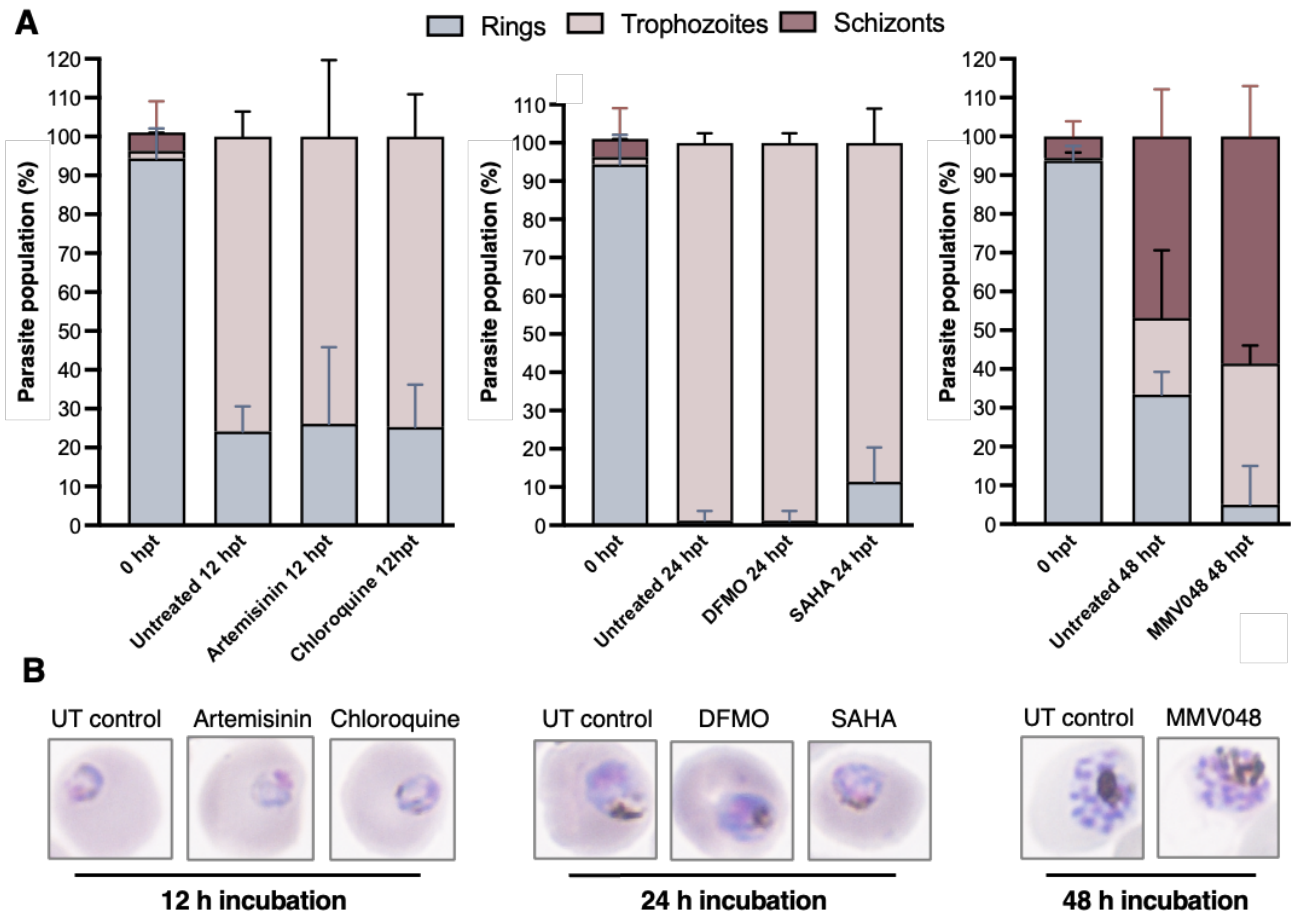


Figure 25. The parasite populations after treatment with the control compounds. (A) The stage distribution and **(B)** the morphology of the parasite populations after compound treatment. Compound-treated samples were

© University of Pretoria

treated at their respective IC_{90} values as determined by the SYBR Green I asexual proliferation assay. The Rapi-Diff-stained thin blood smears were viewed using a light microscope and the 100× oil immersion objective. The data represents the mean ± SD between microscope fields counted (a minimum of 60 parasites across three or more microscope fields of 200-300 erythrocytes(~1000 erythrocytes total)). DFMO = difluoromethylornithine, SAHA = suberoylanilide hydroxamic acid, UT = untreated, hpt = hours post-treatment, and IC_{90} = 90 % of the maximum inhibition concentration.

From the stage distribution of the parasite populations after the compound treatments, there were no statistically significant ($P > 0.05$, Student's t-test) deviations in the parasite populations between the compound-treated and untreated control samples of the 12 and 24 h incubation periods, respectively (Figure 25A and Supplementary Table 5a and b). However, for the MMV048-treated parasite population, there was a statistically significant difference ($P < 0.001$, Student's t-test) in the number of ring-stage parasites present compared to the untreated control, which is expected as the untreated parasites would have completed their life cycle and new ring-stages parasites in the next life cycle is evident, which is not present in the MM048-treated samples (Supplementary Table 5c). However, it is unlikely this would affect the biomarkers' expression as they were originally selected because their expression is not adversely affected across different treatment time points.

3.9 Evaluation of the concentration dependency of the chemo-transcriptomic fingerprints

MMV048 was used to evaluate the effect of the compound concentration on the transcriptional response of the predicted biomarker set. A 6-9 hpi ring-stage NF54 *P. falciparum* parasite culture was drug-treated with MMV048 at IC_{90} (22 nM) and ~6× IC_{90} (140 nM) and incubated for 48 h.

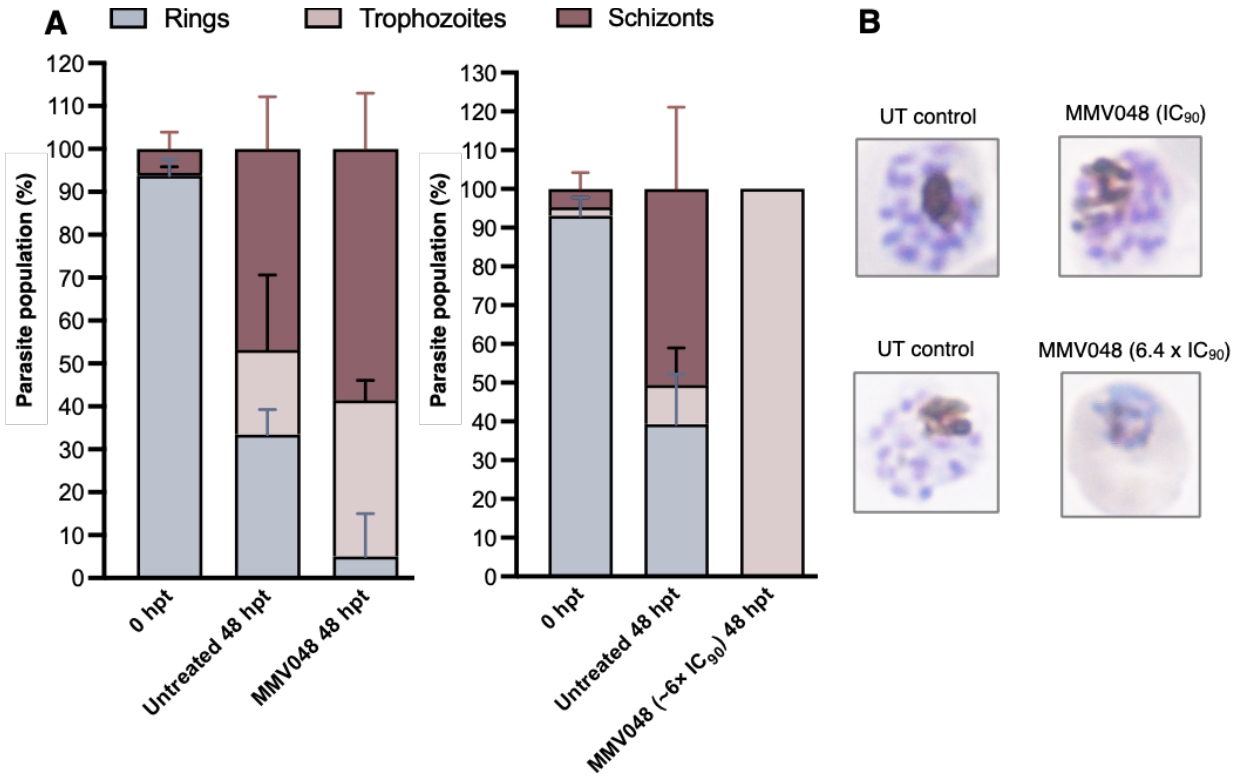


Figure 26. The concentration effect on the parasite population. (A) The stage distribution and **(B)** the morphology of the parasite population after the compound treatments. The stained thin blood smears were viewed using a light microscope and the 100× oil immersion objective. The data represents the mean ± SD between

© University of Pretoria

microscope fields counted (a minimum of 30 parasites across three or more microscope fields of 200-300 erythrocytes (~1000 erythrocytes total)). Where not visible, the error bars fall within the column.

The stage distribution and morphology of the parasites after compound treatment (Figure 26A and B) show that a higher compound concentration prevented the parasites from entering schizogony and affected the trophozoites, indicative of cell stress and pyknosis. This translated to a severe change in the chemo-transcriptomic fingerprint for MMV048-treated parasites to a decreased transcript abundance compared to the fingerprint at IC_{90} (Figure 27). This is indicative of a global cellular stress event associated with cell death. However, by contrast, parasites treated with MMV048 at IC_{90} can still progress to schizonts, and the evaluation of the biomarkers indicates a compound-induced change in the transcriptome, distinct from a global death response, as seen when using a higher concentration. This solidifies the use of IC_{90} in the platform, and all subsequent treatments were also at IC_{90}.

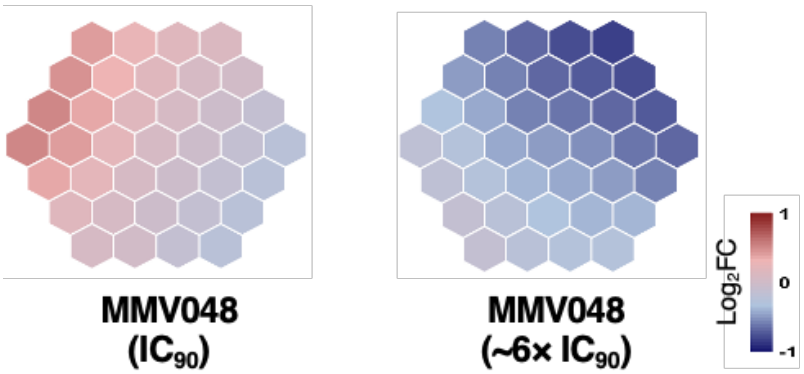


Figure 27. Concentration effect on the transcriptional responses of the biomarkers. The effect of different compound concentrations on the chemo-transcriptomic fingerprints. The colour of the individual hexagons represents the log_2 fold change in gene expression. Data are from one biological repeat (n = 1).

3.10 Evaluating the biomarkers' ability to distinguish the MoA of control compounds

The biomarkers were subsequently used to assess if they could distinguish the MoA of control compounds, each with known targets or biological profiles when they were used in qPCR analysis. *P. falciparum* parasites were treated with artemisinin, chloroquine, DMFO, SAHA, and MMV048 for a specific time period associated with their rate of action, RNA isolated, and the qPCR performed with the set of 43 biomarkers on each sample. The relative abundance of the biomarkers in each of these samples was thus determined. However, we evaluated two different relative quantification methods: 1) the well-established $2^{-\Delta\Delta Cq}$ method that assumes equal primer efficiencies and 2) the Pfaffl method that takes the primer efficiency of each biomarker into account. This allowed the assessment of the effect of primer efficiency differences on the platform's ability to distinguish between compounds with different MoAs (Figure 28).

The two methods of relative quantification show a positive correlation between the quantified values of the biomarkers for all compounds (Pearson r > 0.7), with most of the compound treatments having Pearson r values > 0.9 (Figure 28). Treatment with artemisinin (Pearson r = 0.78) and chloroquine (Pearson r = 0.91) had the lowest correlation between the quantified values of the biomarkers analysed

© University of Pretoria

using the two methods of relative quantification. However, both treatments still show a positive correlation (Pearson r > 0.7).

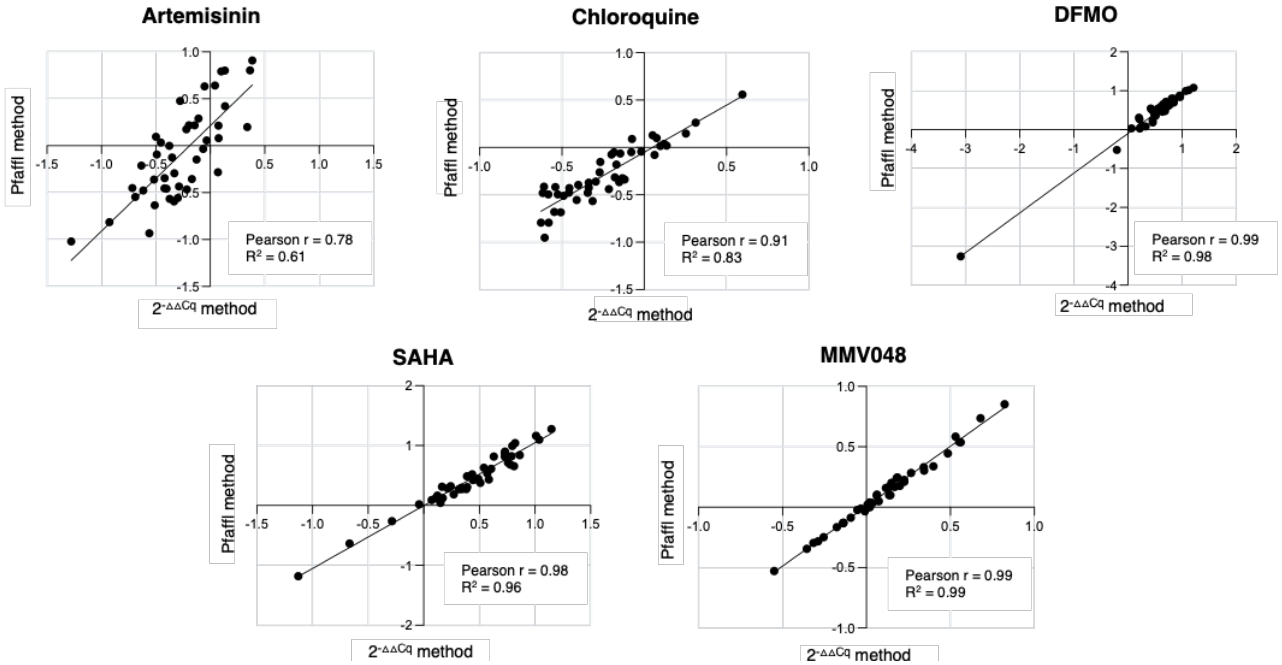


Figure 28. The correlation between the quantified values of the biomarkers analysed using different methods of relative quantification. The Pearson correlation between the quantified values of the biomarkers obtained using the $2^{-\Delta\Delta Cq}$ method of relative quantification and the Pfaffl method of relative quantification (95 % confidence interval), respectively. Data are from one biological repeat (n = 1).

The relative quantification levels obtained for each biomarker with the two methods above were subsequently used to generate chemo-transcriptomic fingerprints followed by hierarchical clustering to identify related MoAs as before (Figure 29). K-means clustering was also evaluated as a clustering method, however, when the silhouette analysis was performed to identify the optimum number of clusters, k = 2 was identified (Supplementary Figure 1). Grouping the compounds into two main clusters would not create the necessary separation to allow insight into their MoA and, therefore, hierarchical clustering would be more informative. Unique chemo-transcriptomic fingerprints were obtained after IDC *P. falciparum* parasites were treated with the control compounds, as is evident from specific fingerprints for each compound evaluated (Figure 29). Additionally, the biomarker analysis successfully separated these fingerprints into specific categories that correlate with their known MoA classifications.

© University of Pretoria

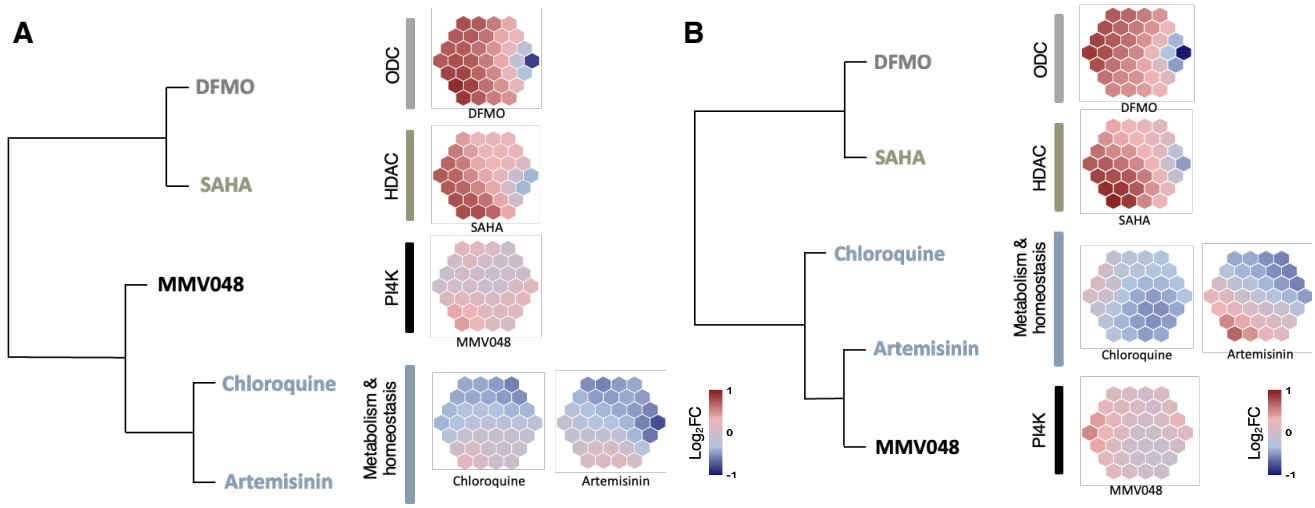


Figure 29. Chemo-transcriptomic fingerprints of the control compound-treated parasites when using different methods of relative quantification. (A) The $2^{-\Delta\Delta Cq}$ method of relative quantification that does not account for differences in primer efficiencies and **(B)** the Pfaffl method of relative quantification that considers the primer efficiency of each primer set. Hierarchical clustering of the compound treatments was performed using Ward linkage on Euclidian distance of expression. The chemo-transcriptomic fingerprints were created using self-organising maps (SOMs) that reduce the dimensionality of the transcriptional responses. The colour of the individual hexagons represents the $\log_2$ fold change in gene expression. Positive values indicate increased transcript abundance, while negative values indicate decreased transcript abundance (data was capped at -1 and 1 for the SOMs). Data are from one biological repeat (n = 1). DFMO = difluoromethylornithine, ODC = ornithine decarboxylase, SAHA = suberoylanilide hydroxamic acid, HDAC = histone deacetylase, and PI4K = phosphatidylinositol 4-kinase.

The biomarkers facilitated distinctive chemo-transcriptomic fingerprints for each of the compound treatments, with the target-specific chemo-transcriptomic fingerprints (DFMO, SAHA, and MMV048 treatments) separating from the non-specific target chemo-transcriptomic fingerprints (chloroquine and artemisinin treatments). The chemo-transcriptomic fingerprints resulting from the two different relative quantification methods were not substantially affected by deviations in primer efficiencies between biomarkers since the separation between the fingerprint classes was still evident. The chemo-transcriptomic fingerprints for the artemisinin- and chloroquine-treated parasites showed the greatest reorganisation. When using the $2^{-\Delta\Delta Cq}$ method, more logical clustering of the chemo-transcriptomic fingerprints was obtained for treatment with artemisinin and chloroquine, clustering the fingerprints together (both compounds affect metabolism and homeostasis), with the next closest cluster being that associated with MMV048 treatment, which has an MoA partly associated with affecting haemoglobin catabolism (86). This showed that the distinctive transcriptional responses observed when using the $2^{-\Delta\Delta Cq}$ method are compound-induced and not an artefact caused by differences in primer efficiencies and data analysis. The $2^{-\Delta\Delta Cq}$ method was preferable to the Pfaffl method, which had more complicated data analysis than the $2^{-\Delta\Delta Cq}$ method. The $2^{-\Delta\Delta Cq}$ method allows for a more scalable, high-throughput approach to data analysis while still showing distinctive chemo-transcriptomic fingerprints and was applied to the remainder of the data.

The data indicated that the 43 predictive biomarkers could be successfully evaluated with qPCR to reveal the MoA of a set of control compounds and showed desirable results, such as grouping the

© University of Pretoria

chemo-transcriptomic fingerprints of parasites treated with compounds that share an MoA together and separating the non-related MoAs.

3.11 Evaluating the biomarkers' ability to distinguish the MoA of clinical candidates

To investigate whether the qPCR-based evaluation of the abundance profiles of the 43 biomarkers can distinguish between compound treatments that were not included in the original rational selection transcript model from van Heerden *et al.* (2021), we evaluated the transcriptional responses of the biomarkers brought on by an additional set of compounds. This included known antimalarials or clinical candidates with known MoAs such as DSM265 (a DHODH inhibitor), pyrimethamine (a DHFR inhibitor), MMV084 (a PKG inhibitor), and M5717 (an EF2 inhibitor). The stage distribution and morphology of the parasite populations were again examined after parasites were treated with these compounds for specific time periods associated with their rate of action (Table 2 section 2.7) to ensure that life cycle stage differences would not complicate the interpretation of the data (Figure 30).

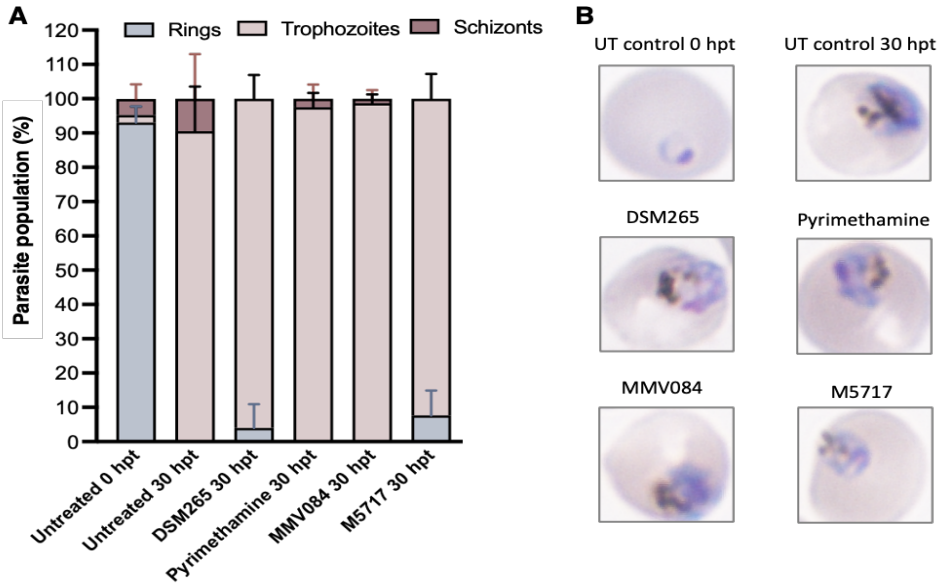


Figure 30. The parasite populations after treatment with the clinical candidates. (A) The stage distribution and **(B)** the morphology of the parasite populations after the compound treatments. Compound-treated samples were treated at their respective IC_{90} concentration as determined by the SYBR Green I asexual proliferation assay. The Rapi-Diff-stained thin blood smears were viewed using a light microscope and the 100× oil immersion objective. The data represents the mean ± SD between microscope fields counted (a minimum of 40 parasites across three or more microscope fields of 200-300 erythrocytes (~1000 erythrocytes total)). UT = untreated, hpt = hours post-treatment, and IC_{90} = 90 % of the maximum inhibition concentration.

The parasite populations treated with the clinical candidates showed no statistically significant difference ($P > 0.05$, unpaired Student's t-test) in the majority-stage parasites (> 90 % trophozoites) between all samples, including the untreated control parasite population (Figure 30A and Supplementary Table 6a). This indicated that the parasites were affected due to compound treatment but were still alive and able to progress to trophozoites. A statistically significant difference in the ring populations for the untreated control 30 hpt and M5717-treated parasites ($P < 0.05$, unpaired Student's t-test), as well as pyrimethamine- and M5717-treated parasites ($P < 0.05$, unpaired Student's t-test),

© University of Pretoria

were observed (Figure 30A and Supplementary Table 6a). However, the M5717-treated parasites had a ring-stage parasite population of 7.7 ± 7.2 %, while the untreated 30 hpt and the pyrimethamine-treated parasite populations did not have any ring-stage parasites. The deviation in stage distributions is, therefore, not biologically relevant as there was a small subset of trophozoites present in the starting population (untreated 0 hpt) at the time of compound treatment that would have developed into ring-stage parasites by the conclusion of the experiment. The parasites were affected by the compound treatments but could still progress to trophozoites, indicating that the fingerprints would not reflect a death phenotype (Figure 30B). The stage distribution of the parasites would, therefore, not influence the biomarkers' transcriptional response due to compound treatment in downstream expression analysis.

To investigate whether the qPCR-based analysis of the 43 biomarkers' expression maintained the ability to create the necessary separation between the chemo-transcriptomic fingerprints of parasites treated with compounds with different MoAs the data were hierarchically clustered (Figure 31).

The newly introduced treatments with clinical candidates with distinct MoAs show unique transcriptional responses due to the compound treatment and well-separated chemo-transcriptomic fingerprints are observed for each MoA (Figure 31). The clustering also shows a clear separation between the different compound treatments and their respective MoAs. This indicates that the 43 predictive biomarkers evaluated using qPCR could successfully differentiate between treatments with compounds that the rational selection model had no previous interaction with and distinguish between MoAs not previously integrated into the model. This solidified the use of the 43 predicted biomarkers to distinguish between MoAs even when the MoA is novel to the model.

© University of Pretoria

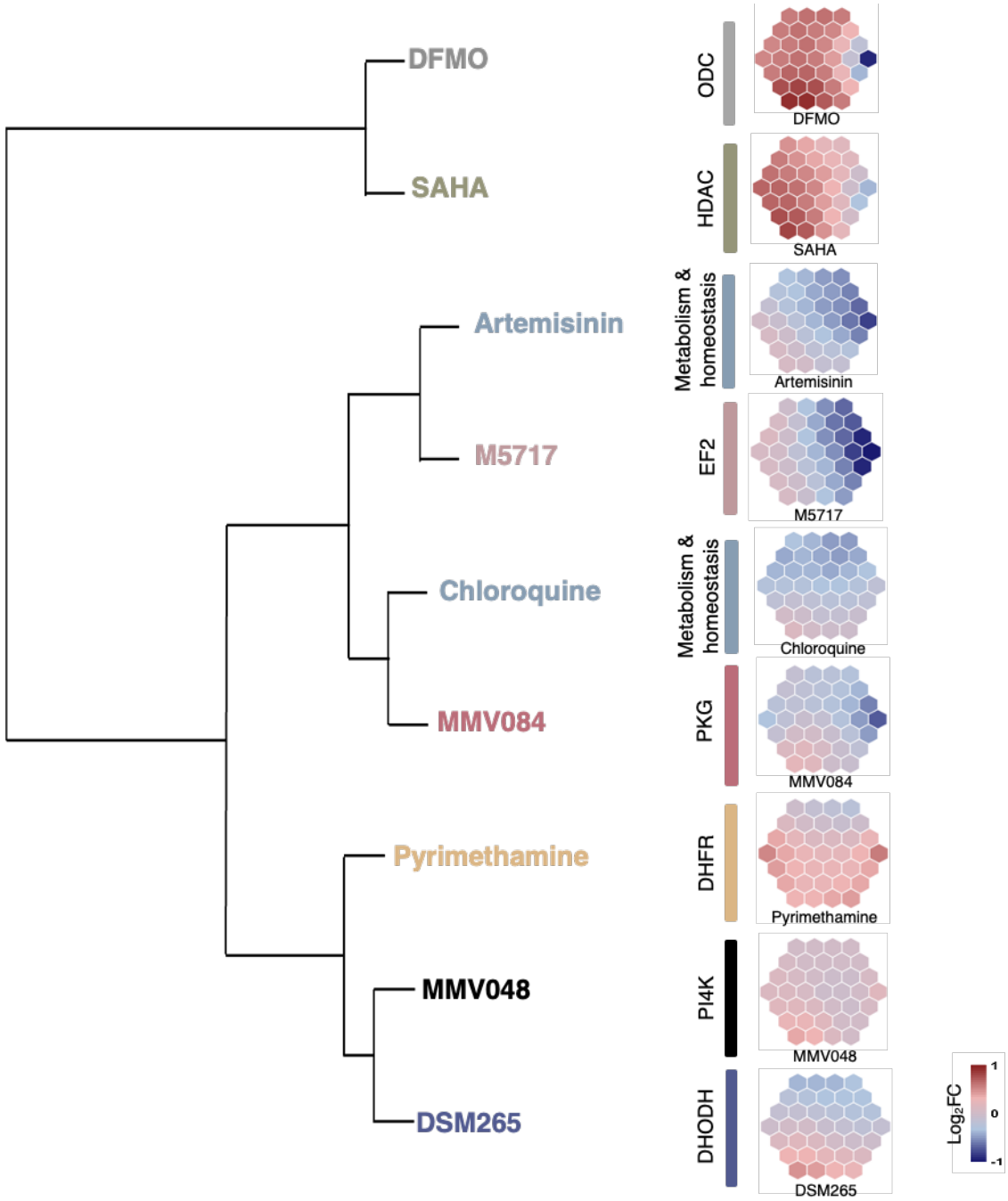


Figure 31. Evaluating the platform's ability to distinguish between MoAs upon adding clinical candidate treatments. The chemo-transcriptomic fingerprints are created using self-organising maps (SOMs) that reduce the dimensionality of the transcriptional responses. The colour of the individual hexagons represents the log2 fold change in gene expression. Hierarchical clustering of the compound treatments was performed using Ward linkage on Euclidian distance of expression. Data are from one biological repeat (n = 1). DFMO = difluoromethylornithine, ODC = ornithine decarboxylase, SAHA = suberoylanilide hydroxamic acid, HDAC = histone deacetylase, PI4K = phosphatidylinositol 4-kinase, DHODH = dihydroorotate dehydrogenase, DHFR = dihydrofolate reductase, PKG = cGMP-dependent protein kinase, and EF2 = translation elongation factor 2.

3.12 Evaluating the biomarkers' ability to show similarities between compounds with the same MoA

Treatment with the clinical candidate UCT943 was additionally introduced to assess how the qPCR-based biomarker transcription analysis platform deals with compounds with the same MoA. UCT943 has the same target as MMV048, both inhibit PI4K. Both compounds underwent the same treatment

© University of Pretoria

protocol as they inherently act against the same parasite development stage. The stage distribution and morphology of the parasite populations were evaluated after 48 h of incubation with the compounds (Figure 32).

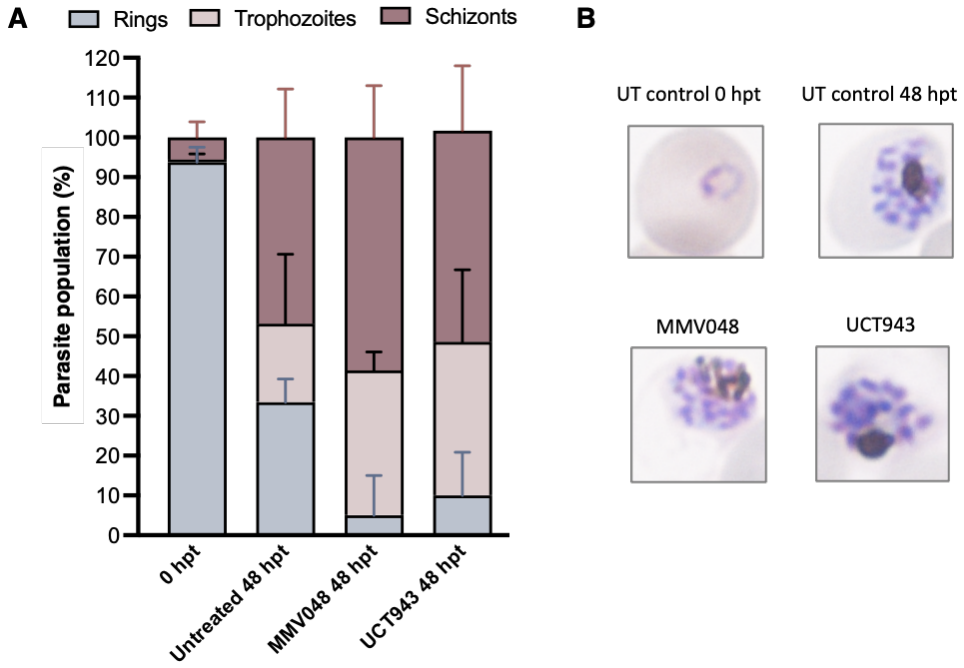


Figure 32. Evaluation of parasite populations treated with compounds with a shared MoA. (A) The stage distribution and **(B)** the morphology of the parasite populations after the compound treatments. Cultures were treated at IC_{90}. The Rapi-Diff-stained thin blood smears were viewed using a light microscope and the 100× oil immersion objective. The data represents the mean ± SD between microscope fields counted (a minimum of 50 parasites across three or more microscope fields of 200-300 erythrocytes (~1000 erythrocytes total)). UT = untreated, hpt = hours post-treatment, and IC_{90} = 90 % of the maximum inhibition concentration.

There was no statistically significant difference ($P > 0.05$, unpaired Student's t-test) in the stage distributions of the parasite populations between the MMV048 and UCT943 treatments (Figure 32A and Supplementary Table 6b). A statistically significant difference in ring-stage parasites in the populations of the untreated 48 hpt and the MMV048 treatment ($P < 0.001$, unpaired Student's t-test) and the untreated 48 hpt population and the UCT943 treatment ($P < 0.01$, unpaired Student's t-test) were observed. This is due to the compounds' greatest activity on schizont-stage parasites, leading to fewer rings in the next IDC, as previously established for MMV048 (78). Both compound-treated parasite populations progressed to schizonts, indicating that the parasites were affected but not dead (Figure 32B). Stage-specific transcription would not lead to differences in the chemo-transcriptomic fingerprints resulting from compound-treated parasites in downstream qPCR analysis.

The metabolism and homeostasis classification defined by van Heerden *et al.* (2021) typically contains chemo-transcriptomic fingerprints of parasites treated with compounds that influence multiple biological processes or have less clearly defined targets. This is observed for the chemo-transcriptomic fingerprints of artemisinin- and chloroquine-treated parasites, which do not share a target but have a broader compound-induced effect (Figure 33). The clustering of these two compound treatments mimics what was shown by van Heerden *et al.* (2021), who saw a similar clustering for artemisinin and

© University of Pretoria

chloroquine treatments when the full set of 50 predicted biomarkers was used (Figure 9 from section 1.7). The chemo-transcriptomic fingerprint that represents treatment with M5717 clusters separately from the other compound treatments, reflecting what was seen in metabolic fingerprinting, where M5717 treatment was listed as unclassified as it affects translation by inhibiting EF2, an MoA that does not overlap directly with the other compounds in this dataset (108).

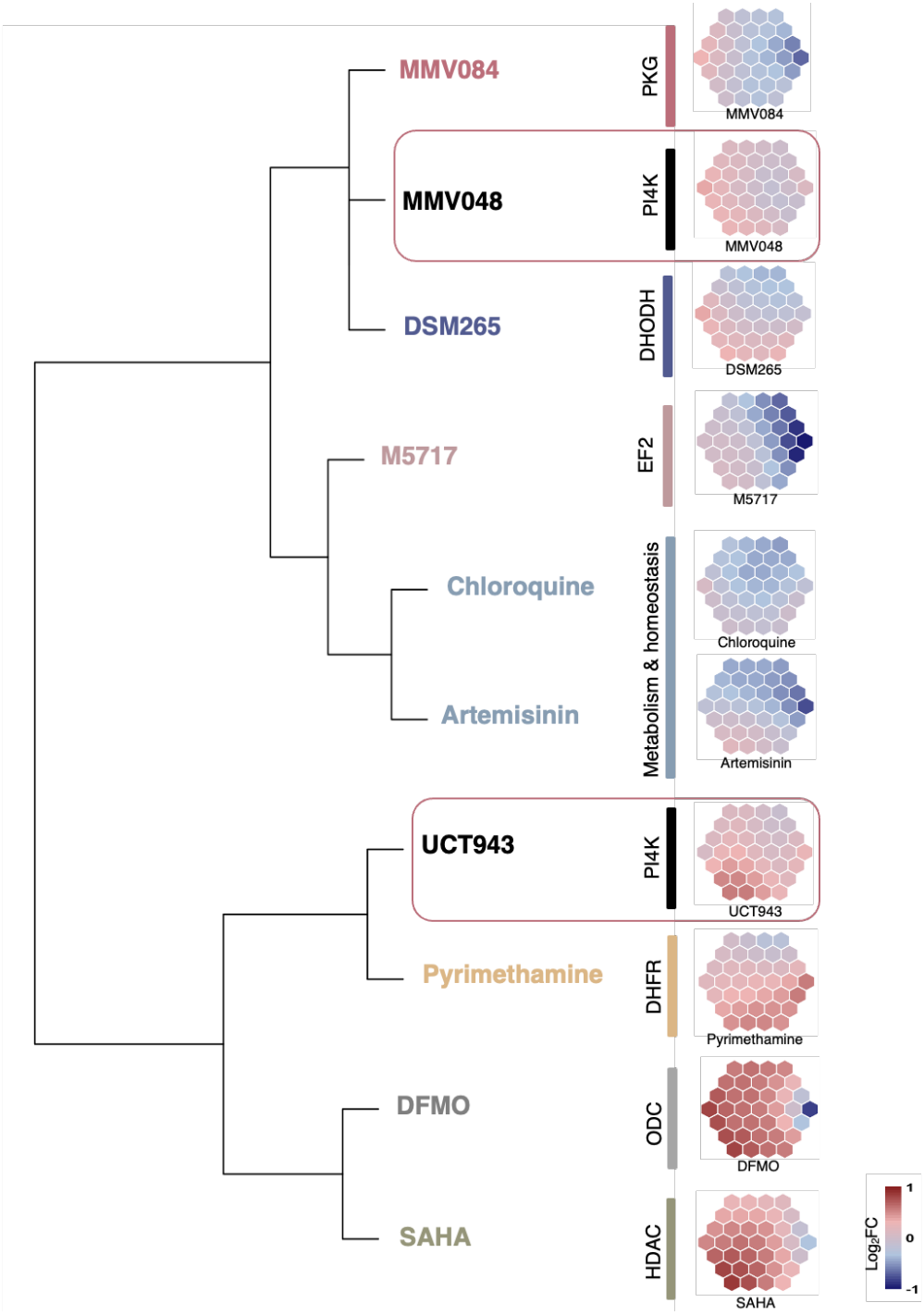


Figure 33. The predicted biomarker set shows distinctive chemo-transcriptomic fingerprints and early signs of clustering by MoA. Red blocks indicate the two compounds, MMV048 and UCT943, that share an MoA. The chemo-transcriptomic fingerprints are created using self-organising maps (SOMs) that reduce the dimensionality of the transcriptional responses. The colour of the individual hexagons represents the $\log_2$ fold change in gene expression. Hierarchical clustering of the compound treatments was performed using Ward linkage on Euclidian distance of expression. Data are from one biological repeat (n = 1). PKG = cGMP-dependent protein kinase, PI4K = phosphatidylinositol 4-kinase, DHODH = dihydroorotate dehydrogenase, EF2 = translation elongation factor 2, DHFR = dihydrofolate reductase, DFMO = difluoromethylornithine, ODC = ornithine decarboxylase, SAHA = suberoylanilide hydroxamic acid, and HDAC = histone deacetylase.

© University of Pretoria

The chemo-transcriptomic fingerprints resulting from treatment with MMV048 and UCT943 cluster apart. The limited size of this compound data set could exacerbate this effect, along with the removal of a biomarker specific to the PI4K inhibition MoA. This did, however, raise the question of whether the transcriptional responses due to the two compound treatments show similarities. Treatment with compounds that share a target or MoA should ideally produce chemo-transcriptomic fingerprints that bear similarities in transcriptional responses. Therefore, the chemo-transcriptomic fingerprints resulting from treatment with the two compounds were investigated (Figure 34).

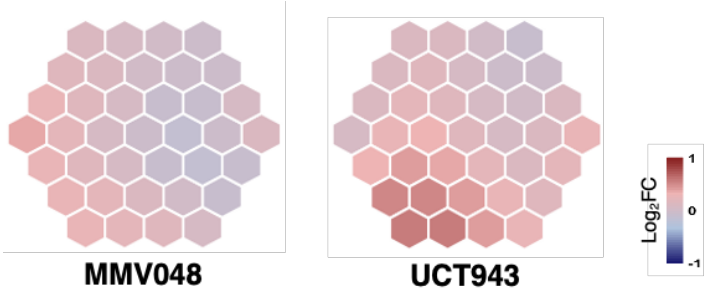


Figure 34. Evaluation of chemo-transcriptomic fingerprints due to treatment with compounds with a shared MoA. The chemo-transcriptomic fingerprints of the MMV048 and UCT943 treatments, both targeting phosphatidylinositol 4-kinase. Samples were treated at their respective IC_{90} concentration as determined by the SYBR Green I asexual proliferation assay. The colour of the individual hexagons represents the log_2 fold change in gene expression. Positive values indicate increased transcript abundance, while negative values indicate decreased transcript abundance. Data are from one biological repeat (n = 1). IC_{90} = 90 % of the maximum inhibition concentration.

The chemo-transcriptomic fingerprints resulting from treatment with MMV048 and UCT943 were similar, with both showing a majority increased transcript abundance (pink) and a region of decreased transcript abundance (light blue) on the right (Figure 34). Treatment with UCT943 results in a slightly more prominent transcriptional response, possibly due to its increased potency compared to treatment with MMV048. This solidifies that treatment with compounds with similar MoAs shows similar chemo-transcriptomic fingerprints when using 43 biomarkers for MoA classification.

© University of Pretoria

Chapter 4: Discussion

Transcriptomics provides a dynamic insight into cellular environmental changes and the evaluation of transcriptional responses due to drug pressure offers new insight into a compound's MoA. In this study, a biomarker set from van Heerden *et al.* (2021) was investigated for its biological relevance as a tool to predict compound MoA. This was accomplished by building a qPCR-based biomarker transcription analysis platform that provided a medium-throughput approach to MoA deconvolution. In this platform, the parasite's transcriptional responses to specific compound treatments were monitored relative to an untreated control population using qPCR and the established $2^{-\Delta\Delta Cq}$ relative quantification method.

Several studies have indicated that even though *P. falciparum* parasites possess tightly regulated transcription, it is possible to elicit a measurable transcriptional response due to compound treatment (104, 112). Hu *et al.* (2010) have shown that *P. falciparum* parasites' transcriptome is sensitive to chemical probing and that genes related by function show similar transcriptional responses due to chemical stimulus (112). This led them to conclude that functionally related genes have the same regulatory mechanisms when exposed to external stimuli, allowing insight into *P. falciparum* parasite's functional genomics. Additionally, Siwo *et al.* (2015) showed that both similarities in chemical structures and MoAs contributed to similarities in transcriptional responses due to compound exposure (104).

Transcriptional profiling is an established and invaluable resource for functional studies and drug discovery in the malaria field, however, applying this whole-genome approach is not conducive to a high-throughput environment. The use of a limited set of transcripts to circumvent the challenges presented regarding throughput has been established in other fields. This especially regards tuberculosis, where Murima *et al.* (2013) identified a minimum set of 90 transcripts that proved sufficient to elucidate the MoA of compounds identified using phenotypic screens (95). This provided the groundwork to establish a similar platform to be used in antimalarial drug discovery.

van Heerden *et al.* (2021) used machine learning and already available compound-induced transcriptional response data to identify biomarkers for MoA classification. The predicted biomarkers additionally ensured compound-specific responses and that the chemo-transcriptomic fingerprints do not merely reflect general drug stress that would be common to most compounds. The rational selection approach of the model ensured that the top 50 biomarkers did not contain transcripts shared between multiple MoAs (107). The high level of co-regulation among functionally related genes that were established by Hu *et al.* (2010) may suggest that the 50 transcripts could additionally serve as a representative set of all the genes that show functional relations to the biomarkers. This could also give an indication of the responsiveness of all the co-regulated genes. The connectedness of the biomarkers to other functionally related genes would have to be investigated to ascertain the full implication of this.

© University of Pretoria

In the pursuit of building a qPCR-based biomarker transcription analysis platform that functions under uniform conditions and allows for swifter and scalable MoA classification, there was the need to remove some of the predicted biomarkers from the set. These biomarkers performed unfavourably in this platform and often showed limited availability of primers that fit the parameters. It was, however, established by van Heerden *et al*. (2021) that 50 biomarkers were the minimum number of biomarkers needed to allow for robust prediction of a compound's MoA. The removal of seven biomarkers (3. PF3D7_0206700, 28. PF3D7_1425800, 31. PF3D7_1013500, 32. PF3D71127900, 45. PF3D7_1458900, 46. PF3D7_1038400, and 49. PF3D7_1340900) throughout the process of primer validation, made it essential to verify that the SOMs used to construct the chemo-transcriptomic fingerprints could still display clear separation between compounds using the original microarray data. Herewith it proved that even 43 biomarkers were sufficient to separate between compounds and their MoAs using chemo-transcriptomic fingerprints, displaying a minimal shift in the appearance of the chemo-transcriptomic fingerprints after removing the seven biomarkers.

A set of five control compounds (artemisinin, chloroquine, DFMO, SAHA, and MMV048) were introduced to the qPCR-based biomarker transcription analysis platform. These compounds are well characterised and familiar to the rational selection 50 transcript model with specific biomarkers pertaining to their MoAs. This showed that the predicted biomarkers maintained their ability to distinguish between treatments with compounds that have dissimilar MoAs *in vitro.* Additionally, this helped clarify that the biomarkers stand when translated to qPCR and that the between-platform variance created when microarray data was replaced by qPCR was not concerning. As expected, the predicted biomarkers could show compound-specific responses not merely due to deviations in life cycle stages, proving that they are robust at any stage in the IDC. This was ensured by van Heerden *et al*. (2021) as the transcripts selected were differentially expressed throughout the treatment period of individual compounds. This served to prove that transcriptional fingerprinting using a limited set of biomarkers can be applied to *P. falciparum* parasites, therefore validating the machine-learning model developed by van Heerden *et al*. (2021).

The introduction of clinical candidates that were not previously exposed to the rational selection 50 transcript model verified the biomarkers' ability to distinguish between MoAs as the newly introduced clinical candidate treatments showed separation from the control compound treatments. The clinical candidates exhibit MoAs distinct from any of the control compounds. This meant that the chemo-transcriptomic fingerprints could discern compounds MoAs not showing similarities to MoAs already established within the model, allowing the identification of novel MoAs. This is impressive as the original model was built from biomarkers that are associated with specific MoAs, further solidifying that this limited set of biomarkers provides robust chemo-transcriptomic fingerprints pertaining to a compound-specific response.

© University of Pretoria

The compound UCT943 is a derivative of MMV048 with improved solubility and anti-plasmodial activity, and the compounds share chemical characteristics (137). The chemo-transcriptomic fingerprints of the two compound treatments show shared trends in transcriptional response. Treatment with UCT943 does induce a marginally stronger transcriptional response in the parasites compared to treatment with MMV048, likely due to UCT943's increased potency and pharmacodynamic properties (137). This same phenomenon has been observed by Siwo *et al*. (2015), where small molecule inhibitors that share chemical substructures and MoAs induce similar transcriptional responses, further validating this approach (104). This could be a beneficial characteristic as it can prove useful when compounds are modified for better potency, selectivity, and pharmacodynamic properties during the drug discovery process to ensure that there is no deviation from the compound's original MoA or the development of off-target effects.

It is important to note that while this platform provides a promising medium-throughput approach, there are limitations in its current state. Removing biomarkers from the limited set can compromise the platform's separation ability for specific MoAs including inhibition of CDPK, ODC, merozoite invasion, serine protease, calcium-binding ionophore, and PI4K. However, there are multiple biomarkers associated with each MoA within the rational selection 50 transcript model (except for inhibition of calcium-binding ionophores) that should compensate for the removed biomarkers. Extending the platform to the set of 100 potential biomarkers identified by van Heerden *et al*. (2021) could contribute to much greater separation between treatments with different compounds and the compounds' respective MoAs. Additionally, the limited set of compound treatments currently introduced to the platform also means that hierarchical clustering could infer a relationship between compound treatments where there is none. This, however, should start to resolve as a larger compound treatment pool is introduced and the number of MoAs is expanded. Machine learning models have the added benefit of being dynamic and with the generation of new data, such as introducing new MoAs and chemical classes, they can be continuously updated, improving the model and its predictive capabilities.

The ability to condense the compound effects into chemo-transcriptomic fingerprints with the combination of machine learning and transcriptional profiling can substantially accelerate antimalarial drug discovery by providing a greater throughput approach than the genome-wide approach often employed by microarrays. Additionally, a multiplexing qPCR-based approach can be considered to further improve the platform's throughput. This streamlined approach to MoA deconvolution allows for enhanced predictive capabilities in the early stages of drug development that would ensure the prioritisation of compounds that meet the requirements for clinical implementation. A severe impediment to the drug discovery process is the challenges inherent to the deconvolution of a compound's MoA. Knowing the MoA is, however, extremely beneficial to ensure that a clinical candidate does not progress throughout most of the pipeline only to be excluded due to an unfavourable MoA or off-target effects. This minimises the waste of time, resources, and often substantial funds linked to developing new

© University of Pretoria

antimalarials. Therefore, the collaboration of computational approaches and experimental validation ensures that only the most viable candidates progress through the drug discovery pipeline, ultimately reducing late-stage failure.

Herewith we contribute a medium-throughput, scalable approach to MoA stratification using a limited set of biomarkers and eliminate the uncertainties and ambiguities that often accompany MoA and target identification pursuits. This platform would, in turn, also assist in streamlining antimalarial drug development and provide a more efficient and economical approach to the identification of a compound's MoA.

© University of Pretoria

Chapter 5: Conclusion

With the persistent development of drug resistance to current antimalarials, there is a constant need to investigate new compounds with novel MoAs. During the drug discovery process, compounds progress through distinct stages before they can be investigated by clinical trials. The hit-to-lead and lead optimisation stages tend to be more successful if the compound's target or MoA is known. Target and MoA deconvolution are, however, challenging processes and have low success rates. To address this, van Heerden *et al.* (2021) developed a rational selection 50 transcript model that can stratify compounds based on their MoA with an accuracy of 77 %.

Using the treatment parameters established above, we demonstrated that the identified biomarkers function as indicators of compound MoA, creating a clear distinction between different MoAs and showing similar chemo-transcriptomic fingerprints for compounds that share an MoA. This proved that the predicted biomarkers identified using machine learning are biologically relevant for MoA stratification. This study offers a medium throughput approach to antimalarial drug MoA deconvolution using a limited set of predictive biomarkers. It can also guide future, more in-depth studies into a compound's target and MoA, possibly reducing ambiguity and providing a starting point for MoA and target deconvolution studies.

© University of Pretoria

References

1. WHO. World Malaria Report 2023. Geneva: World Health Organisation; 2023.
2. Wells TN, Alonso PL, Gutteridge WE. New medicines to improve control and contribute to the eradication of malaria. Nat Rev Drug Discov. 2009;8(11):879-91.
3. Phillips MA, Burrows JN, Manyando C, van Huijsduijnen RH, Van Voorhis WC, Wells TNC. Malaria. Nat Rev Dis Primers. 2017;3:17050.
4. Acharya P, Garg M, Kumar P, Munjal A, Raja KD. Host-Parasite Interactions in Human Malaria: Clinical Implications of Basic Research. Front Microbiol. 2017;8:889.
5. Ngotho P, Soares AB, Hentzschel F, Achcar F, Bertuccini L, Marti M. Revisiting gametocyte biology in malaria parasites. FEMS Microbiol Rev. 2019;43(4):401-14.
6. Josling GA, Llinás M. Sexual development in *Plasmodium* parasites: knowing when it's time to commit. Nat Rev Microbiol. 2015;13(9):573-87.
7. Lampe L, Levashina EA. MicroRNA Tissue Atlas of the Malaria Mosquito *Anopheles gambiae*. G3 (Bethesda). 2018;8(1):185-93.
8. Venugopal K, Hentzschel F, Valkiūnas G, Marti M. *Plasmodium* asexual growth and sexual development in the haematopoietic niche of the host. Nat Rev Microbiol. 2020;18(3):177-89.
9. Meibalan E, Marti M. Biology of Malaria Transmission. Cold Spring Harb Perspect Med. 2017;7(3).
10. Connacher J, von Grüning H, Birkholtz L. Histone Modification Landscapes as a Roadmap for Malaria Parasite Development. Front Cell Dev Biol. 2022;10:848797.
11. Cowman AF, Healer J, Marapana D, Marsh K. Malaria: Biology and Disease. Cell. 2016;167(3):610-24.
12. Maier AG, Matuschewski K, Zhang M, Rug M. *Plasmodium falciparum*. Trends Parasitol. 2019;35(6):481-2.
13. Beck JR, Ho CM. Transport mechanisms at the malaria parasite-host cell interface. PLOS Pathog. 2021;17(4):e1009394.
14. Lew VL, Tiffert T, Ginsburg H. Excess hemoglobin digestion and the osmotic stability of *Plasmodium falciparum*-infected red blood cells. Blood. 2003;101(10):4189-94.
15. Introini V, Govendir MA, Rayner JC, Cicuta P, Bernabeu M. Biophysical Tools and Concepts Enable Understanding of Asexual Blood Stage Malaria. Front Cell Infect Microbiol. 2022;12:908241.
16. Coronado LM, Nadovich CT, Spadafora C. Malarial hemozoin: from target to tool. Biochim Biophys Acta. 2014;1840(6):2032-41.
17. Ganter M, Goldberg JM, Dvorin JD, Paulo JA, King JG, Tripathi AK, *et al*. *Plasmodium falciparum* CRK4 directs continuous rounds of DNA replication during schizogony. Nat Microbiol. 2017;2:17017.
18. Voß Y, Klaus S, Guizetti J, Ganter M. *Plasmodium* schizogony, a chronology of the parasite's cell cycle in the blood stage. PLOS Pathog. 2023;19(3):e1011157.
19. Dixon MWA, Tilley L. *Plasmodium falciparum* goes bananas for sex. Mol Biochem Parasitol. 2021;244:111385.
20. Dash M, Sachdeva S, Bansal A, Sinha A. Gametogenesis in *Plasmodium*: Delving Deeper to Connect the Dots. Front Cell Infect Microbiol. 2022;12:877907.
21. Talman AM, Domarle O, McKenzie FE, Ariey F, Robert V. Gametocytogenesis : the puberty of *Plasmodium falciparum*. Malar J. 2004;3(1):24.
22. Volohonsky G, Paul-Gilloteaux P, Štáfková J, Soichot J, Salamero J, Levashina EA. Kinetics of *Plasmodium* midgut invasion in *Anopheles* mosquitoes. PLOS Pathog. 2020;16(9):e1008739.
23. Usui M, Williamson KC. Stressed Out About *Plasmodium falciparum* Gametocytogenesis. Front Cell Infect Microbiol. 2021;11:790067.
24. de Jong RM, Tebeje SK, Meerstein-Kessel L, Tadesse FG, Jore MM, Stone W, *et al.* Immunity against sexual stage *Plasmodium falciparum* and *Plasmodium vivax* parasites. Immunol Rev. 2020;293(1):190-215.
25. White NJ, Pukrittayakamee S, Hien TT, Faiz MA, Mokuolu OA, Dondorp AM. Malaria. Lancet. 2014;383(9918):723-35.
26. Savi MK. An Overview of Malaria Transmission Mechanisms, Control, and Modeling. Med Sci (Basel). 2022;11(1).
27. Wangdi K, Furuya-Kanamori L, Clark J, Barendregt JJ, Gatton ML, Banwell C, *et al.* Comparative effectiveness of malaria prevention measures: a systematic review and network meta-analysis. Parasit Vectors. 2018;11(1):210.

© University of Pretoria

28. Birkholtz L-M, Bornman R, Focke W, Mutero C, de Jager C. Sustainable malaria control: transdisciplinary approaches for translational applications. Malar J. 2012;11(1):431.
29. Kudom AA, Anane LN, Afoakwah R, Adokoh CK. Relating High Insecticide Residues in Larval Breeding Habitats in Urban Residential Areas to the Selection of Pyrethroid Resistance in *Anopheles gambiae* s.l. (Diptera: Culicidae) in Akim Oda, Ghana. J Med Entomol. 2018;55(2):490-5.
30. Moshi IR, Ngowo H, Dillip A, Msellemu D, Madumla EP, Okumu FO, *et al*. Community perceptions on outdoor malaria transmission in Kilombero Valley, Southern Tanzania. Malar J. 2017;16(1):274.
31. Geerligs PD, Brabin BJ, Eggelte TA. Analysis of the effects of malaria chemoprophylaxis in children on haematological responses, morbidity and mortality. Bull World Health Organ. 2003;81(3):205-16.
32. Burrows JN, Duparc S, Gutteridge WE, Hooft van Huijsduijnen R, Kaszubska W, Macintyre F, *et al.* New developments in anti-malarial target candidate and product profiles. Malar J. 2017;16(1):26.
33. Greenwood B. Anti-malarial drugs and the prevention of malaria in the population of malaria endemic areas. Malar J. 2010;9 Suppl 3(Suppl 3):S2.
34. Ashley EA, Pyae Phyo A, Woodrow CJ. Malaria. Lancet. 2018;391(10130):1608-21.
35. Fernando SD, Rodrigo C, Rajapakse S. Chemoprophylaxis in malaria: drugs, evidence of efficacy and costs. Asian Pac J Trop Med. 2011;4(4):330-6.
36. Littmann J, Achu D, Laufer MK, Karema C, Schellenberg D. Making the most of malaria chemoprevention. Malar J. 2024;23(1):51.
37. Datoo MS, Natama MH, Somé A, Traoré O, Rouamba T, Bellamy D, *et al.* Efficacy of a low-dose candidate malaria vaccine, R21 in adjuvant Matrix-M, with seasonal administration to children in Burkina Faso: a randomised controlled trial. Lancet. 2021;397(10287):1809-18.
38. Ngou O, Nzoumbou-Boko R, Bakamba P, Boum Y, 2nd. R21/Matrix-M malaria vaccine: a vital tool in the arsenal against malaria, not a silver bullet. Lancet Infect Dis. 2024;24(5):438-9.
39. Hanboonkunupakarn B, White NJ. Advances and roadblocks in the treatment of malaria. Br J Clin Pharmacol. 2022;88(2):374-82.
40. Dondorp A, Nosten F, Stepniewska K, Day N, White N. Artesunate versus quinine for treatment of severe *falciparum* malaria: a randomised trial. Lancet. 2005;366(9487):717-25.
41. Davis TM, Phuong HL, Ilett KF, Hung NC, Batty KT, Phuong VD, *et al.* Pharmacokinetics and pharmacodynamics of intravenous artesunate in severe *falciparum* malaria. Antimicrob Agents Chemother. 2001;45(1):181-6.
42. Poespoprodjo JR, Douglas NM, Ansong D, Kho S, Anstey NM. Malaria. Lancet. 2023;402(10419):2328-45.
43. Muangphrom P, Seki H, Fukushima EO, Muranaka T. Artemisinin-based antimalarial research: application of biotechnology to the production of artemisinin, its mode of action, and the mechanism of resistance of *Plasmodium* parasites. J Nat Med. 2016;70(3):318-34.
44. Maafoh C, Onyedibe K. Alternative first-line malaria treatment. Ann Afr Med. 2024;23(1):5-12.
45. Umumararungu T, Nkuranga JB, Habarurema G, Nyandwi JB, Mukazayire MJ, Mukiza J, *et al.* Recent developments in antimalarial drug discovery. Bioorg Med Chem. 2023;88-89:117339.
46. Conrad MD, Rosenthal PJ. Antimalarial drug resistance in Africa: the calm before the storm? Lancet Infect Dis. 2019;19(10):e338-e51.
47. Mbugi EV, Mutayoba BM, Malisa AL, Balthazary ST, Nyambo TB, Mshinda H. Drug resistance to sulphadoxine-pyrimethamine in *Plasmodium falciparum* malaria in Mlimba, Tanzania. Malar J. 2006;5(1):94.
48. Bruxvoort K, Goodman C, Kachur SP, Schellenberg D. How patients take malaria treatment: a systematic review of the literature on adherence to antimalarial drugs. PLOS ONE. 2014;9(1):e84555.
49. Ariey F, Witkowski B, Amaratunga C, Beghain J, Langlois AC, Khim N, *et al.* A molecular marker of artemisinin-resistant *Plasmodium falciparum* malaria. Nature. 2014;505(7481):50-5.
50. Ashley EA, Dhorda M, Fairhurst RM, Amaratunga C, Lim P, Suon S, *et al*. Spread of artemisinin resistance in *Plasmodium falciparum* malaria. N Engl J Med. 2014;371(5):411-23.
51. Siqueira-Neto JL, Wicht KJ, Chibale K, Burrows JN, Fidock DA, Winzeler EA. Antimalarial drug discovery: progress and approaches. Nat Rev Drug Discov. 2023;22(10):807-26.
52. Burrows JN, van Huijsduijnen RH, Möhrle JJ, Oeuvray C, Wells TN. Designing the next generation of medicines for malaria control and eradication. Malar J. 2013;12:187.
53. Yahiya S, Rueda-Zubiaurre A, Delves MJ, Fuchter MJ, Baum J. The antimalarial screening landscape-looking beyond the asexual blood stage. Curr Opin Chem Biol. 2019;50:1-9.

© University of Pretoria

54. Madhav H, Hoda N. An insight into the recent development of the clinical candidates for the treatment of malaria and their target proteins. Eur J Med Chem. 2021;210:112955.
55. McCarthy JS, Donini C, Chalon S, Woodford J, Marquart L, Collins KA, *et al*. A Phase 1, Placebo-controlled, Randomized, Single Ascending Dose Study and a Volunteer Infection Study to Characterize the Safety, Pharmacokinetics, and Antimalarial Activity of the *Plasmodium* Phosphatidylinositol 4-Kinase Inhibitor MMV390048. Clin Infect Dis. 2020;71(10):e657-e64.
56. Demarta-Gatsi C, Donini C, Duffy J, Sadler C, Stewart J, Barber JA, *et al.* Malarial PI4K inhibitor induced diaphragmatic hernias in rat: Potential link with mammalian kinase inhibition. Birth Defects Res. 2022;114(10):487-98.
57. Jørgensen R, Merrill AR, Andersen GR. The life and death of translation elongation factor 2. Biochem Soc Trans. 2006;34(Pt 1):1-6.
58. McCarthy JS, Lotharius J, Rückle T, Chalon S, Phillips MA, Elliott S, *et al.* Safety, tolerability, pharmacokinetics, and activity of the novel long-acting antimalarial DSM265: a two-part first-in-human phase 1a/1b randomised study. Lancet Infect Dis. 2017;17(6):626-35.
59. Duffey M, Blasco B, Burrows JN, Wells TNC, Fidock DA, Leroy D. Assessing risks of *Plasmodium falciparum* resistance to select next-generation antimalarials. Trends Parasitol. 2021;37(8):709-21.
60. White J, Dhingra SK, Deng X, El Mazouni F, Lee MCS, Afanador GA, *et al.* Identification and Mechanistic Understanding of Dihydroorotate Dehydrogenase Point Mutations in *Plasmodium falciparum* that Confer *in Vitro* Resistance to the Clinical Candidate DSM265. ACS Infect Dis. 2019;5(1):90-101.
61. Baker DA, Matralis AN, Osborne SA, Large JM, Penzo M. Targeting the Malaria Parasite cGMP-Dependent Protein Kinase to Develop New Drugs. Front Microbiol. 2020;11:602803.
62. Vanaerschot M, Murithi JM, Pasaje CFA, Ghidelli-Disse S, Dwomoh L, Bird M, *et al.* Inhibition of Resistance-Refractory *P. falciparum* Kinase PKG Delivers Prophylactic, Blood Stage, and Transmission-Blocking Antiplasmodial Activity. Cell Chem Biol. 2020;27(7):806-16.e8.
63. Müller IB, Das Gupta R, Lüersen K, Wrenger C, Walter RD. Assessing the polyamine metabolism of *Plasmodium falciparum* as chemotherapeutic target. Mol Biochem Parasitol. 2008;160(1):1-7.
64. McCann PP, Bacchi CJ, Clarkson AB, Jr., Seed JR, Nathan HC, Amole BO, *et al.* Further studies on difluoromethylornithine in African trypanosomes. Med Biol. 1981;59(5-6):434-40.
65. Meyskens FL, Jr., Gerner EW. Development of difluoromethylornithine (DFMO) as a chemoprevention agent. Clin Cancer Res. 1999;5(5):945-51.
66. Marks PA. Discovery and development of SAHA as an anticancer agent. Oncogene. 2007;26(9):1351-6.
67. Marfurt J, Chalfein F, Prayoga P, Wabiser F, Kenangalem E, Piera KA, *et al*. *Ex vivo* activity of histone deacetylase inhibitors against multidrug-resistant clinical isolates of *Plasmodium falciparum* and *P. vivax*. Antimicrob Agents Chemother. 2011;55(3):961-6.
68. Basore K, Cheng Y, Kushwaha AK, Nguyen ST, Desai SA. How do antimalarial drugs reach their intracellular targets? Front Pharmacol. 2015;6:91.
69. Croston GE. The utility of target-based discovery. Expert Opin Drug Discov. 2017;12(5):427-9.
70. Wagner BK. The resurgence of phenotypic screening in drug discovery and development. Expert Opin Drug Discov. 2016;11(2):121-5.
71. Schenone M, Dančík V, Wagner BK, Clemons PA. Target identification and mechanism of action in chemical biology and drug discovery. Nat Chem Biol. 2013;9(4):232-40.
72. Yang T, Ottilie S, Istvan ES, Godinez-Macias KP, Lukens AK, Baragaña B, *et al.* MalDA, Accelerating Malaria Drug Discovery. Trends Parasitol. 2021;37(6):493-507.
73. Zheng W, Thorne N, McKew JC. Phenotypic screens as a renewed approach for drug discovery. Drug Discov Today. 2013;18(21-22):1067-73.
74. Okombo J, Kanai M, Deni I, Fidock DA. Genomic and Genetic Approaches to Studying Antimalarial Drug Resistance and *Plasmodium* Biology. Trends Parasitol. 2021;37(6):476-92.
75. Cowell AN, Winzeler EA. Advances in omics-based methods to identify novel targets for malaria and other parasitic protozoan infections. Genome Med. 2019;11(1):63.
76. Challis MP, Devine SM, Creek DJ. Current and emerging target identification methods for novel antimalarials. Int J Parasitol Drugs Drug Resist. 2022;20:135-44.
77. Luth MR, Gupta P, Ottilie S, Winzeler EA. Using in Vitro Evolution and Whole Genome Analysis To Discover Next Generation Targets for Antimalarial Drug Discovery. ACS Infect Dis. 2018;4(3):301-14.
78. Paquet T, Le Manach C, Cabrera DG, Younis Y, Henrich PP, Abraham TS, *et al.* Antimalarial efficacy of MMV390048, an inhibitor of *Plasmodium* phosphatidylinositol 4-kinase. Sci Transl Med. 2017;9(387).

© University of Pretoria

79. Azad MA, Wright GD. Determining the mode of action of bioactive compounds. Bioorg Med Chem. 2012;20(6):1929-39.
80. Dziekan JM, Wirjanata G, Dai L, Go KD, Yu H, Lim YT, *et al.* Cellular thermal shift assay for the identification of drug-target interactions in the *Plasmodium falciparum* proteome. Nat Protoc. 2020;15(6):1881-921.
81. Caballero IM, Lundgren S. A Shift in Thinking: Cellular Thermal Shift Assay-Enabled Drug Discovery. ACS Med Chem Lett. 2023;14(4):369-75.
82. Huang L, Wang D, Zhang C. Drug Affinity Responsive Target Stability (DARTS) Assay to Detect Interaction Between a Purified Protein and a Small Molecule. Methods Mol Biol. 2021;2213:175-82.
83. Pai MY, Lomenick B, Hwang H, Schiestl R, McBride W, Loo JA, *et al*. Drug affinity responsive target stability (DARTS) for small-molecule target identification. Methods Mol Biol. 2015;1263:287-98.
84. Lomenick B, Hao R, Jonai N, Chin RM, Aghajan M, Warburton S, *et al.* Target identification using drug affinity responsive target stability (DARTS). Proc Natl Acad Sci U S A. 2009;106(51):21984-9.
85. Tulloch LB, Menzies SK, Coron RP, Roberts MD, Florence GJ, Smith TK. Direct and indirect approaches to identify drug modes of action. IUBMB Life. 2018;70(1):9-22.
86. Allman EL, Painter HJ, Samra J, Carrasquilla M, Llinás M. Metabolomic Profiling of the Malaria Box Reveals Antimalarial Target Pathways. Antimicrob Agents Chemother. 2016;60(11):6635-49.
87. Perez de Souza L, Alseekh S, Scossa F, Fernie AR. Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. Nat Methods. 2021;18(7):733-46.
88. Iwata M, Sawada R, Iwata H, Kotera M, Yamanishi Y. Elucidating the modes of action for bioactive compounds in a cell-specific manner by large-scale chemically-induced transcriptomics. Sci Rep. 2017;7:40164.
89. Lamb J, Crawford ED, Peck D, Modell JW, Blat IC, Wrobel MJ, *et al.* The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. Science. 2006;313(5795):1929-35.
90. Vásquez-Bochm LX, Velázquez-Paniagua M, Castro-Vázquez SS, Guerrero-Rodríguez SL, Mondragon-Peralta A, De La Fuente-Granada M, *et al.* Transcriptome-based identification of lovastatin as a breast cancer stem cell-targeting drug. Pharmacol Rep. 2019;71(3):535-44.
91. Baillif B, Wichard J, Méndez-Lucio O, Rouquié D. Exploring the Use of Compound-Induced Transcriptomic Data Generated From Cell Lines to Predict Compound Activity Toward Molecular Targets. Front Chem. 2020;8:296.
92. Iorio F, Bosotti R, Scacheri E, Belcastro V, Mithbaokar P, Ferriero R, *et al*. Discovery of drug mode of action and drug repositioning from transcriptional responses. Proc Natl Acad Sci U S A. 2010;107(33):14621-6.
93. Mashima T, Ushijima M, Matsuura M, Tsukahara S, Kunimasa K, Furuno A, et al. Comprehensive transcriptomic analysis of molecularly targeted drugs in cancer for target pathway evaluation. Cancer Sci. 2015;106(7):909-20.
94. Boshoff HI, Myers TG, Copp BR, McNeil MR, Wilson MA, Barry CE, 3rd. The transcriptional responses of *Mycobacterium tuberculosis* to inhibitors of metabolism: novel insights into drug mechanisms of action. J Biol Chem. 2004;279(38):40174-84.
95. Murima P, de Sessions PF, Lim V, Naim AN, Bifani P, Boshoff HI, *et al.* Exploring the mode of action of bioactive compounds by microfluidic transcriptional profiling in mycobacteria. PLOS ONE. 2013;8(7):e69191.
96. Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. PLOS Biol. 2003;1(1):E5.
97. Hollin T, Le Roch KG. From Genes to Transcripts, a Tightly Regulated Journey in *Plasmodium*. Front Cell Infect Microbiol. 2020;10:618454.
98. Bonnell VA, Zhang Y, Brown AS, Horton J, Josling GA, Chiu TP, *et al.* DNA sequence and chromatin differentiate sequence-specific transcription factor binding in the human malaria parasite *Plasmodium falciparum*. Nucleic Acids Res. 2024;52(17):10161-79.
99. Watzlowik MT, Das S, Meissner M, Längst G. Peculiarities of *Plasmodium falciparum* Gene Regulation and Chromatin Structure. Int J Mol Sci. 2021;22(10).
100. Tuteja R, Ansari A, Chauhan VS. Emerging Functions of Transcription Factors in Malaria Parasite. Biomed Res Int. 2011;2011(1):461979.
101. Kensche PR, Hoeijmakers WA, Toenhake CG, Bras M, Chappell L, Berriman M, *et al.* The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. Nucleic Acids Res. 2016;44(5):2110-24.
102. Sepulveda LFO, de Oliveira GM, Chagas EHN, Wild N, da Paz FS, Wrenger C, *et al.* Plasmodial Transcription Factors and Chromatin Modifiers as Drug Targets. Future pharmacol. 2023;3(4):846-61.

© University of Pretoria

103. Hu G, Cabrera A, Kono M, Mok S, Chaal BK, Haase S, *et al.* Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. Nat Biotechnol. 2010;28(1):91-8.
104. Siwo GH, Smith RS, Tan A, Button-Simons KA, Checkley LA, Ferdig MT. An integrative analysis of small molecule transcriptional responses in the human malaria parasite *Plasmodium falciparum*. BMC Genomics. 2015;16:1030.
105. Niemand J, van Biljon R, van der Watt M, van Heerden A, Reader J, van Wyk R, *et al*. Chemogenomic Fingerprints Associated with Stage-Specific Gametocytocidal Compound Action against Human Malaria Parasites. ACS Infect Dis. 2021;7(10):2904-16.
106. van der Watt ME, Reader J, Churchyard A, Nondaba SH, Lauterbach SB, Niemand J, *et al.* Potent *Plasmodium falciparum* gametocytocidal compounds identified by exploring the kinase inhibitor chemical space for dual active antimalarials. J Antimicrob Chemother. 2018;73(5):1279-90.
107. van Heerden A, van Wyk R, Birkholtz L-M. Machine Learning Uses Chemo-Transcriptomic Profiles to Stratify Antimalarial Compounds With Similar Mode of Action. Front cell infect microbiol. 2021;11.
108. Andrews KT, Gupta AP, Tran TN, Fairlie DP, Gobert GN, Bozdech Z. Comparative gene expression profiling of *P. falciparum* malaria parasites exposed to three different histone deacetylase inhibitors. PLOS ONE. 2012;7(2):e31847.
109. van Brummelen AC, Olszewski KL, Wilinski D, Llinás M, Louw AI, Birkholtz LM. Co-inhibition of *Plasmodium falciparum* S-adenosylmethionine decarboxylase/ornithine decarboxylase reveals perturbation-specific compensatory mechanisms by transcriptome, proteome, and metabolome analyses. J Biol Chem. 2009;284(7):4635-46.
110. Cheemadan S, Ramadoss R, Bozdech Z. Role of calcium signaling in the transcriptional regulation of the apicoplast genome of *Plasmodium falciparum*. Biomed Res Int. 2014;2014:869401.
111. van der Watt ME, Reader J, Churchyard A, Nondaba SH, Lauterbach SB, Niemand J, *et al.* Potent *Plasmodium falciparum* gametocytocidal compounds identified by exploring the kinase inhibitor chemical space for dual active antimalarials. J Antimicrob Chemother. 2018;73(5):1279-90.
112. Hu G, Cabrera A, Kono M, Mok S, Chaal BK, Haase S, *et al.* Transcriptional profiling of growth perturbations of the human malaria parasite *Plasmodium falciparum*. Nat Biotechnol. 2010;28(1):91-8.
113. Shieh S-L, Liao IE. A new approach for data clustering and visualization using self-organizing maps. Expert Syst Appl. 2012;39(15):11924-33.
114. Jensen JB, Trager W. *Plasmodium falciparum* in culture: use of outdated erythrocytes and description of the candle jar method. J Parasitol. 1977;63(5):883-6.
115. Reader J, Botha M, Theron A, Lauterbach SB, Rossouw C, Engelbrecht D, *et al.* Nowhere to hide: interrogating different metabolic parameters of *Plasmodium falciparum* gametocytes in a transmission blocking drug discovery pipeline towards malaria elimination. Malar J. 2015;14:213.
116. van Biljon R, Niemand J, van Wyk R, Clark K, Verlinden B, Abrie C, *et al.* Inducing controlled cell cycle arrest and re-entry during asexual proliferation of *Plasmodium falciparum* malaria parasites. Sci Rep. 2018;8(1):16581.
117. Lambros C, Vanderberg JP. Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. J Parasitol. 1979;65(3):418-20.
118. Radfar A, Méndez D, Moneriz C, Linares M, Marín-García P, Puyet A, *et al.* Synchronous culture of *Plasmodium falciparum* at high parasitemia levels. Nat Protoc. 2009;4(12):1899-915.
119. Mbengue A, Yam XY, Braun-Breton C. Human erythrocyte remodelling during *Plasmodium falciparum* malaria parasite growth and egress. Br J Haematol. 2012;157(2):171-9.
120. Smilkstein M, Sriwilaijaroen N, Kelly JX, Wilairat P, Riscoe M. Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. Antimicrob Agents Chemother. 2004;48(5):1803-6.
121. Zhang ZW, Cheng J, Xu F, Chen YE, Du JB, Yuan M, *et al.* Red blood cell extrudes nucleus and mitochondria against oxidative stress. IUBMB Life. 2011;63(7):560-5.
122. Painter HJ, Chung NC, Sebastian A, Albert I, Storey JD, Llinás M. Genome-wide real-time in vivo transcriptional dynamics during *Plasmodium falciparum* blood-stage development. Nat Commun. 2018;9(1):2656.
123. Salanti A, Staalsoe T, Lavstsen T, Jensen AT, Sowa MP, Arnot DE, *et al.* Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. Mol Microbiol. 2003;49(1):179-91.

© University of Pretoria

124. Jensen AT, Magistrado P, Sharp S, Joergensen L, Lavstsen T, Chiucchiuini A, *et al. Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes. J Exp Med. 2004;199(9):1179-90.
125. Rodríguez A, Rodríguez M, Córdoba JJ, Andrade MJ. Design of primers and probes for quantitative real-time PCR methods. Methods Mol Biol. 2015;1275:31-56.
126. Walz A, Duffey M, Aljayyoussi G, Sax S, Leroy D, Besson D, *et al.* The Parasite Reduction Ratio (PRR) Assay Version 2: Standardized Assessment of *Plasmodium falciparum* Viability after Antimalarial Treatment In Vitro. Pharmaceuticals (Basel). 2023;16(2).
127. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402-8.
128. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. 2001;29(9):e45.
129. Fang H, Gough J. supraHex: an R/Bioconductor package for tabular omics data analysis using a supra-hexagonal map. Biochem Biophys Res Commun. 2014;443(1):285-9.
130. WHO. Giemsa Staining of Malaria Blood Films. Geneva; 2016.
131. Rajendran V, Gurukkalot K. In vitro drug interaction of ionophores with artemisinin and chloroquine against *Plasmodium falciparum* 3D7 blood-stage infection. Front Drug Discov. 2023;3.
132. de Lange C, Coertzen D, Smit FJ, Wentzel JF, Wong HN, Birkholtz LM, *et al.* Synthesis, antimalarial activities and cytotoxicities of amino-artemisinin-1,2-disubstituted ferrocene hybrids. Bioorg Med Chem Lett. 2018;28(19):3161-3.
133. Llanos-Cuentas A, Casapia M, Chuquiyauri R, Hinojosa JC, Kerr N, Rosario M, *et al.* Antimalarial activity of single-dose DSM265, a novel *plasmodium* dihydroorotate dehydrogenase inhibitor, in patients with uncomplicated P*lasmodium falciparum* or *Plasmodium vivax* malaria infection: a proof-of-concept, open-label, phase 2a study. Lancet Infect Dis. 2018;18(8):874-83.
134. Sandefur CI, Wooden JM, Quaye IK, Sirawaraporn W, Sibley CH. Pyrimethamine-resistant dihydrofolate reductase enzymes of *Plasmodium falciparum* are not enzymatically compromised in vitro. Mol Biochem Parasitol. 2007;154(1):1-5.
135. Phillips MA, Lotharius J, Marsh K, White J, Dayan A, White KL, *et al*. A long-duration dihydroorotate dehydrogenase inhibitor (DSM265) for prevention and treatment of malaria. Sci Transl Med. 2015;7(296):296ra111.
136. Andriantsoanirina V, Durand R, Pradines B, Baret E, Bouchier C, Ratsimbasoa A, *et al*. *In vitro* susceptibility to pyrimethamine of DHFR I164L single mutant *Plasmodium falciparum*. Malar J. 2011;10:283.
137. Brunschwig C, Lawrence N, Taylor D, Abay E, Njoroge M, Basarab GS, *et al.* UCT943, a Next-Generation *Plasmodium falciparum* PI4K Inhibitor Preclinical Candidate for the Treatment of Malaria. Antimicrob Agents Chemother. 2018;62(9).
138. Batugedara G, Lu XM, Hristov B, Abel S, Chahine Z, Hollin T, *et al*. Novel insights into the role of long non-coding RNA in the human malaria parasite, *Plasmodium falciparum*. Nat Commun. 2023;14(1):5086.
139. Masotti A, Caputo V, Da Sacco L, Pizzuti A, Dallapiccola B, Bottazzo GF. Quantification of small non-coding RNAs allows an accurate comparison of miRNA expression profiles. J Biomed Biotechnol. 2009;2009:659028.
140. Chappell L, Ross P, Orchard L, Russell TJ, Otto TD, Berriman M, *et al.* Refining the transcriptome of the human malaria parasite *Plasmodium falciparum* using amplification-free RNA-seq. BMC Genomics. 2020;21(1):395.
141. Lucena-Aguilar G, Sánchez-López AM, Barberán-Aceituno C, Carrillo-Ávila JA, López-Guerrero JA, Aguilar-Quesada R. DNA Source Selection for Downstream Applications Based on DNA Quality Indicators Analysis. Biopreserv Biobank. 2016;14(4):264-70.
142. von Ahsen N, Wittwer CT, Schütz E. Oligonucleotide melting temperatures under PCR conditions: nearest-neighbor corrections for Mg(2+), deoxynucleotide triphosphate, and dimethyl sulfoxide concentrations with comparison to alternative empirical formulas. Clin Chem. 2001;47(11):1956-61.

© University of Pretoria

Supplementary

Supplementary Table 1. Primer information.

Primer name:	Sequence (5'→3')	IR	Tm (°C)	ΔG (kJ) H	ΔG (kJ) M	GC content (%)	Length (bp)	Ta (°C)*
1. PF3D7_0108700 (FWD)	TGATACGGAACAAAGCGTGGAGAAC	39	59.4	-15.90	0.00	48	25	62.98
1. PF3D7_0108700 (REV)	AGGGGGTGGCGCATAAGAAAAAG	39	59.6	-24.85	0.00	52	23	62.36
2. PF3D7_0203700 (FWD)	TGATAAAGACAGCGTTGAAGAAAGTACCTC	50	58.9	-17.87	0.00	40	30	64.03
2. PF3D7_0203700 (REV)	GTCCATTGCCAGCTTCCCATTTG	50	58.8	-15.31	0.00	52	23	62.36
3. PF3D7_0206700 (FWD)	ACATGGACAACCTGCATCACCAAC	13	59.8	-16.48	0.00	50	24	62.72
3. PF3D7_0206700 (REV)	CAATCAATATTAGGATCACAGACTTTATGCGC	13	58.5	-22.93	0.00	37.5	32	64.36
4. PF3D7_0508700 (FWD)	TTTAATAGCGACATCTGTTATGGCTAGAGG	23	58.3	-15.82	0.00	40	30	64.03
4. PF3D7_0508700 (REV)	AGATCTTCCGGTTCTCCCTATTCGATG	23	59.4	-22.84	0.00	48	27	64.91
5. PF3D7_0618100 (FWD)	ACGTTTCGCATCATGAGGCCTATG	10	59.5	-32.84	0.00	50	24	62.72
5. PF3D7_0618100 (REV)	ACCATTTTGGCTAGTTTGAGCTAGCG	10	59.7	-31.67	-11.55	46.2	26	63.24
6. PF3D7_1308500 (FWD)	TGACTAAGAGGGATGAGGAAGATAAATTCG	41	57.8	-8.66	0.00	40	30	64.03
6. PF3D7_1308500 (REV)	CTTTTTCCACAACCTGAACAACCGC	41	59.5	-6.28	0.00	48	25	62.98
7. PF3D7_0322100 (FWD)	AATTGAAATTGAAGGAAGGGTGGGGTTAG	44	59.2	-10.59	0.00	41	29	63.70
7. PF3D7_0322100 (REV)	ACGATTCTCTATCAATTCCGGCCTGG	44	60.3	-31.51	0.00	50	26	64.80
8. PF3D7_0614300 (FWD)	TTATTGCTTTAACGGCACATGCAGATATG	25	58.6	-20.38	0.00	37	29	62.06
8. PF3D7_0614300 (REV)	CGATGGAAGACAAGCATGCGC	25	59.2	-30.33	0.00	57	21	61.72
9. PF3D7_1015500 (FWD)	TGCTGAACCAAAAATACTTCCGTCACTTC	42	59.8	-8.62	0.00	41.4	29	63.86
9. PF3D7_1015500 (REV)	CGTCTTGACAAAATTTGCAATCAACACCC	42	60.1	-29.87	-2.47	41.4	29	63.86
10. PF3D7_1039000 (FWD)	GTGAGGCAGAAAGCAATGGAGGAATTC	40	60.2	-17.11	0.00	48	27	64.91
10. PF3D7_1039000 (REV)	CGCTACCATACCCATTCGCTTCTTC	40	59.6	-8.03	0.00	52	25	64.62
11. PF3D7_1112700 (FWD)	CGTATATGTACGACAACAGGTATACACTCC	11	58.3	-22.97	-3.18	43.3	30	65.39
11. PF3D7_1112700 (REV)	GTCCTTGTCAATCCCTTTTGCTATCTTTAG	11	57.9	-10.63	0.00	40	30	64.03
12. PF3D7_1323800 (FWD)	AAGTTGGATATCAAGAGCAAAAACTCAAGG	38	58	-16.23	0.00	36.7	30	62.68
12. PF3D7_1323800 (REV)	TCCGCTTCGCCATGTAATATATTACATCC	38	58.7	-42.72	-3.77	41.4	29	63.86
13. PF3D7_1438000 (FWD)	TCCAACAAACAGAGGAACCTGTTGATATC	45	58.8	-19.71	-2.55	41.4	29	63.86
13. PF3D7_1438000 (REV)	ACCAGTAGATCTTGGTGCTCTATATATCCC	45	58.5	-30.00	0.00	43.3	30	65.39
14. PF3D7_0623900 (FWD)	AGTAGTAAGTGCAGCGTCAATTTGTGC	3	60.1	-20.29	0.00	44.4	27	63.43
14. PF3D7_0623900 (REV)	GGGTCTCCTGGATAACCCGATCC	3	59.7	-48.83	-5.61	60.9	23	66.01
15. PF3D7_1030600 (FWD)	ACATTAATATGTTACACGAAAGGTCCCGG	31	58.9	-18.49	0.00	41	29	63.70
15. PF3D7_1030600 (REV)	TCCCCATTTCGATATGAGCAATACAATG	31	57.5	-17.15	0.00	39	28	62.08
16. PF3D7_1467400 (FWD)	CATAAAACCTGTTTTCCATACTCTCCAAGC	1	58.3	-15.52	0.00	40	30	64.03
16. PF3D7_1467400 (REV)	ACACGAATATGTGTATGGGGTTCTCTTATG	1	58.4	-12.76	-2.64	40	30	64.03
17. PF3D7_0206100 (FWD)	CTCTTTCAATCGATGCCCAACAGTCC	26	60	-22.43	0.00	50	26	64.80
17. PF3D7_0206100 (REV)	AATCCGTCTGAATGACCCAACCATAC	26	58.7	-9.00	0.00	46	26	63.16
18. PF3D7_0619800 (FWD)	GAAATGGAAGGAGGCATTTTGTTTATTGTTAAAT	14	57.8	-14.48	0.00	29	34	62.07
18. PF3D7_0619800 (REV)	TCTTAATTTTTCTTGTGACATTTTTTGTCCACAC	14	58	-21.00	-5.73	29	34	62.07

© University of Pretoria

Primer name:	Sequence (5'→3')	IR	Tm (°C)	ΔG (kJ) H	ΔG (kJ) M	GC content (%)	Length (bp)	Ta (°C)*
19. PF3D7_0806600 (FWD)	CCTTGCACAACGGCATCGAAAC	43	59.6	-17.99	0.00	54	22	61.89
19. PF3D7_0806600 (REV)	ACATTTTGTTTAGCGACGTAAGGGAGG	43	59.4	-18.74	0.00	44	27	63.27
20.PF3D7_1220400 (FWD)	GTTATGTTATACCAATGGACTTATTAGAGGAAGC	36	57.4	-19.16	0.00	35	34	64.53
20. PF3D7_1220400 (REV)	CAACAAGTGAGAAGTTGACAGATACGTAAG	36	57.9	-23.93	-0.96	40	30	64.03
21. PF3D7_1317100 (FWD)	GAAGTACAAGAACCCGTGACTTGCTC	37	59.6	-23.18	-2.59	50	26	64.80
21. PF3D7_1317100 (REV)	AGAAATTGACTGTGGCGTTTCTCCTTG	37	59.8	-14.10	0.00	44	27	63.27
22. PF3D7_1325400 (FWD)	GTTGTCGCGCTCGCCTG	48	58.9	-24.10	0.00	70	17	59.76
22. PF3D7_1325400 (REV)	ACAAAACATGGAGTTGGTGTCTTCCC	48	59.5	-16.23	-2.18	46	26	63.16
23. PF3D7_1475100 (FWD)	ACAGCGTGGTATAAAGAAAGAGAGGTGTC	8	60.1	-13.51	0.00	44	29	64.93
23. PF3D7_1475100 (REV)	CGTCCCATTTGTATCATCAACTGTTTGTTC	8	59	-13.85	0.00	40	30	64.03
24. PF3D7_0503400 (FWD)	TGCATATGTTGTGTTCGATGCAGTAAGC	21	60.2	-20.59	-2.34	42	28	63.31
24. PF3D7_0503400 (REV)	ATCAGCAACATCTTGTGCACTTTCAATAAC	21	58.9	-27.61	-2.34	36	30	62.39
25. PF3D7_0509100 (FWD)	CGGATGAGGAAGAGCCAATCGTC	46	59.3	-13.31	-0.71	56	23	64.00
25. PF3D7_0509100 (REV)	TGGTCATCATCACCGTCATCTTCATCATC	46	60.3	-11.59	0.00	44	29	64.93
26. PF3D7_1019800 (FWD)	TGATGAATATTTTGATGAAGTGCTTGTACCAATG	17	58.6	-14.56	0.00	32	34	63.30
26. PF3D7_1019800 (REV)	CATTCGTTCTTATCATAGCATAATGTCCAGTAG	17	57.6	-12.76	0.00	36	33	64.36
27. PF3D7_1242700 (FWD)	ACAGAGGAAGTAGCCATAATTCCCTCAAAG	34	59.7	-20.84	-6.82	43	30	65.26
27. PF3D7_1242700 (REV)	ACCACGTACAGGTCCCTTTTGTATTC	34	58.6	-27.20	0.00	46	26	63.16
28. PF3D7_1425800 (FWD)	TGCTGGAAAACTAGCCGAAAAGGAATG	4	59.8	-11.63	-2.05	44	27	63.27
28. PF3D7_1425800 (REV)	TTTCCCTTGTCTTTTTCTAGCACATCTTTG	4	58.2	-10.88	0.00	36	30	62.39
29. PF3D7_1440500 (FWD)	GAAGAAGAAGAGATGTTGGACATGAATGG	47	57.9	-13.43	0.00	41	29	63.70
29. PF3D7_1440500 (REV)	GTTACCAACCTTATCCCGCTCGTTC	47	59.5	-9.79	0.00	52	25	64.62
30. PF3D7_0317300 (FWD)	ACACATCTGGGTTTTATGGGTGCATC	29	59.4	-18.07	0.00	46	26	63.16
30. PF3D7_0317300 (REV)	GTGCTTGTGTTGAATTTCCTTCGGC	29	59.6	-9.00	0.00	48	25	62.98
31. PF3D7_1013500 (FWD)	GTTGGTCAAAACCATCCACTTTCCAAATG	16	59.4	-21.38	-5.86	41.1	29	63.74
31. PF3D7_1013500 (REV)	GCATTTAACTGGAAAGCAGGCGC	16	59.7	-23.93	0.00	52.2	23	62.44
32. PF3D7_1127900 (FWD)	CCTCAGTGATATTAGCGCGTATGCAG	12	59.5	-27.78	0.00	50	26	64.80
32. PF3D7_1127900 (REV)	CCTGAATGGTCAGGTCGTCTTTTTCAC	12	59.9	-29.79	-8.20	48.1	27	64.95
33. PF3D7_1352000 (FWD)	TGAGTCGGAGCAAGAGAGGACG	28	59.9	-11.63	0.00	59.1	22	63.99
33. PF3D7_1352000 (REV)	ACCATCATTGTCAATATCCATGTTGACCC	28	59.1	-22.93	-1.76	41.4	29	63.86
34. PF3D7_1474500 (FWD)	TTTAGTAAATGGTGACGTGGAAGGTTTAAC	24	57.8	-15.40	0.00	36.7	30	62.68
34. PF3D7_1474500 (REV)	GCAAAACGTGCCCTAGGAACAAC	24	59	-28.87	0.00	52.2	23	62.44
35. PF3D7_0612600 (FWD)	ACAGTCTGTGGAGTCTTCAGAAGACAATC	33	59.8	-39.71	-5.77	44.8	29	65.25
35. PF3D7_0612600 (REV)	GGTTTCCGCCAAATCATCAGCATTATG	33	59.2	-12.64	0.00	44.4	27	63.43
36. PF3D7_0704500 (FWD)	TCATACATGTGGAACGCAAAACATGAATAC	5	59.6	-20.38	-1.88	36	30	62.39
36. PF3D7_0704500 (REV)	TTTCCTCATCAAATTGCTCTTTCGTCAAAC	5	58.8	-11.55	0.00	36	30	62.39

© University of Pretoria

Primer name:	Sequence (5'→3')	IR	Tm (°C)	ΔG (kJ) H	ΔG (kJ) M	GC content (%)	Length (bp)	Ta (°C)*
37. PF3D7_1324000 (FWD)	GATGATTCTAATGAAGACGGTGATGAAAGAC	6	57.9	-13.26	0.00	38	31	63.91
37. PF3D7_1324000 (REV)	CCTCCAAATAATTCATCTGCTCCACTACC	6	59	-8.62	0.00	44	29	64.93
38. PF3D7_0604100 (FWD)	GGGGGTATAGCCATAGTAACGACATTTTG	9	59	-28.16	0.00	44	29	64.93
38. PF3D7_0604100 (REV)	TGACTGGTTCTTGGCTACCTATGTGTTC	9	60.1	-9.50	0.00	46	28	64.95
39. PF3D7_1322200 (FWD)	AGGTAAGAAGAGCAACAGGGGTAATTTCC	19	60.1	-9.62	0.00	44	29	64.93
39. PF3D7_1322200 (REV)	TCAACGCATAGTTTGCTCCCTTCAAC	19	59.8	-12.22	0.00	46	26	63.16
40. PF3D7_1427000 (FWD)	AGGTGAATGTGTCAAGAGGATGGGTG	22	60.3	-9.25	0.00	50	26	64.80
40. PF3D7_1427000 (REV)	GTTTATTTTTCGTGAGCTTTCTTGTTTGTTCTC	22	58.4	-17.32	0.00	33	33	63.13
41. PF3D7_0511800 (FWD)	TGCCCTTGCAGCCATTTTAGAAAACTC	18	60.3	-17.91	-1.76	44	27	63.27
41. PF3D7_0511800 (REV)	TCCTGTTCCAAAGTAAAAGTCGAGGAGG	18	60	-15.56	0.00	46	28	64.95
42. PF3D7_0717800 (FWD)	GGACACAAGCTAACGTTCGTTCTGC	20	60.6	-32.80	0.00	52	25	64.62
42. PF3D7_0717800 (REV)	TGTGCTTTCCCAAGAAACGCTACG	20	60.1	-16.65	-5.06	50	24	62.72
43. PF3D7_0823800 (FWD)	TGGGTAGAAAAGCTTCTAATGATAGCTACG	27	58.2	-27.87	-6.23	40	30	64.03
43. PF3D7_0823800 (REV)	CCACCTATCATATTTCCTGAAAAATTACCATCC	27	57.8	-13.35	0.00	36	33	64.36
44. PF3D7_1115400 (FWD)	GAGCAGCACCAAATCATGCGG	49	59.1	-14.60	0.00	57	21	61.72
44. PF3D7_1115400 (REV)	CCTTCTCCCCAATCAGATCCCCATG	49	60.2	-11.00	0.00	56	25	66.26
45. PF3D7_1458900 (FWD)	CTTCTCCATATTTATTTATTTCCTAGGCCCC	35	57.1	-23.93	0.00	38.7	31	64.20
45. PF3D7_1458900 (REV)	CCTGAGTCTATCCACCATATCATTCTTACC	35	57.9	-10.71	0.00	43.3	30	65.39
46. PF3D7_1038400 (FWD)	AGGTGGTACCCGAAGAAGTTGTTGAAG	2	60.4	-22.59	0.00	48	27	64.91
46. PF3D7_1038400 (REV)	CATTTCTTCAACAACCTCTCCCGGTAC	2	59.4	-19.33	0.00	48	27	64.91
47. PF3D7_0213000 (FWD)	GGAAGTGATTATTTAATTTTGCTAAAACAAAAAGTGG	7	57.5	-24.02	-0.88	27	37	62.80
47. PF3D7_0213000 (REV)	CTAAGGTACCATATATATATCCACAACCTATCCC	7	57.2	-25.36	0.00	38	34	65.76
48. PF3D7_0301800 (FWD)	GAAGCAGTTGGTGGAGAAATAGGATTGAC	15	59.5	-9.62	0.00	44	29	64.93
48. PF3D7_0301800 (REV)	CCATTCTGTCAATTTCATGCCATTCAGC	15	59.4	-12.68	0.00	42	28	63.31
49. PF3D7_1340900 (FWD)	TGGGTGCAACGGTTTGGTTAGC	32	60.2	-24.85	0.00	54	22	61.89
49. PF3D7_1340900 (REV)	GATTTCATATTACCTGTCGCTAATCCAAAACC	32	58.4	-8.03	0.00	37	32	64.16
50. PF3D7_1404400 (FWD)	ACCTGTTCGATTAAAAGCCCCGAAG	30	59.4	-15.48	0.00	48	25	62.98
50. PF3D7_1404400 (REV)	CTCCTCTTATTCTTTTTGGACAGGCAGC	30	59.7	-12.84	0.00	46	28	64.95
ECG: PF3D7_0717700 (FWD)	CAAGTAGCAGGTCATCGTGG	N/A	54.8	-2.34	0.00	55	20	-
ECG: PF3D7_0717700 (REV	CAAGTTCGGCACATTCTTCCA	N/A	55.4	-3.24	0.00	47.6	21	-

© University of Pretoria

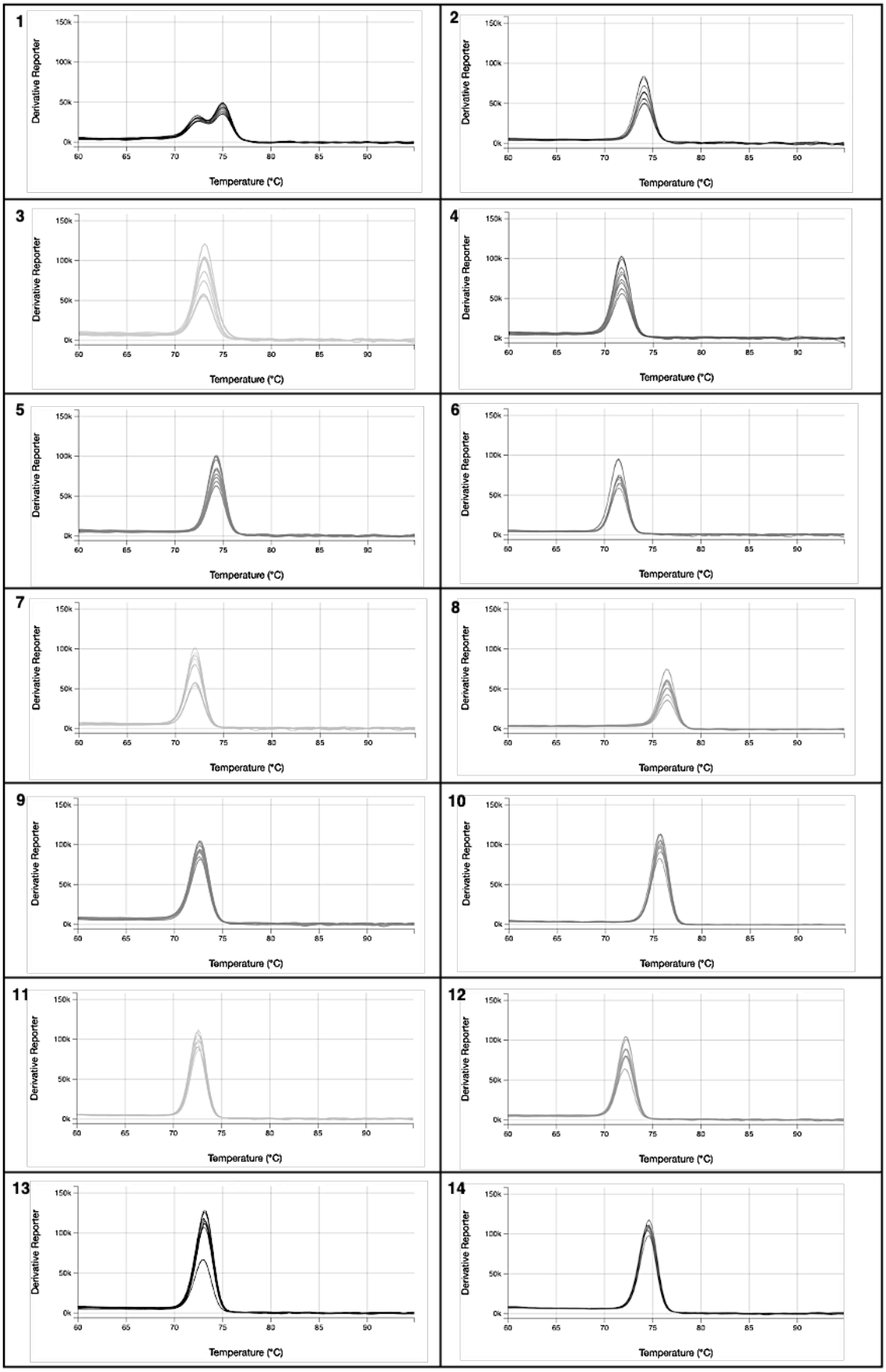
Supplementary Table 2. cDNA NanoDrop One^{C} readings.

Sample	Concentration (ng/μL)	A_{260}/A_{280}	A_{260}/A_{230}
Untreated 12 h	4.8	1.36	0.5
Artemisinin	7.5	1.32	0.74
Chloroquine	10.1	1.37	0.84
Untreated 24 h	11.6	1.62	1.14
DFMO	7.8	1.43	0.63
SAHA	15.4	1.69	0.74
Untreated 48 h	14.7	1.43	1.83
MMV048	16.6	1.51	1.75
UCT943	21.4	1.29	0.32
Untreated 30 h	25.8	1.20	1.00
DSM265	15.9	1.71	0.68
Pyrimethamine	14.5	1.75	1.14
Untreated 30 h (2)	13.9	1.64	1.27
MMV084	18.2	1.61	1.25
M5717	15.5	1.58	1.25

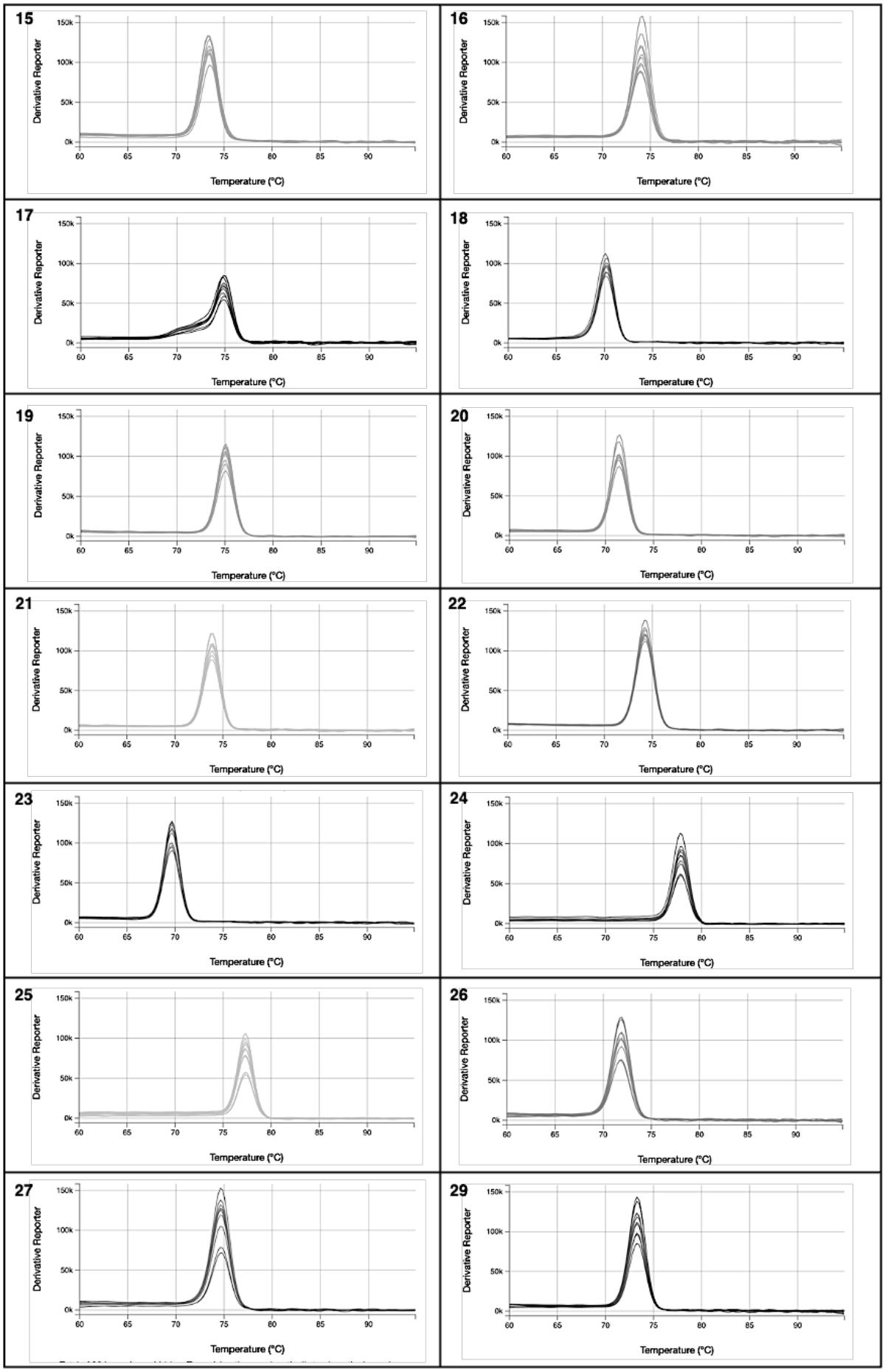
© University of Pretoria

Supplementary Table 3. Melt curves of each biomarker.

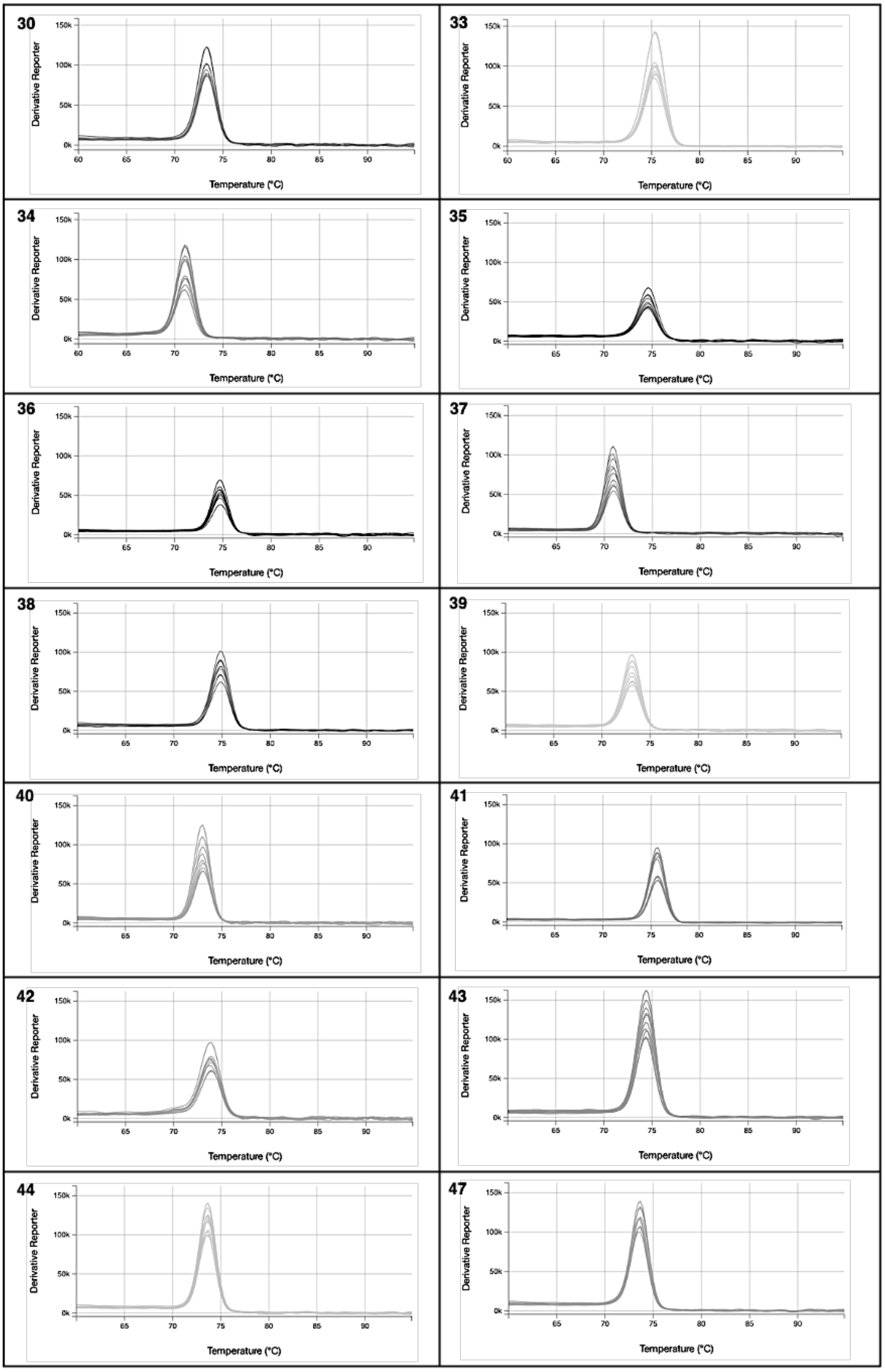
The numbers in the top corners correspond to the biomarker numbers.



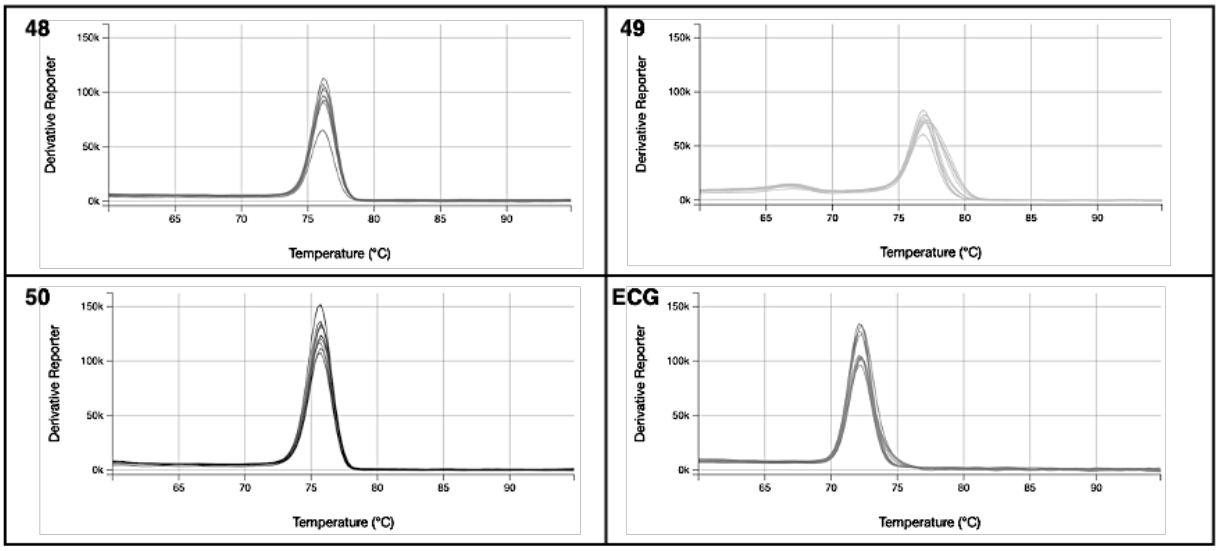
© University of Pretoria



© University of Pretoria

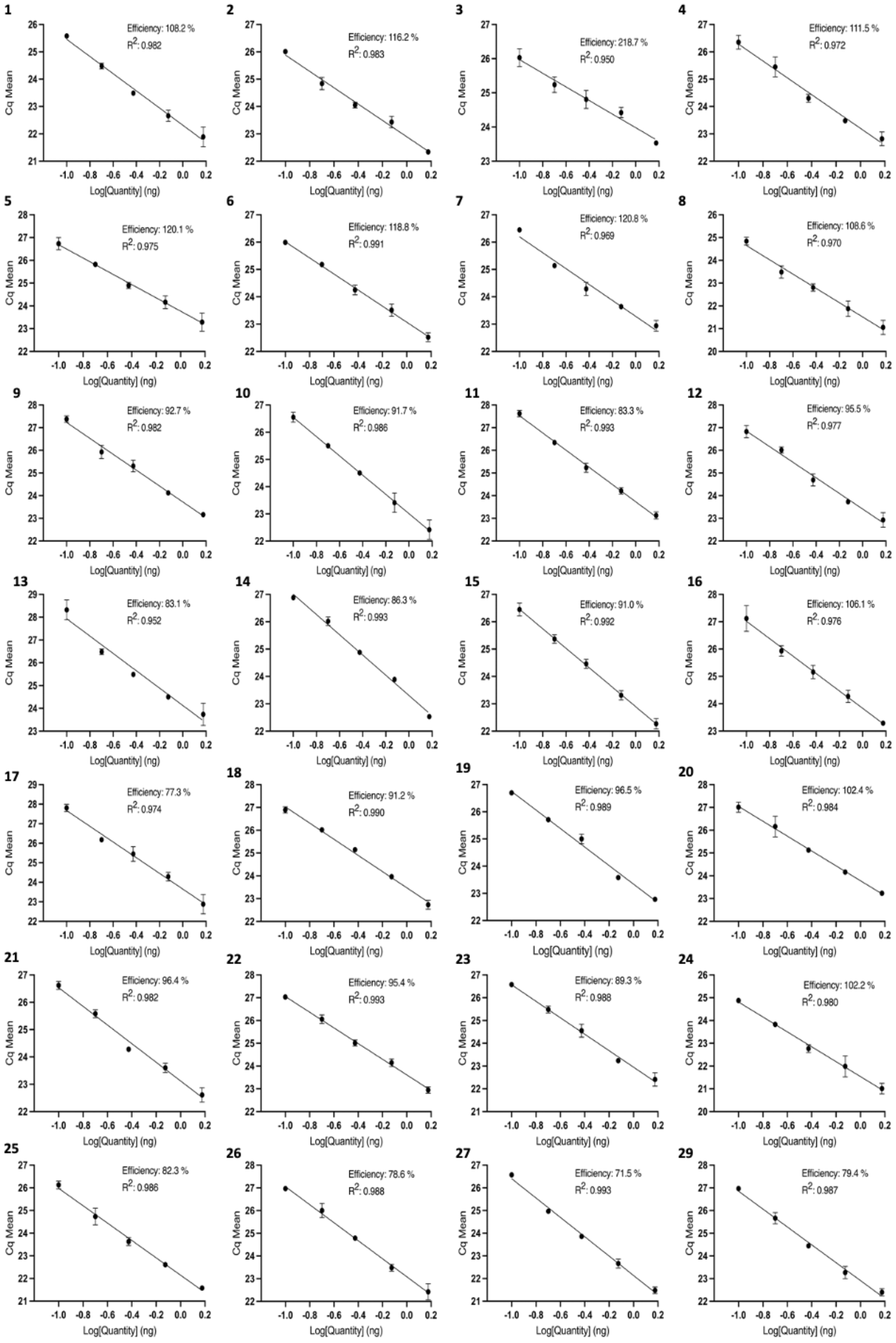


© University of Pretoria



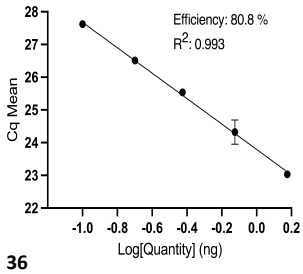
© University of Pretoria

Supplementary Table 4. Standard curves to assess primer efficiencies.

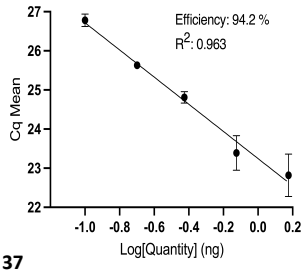


© University of Pretoria

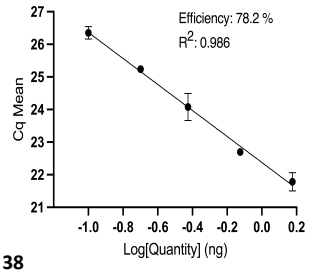
30



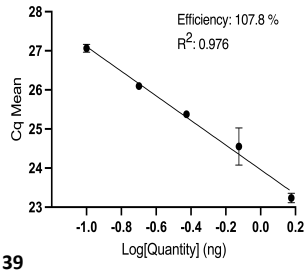
33



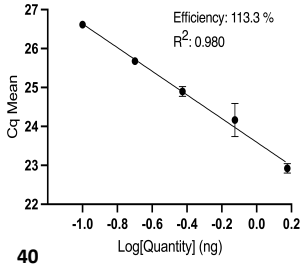
34



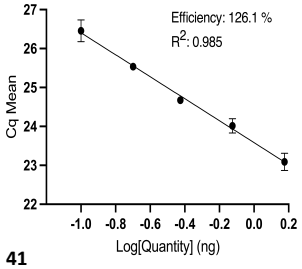
35



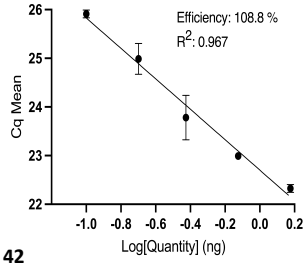
36



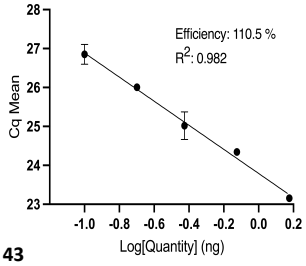
37



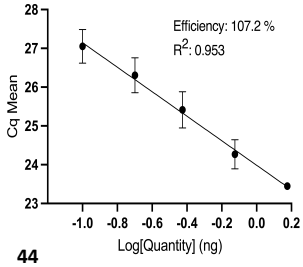
38



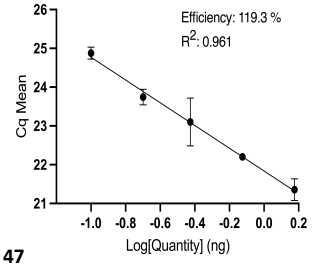
39



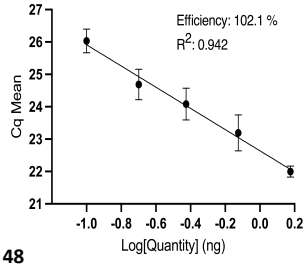
40



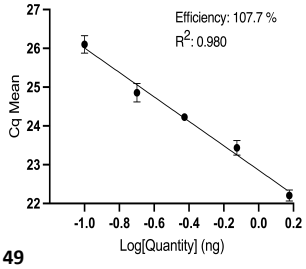
41



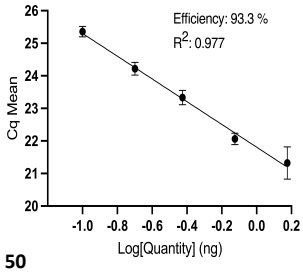
42



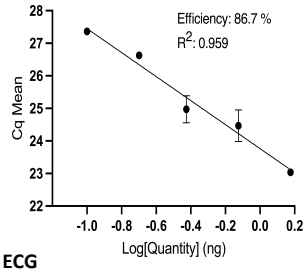
43



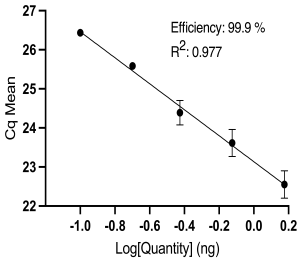
44



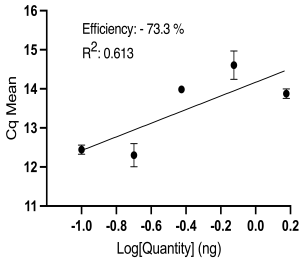
47



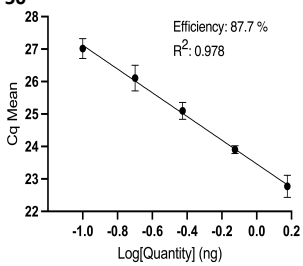
48



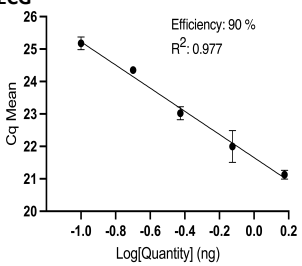
49



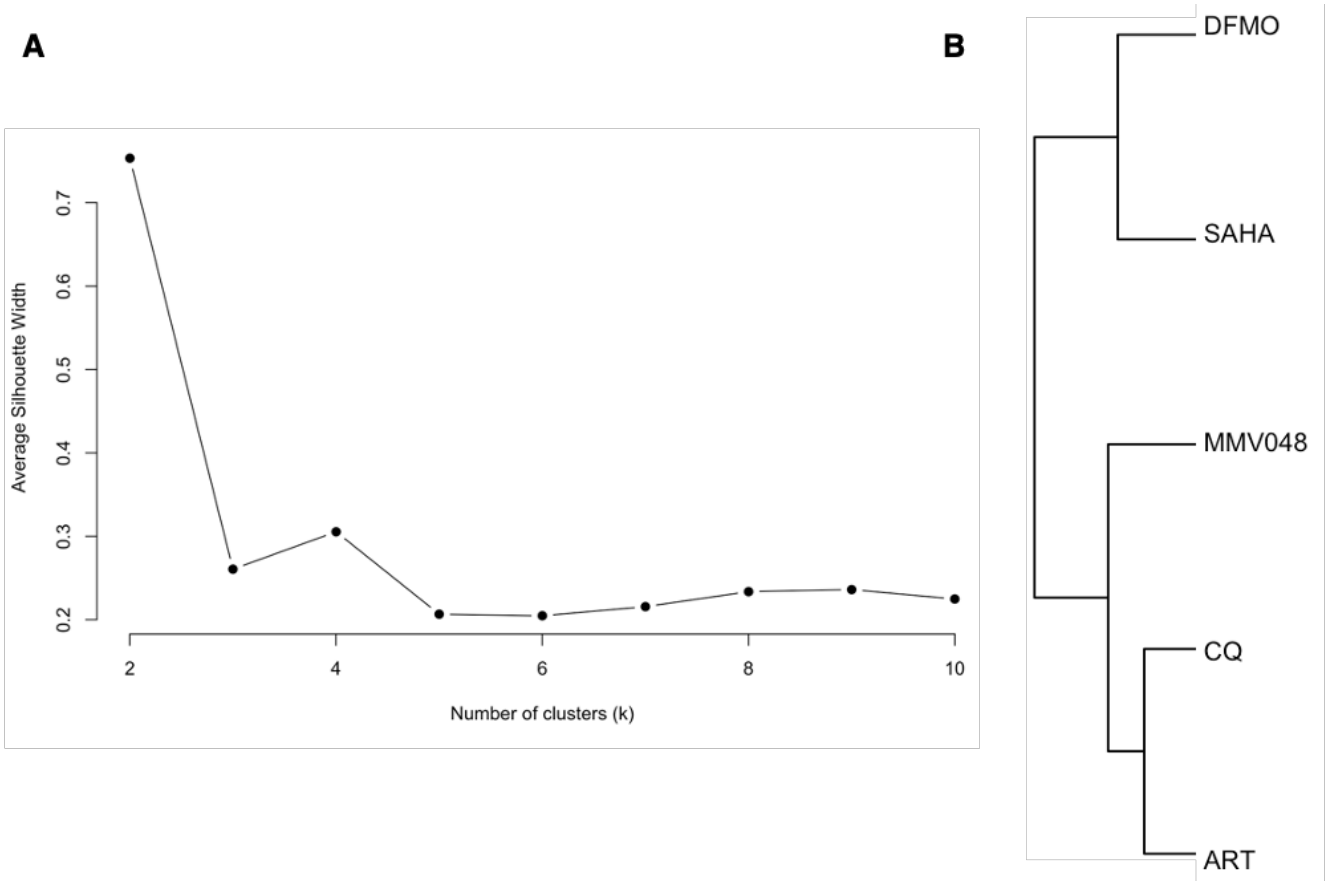
50



ECG



© University of Pretoria



Supplementary Figure 1. K-means clustering. (A) Silhouette analysis for an optimal number of clusters and (B) K-means clustering. The optimal number of clusters was determined as 2 (k = 2) that shows a silhouette score of 0.7. Silhouette scores fall between -1 and 1.

Supplementary Table 5a. Unpaired Student's t-test between parasite developmental stages of the 12 h incubation control compound treatments.

Rings	**Untreated 12 h**	**Chloroquine**
Untreated 12 h		
Chloroquine	0.845	
Artemisinin	0.834	0.933
Trophozoites	**Untreated 12 h**	**Chloroquine**
Untreated 12 h		
Chloroquine	0.845	
Artemisinin	0.834	0.933

© University of Pretoria

Supplementary Table 5b. Unpaired Student's t-test between parasite developmental stages of the 24 h incubation control compound treatments.

Rings	Untreated 24 h	DFMO
Untreated 24 h		
DFMO	1	
SAHA	0.058	0.058
Trophozoites	**Untreated 24 h**	**DFMO**
Untreated 24 h		
DFMO	1	
SAHA	0.058	0.058

Supplementary Table 5c. Unpaired Student's t-test between parasite developmental stages of the 48 h incubation control compound treatments.

Untreated 48h *vs.* MMV048	
Rings	0.00043
Trophozoites	0.067
Schizonts	0.185

Supplementary Table 6a. Unpaired Student's t-test between parasite developmental stages of the 30 h incubation clinical candidate treatments.

Rings	Untreated 30 h	DSM265	Pyrimethamine	MMV084
Untreated 30 h				
DSM265	0.286			
Pyrimethamine	N/A	0.170		
MMV084	N/A	0.286	N/A	
M5717	0.045	0.560	0.026	0.079
Trophozoites	**Untreated 30 h**	**DSM265**	**Pyrimethamine**	**MMV084**
Untreated 30 h				
DSM265	0.534			
Pyrimethamine	0.241	0.684		
MMV084	0.259	0.486	0.617	
M5717	0.838	0.560	0.201	0.151
Schizonts	**Untreated 30 h**	**DSM265**	**Pyrimethamine**	**MMV084**
Untreated 30 h				
DSM265	0.268			
Pyrimethamine	0.241	0.354		
MMV084	0.259	0.437	0.617	
M5717	0.268	N/A	0.354	0.437

© University of Pretoria

Supplementary Table 6b. Unpaired Student's t-test between parasite developmental stages of the 48 h incubation clinical candidate treatments.

Parasite development stage	Untreated 48h *vs.* UCT943	MMV048 *vs.* UCT943
Rings	0.0013	0.501
Trophozoites	0.095	0.825
Schizonts	0.471	0.590

© University of Pretoria