Title

Prioritised identification of structural classes of natural products from higher plants in the expedition of antimalarial drug discovery.

Permalink

https://escholarship.org/uc/item/4gc5j2d8

Journal

Natural Products and Bioprospecting, 13(1)

ISSN

2192-2195

Authors

Moyo, Phanankosi

Invernizzi, Luke

Mianda, Sephora

et al.

Publication Date

2023-10-12

DOI

10.1007/s13659-023-00402-2

ORIGINAL ARTICLE

Prioritised identification of structural classes of natural products from higher plants in the expedition of antimalarial drug discovery

Phanankosi Moyo[1] , Luke Invernizzi[1] , Sephora M. Mianda[1] , Wiehan Rudolph[1] , Andrew W. Andayi[2] , Mingxun Wang[3] , Neil R. Crouch[4,5] and Vinesh J. Maharaj[1*]

Abstract

The emergence and spread of drug-recalcitrant *Plasmodium falciparum* parasites threaten to reverse the gains made in the fight against malaria. Urgent measures need to be taken to curb this impending challenge. The higher plant-derived sesquiterpene, quinoline alkaloids, and naphthoquinone natural product classes of compounds have previously served as phenomenal chemical scaffolds from which integral antimalarial drugs were developed. Historical successes serve as an inspiration for the continued investigation of plant-derived natural products compounds in search of novel molecular templates from which new antimalarial drugs could be developed. The aim of this study was to identify potential chemical scaffolds for malaria drug discovery following analysis of historical data on phytochemicals screened in vitro against *P. falciparum*. To identify these novel scaffolds, we queried an in-house manually curated database of plant-derived natural product compounds and their in vitro biological data. Natural products were assigned to different structural classes using NPClassifier. To identify the most promising chemical scaffolds, we then correlated natural compound class with bioactivity and other data, namely (i) potency, (ii) resistance index, (iii) selectivity index and (iv) physicochemical properties. We used an unbiased scoring system to rank the different natural product classes based on the assessment of their bioactivity data. From this analysis we identified the top-ranked natural product pathway as the alkaloids. The top three ranked super classes identified were (i) pseudoalkaloids, (ii) naphthalenes and (iii) tyrosine alkaloids and the top five ranked classes (i) quassinoids (of super class triterpenoids), (ii) steroidal alkaloids (of super class pseudoalkaloids) (iii) cycloeudesmane sesquiterpenoids (of super class triterpenoids) (iv) isoquinoline alkaloids (of super class tyrosine alkaloids) and (v) naphthoquinones (of super class naphthalenes). Launched chemical space of these identified classes of compounds was, by and large, distinct from that of 'legacy' antimalarial drugs. Our study was able to identify chemical scaffolds with acceptable biological properties that are structurally different from current and previously used antimalarial drugs. These molecules have the potential to be developed into new antimalarial drugs.

Keywords Natural products, Compound classes, Phytochemicals, *Plasmodium falciparum*, Antiplasmodial drug resistance, Malaria, Drug development

*Correspondence:
Vinesh J. Maharaj
vinesh.maharaj@up.ac.za
Full list of author information is available at the end of the article

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 2 of 20

1 Introduction

Despite a plethora of resolute regional and global concerted efforts to curb malaria, this infectious disease continues to be a considerable health burden particularly to the populace in low-income countries of Africa and Asia [1]. In 2021 there were ca. 245 million reported clinical cases of malaria globally. The World Health Organisation (WHO) African Region carried the highest burden of this disease with ca. 234 million cases accounting for ca. 95% of global clinical malaria cases [1]. Significantly, four Sub-Saharan African countries accounted for 50% of the global clinical malaria burden in 2021 [1]. Despite reporting a low number of clinical malaria cases over the years, the South-East Asia region is notorious for being the epicenter of antimalarial drug-resistance development [2]. This comes as earlier chloroquine and sulphadoxine-pyrimethamine drug-resistant *P. falciparum* strains first emerged from this territory before spreading to the rest of the world [2]. Unfortunately, this phenomenon is recurring in South-East Asia as evidenced by the emergence of *P. falciparum* strains resistant to the current first-line WHO recommended antimalarial therapeutics, the artemisinin-based combination therapy (ACT) [2, 3]. There is genuine solicitude that the inadvertent spread of these ACT-resistant *P. falciparum* strains to Africa will result in a catastrophic outbreak of the disease, undermining momentous efforts to eliminate and eradicate malaria [2]. In this context, there is an urgent need to discover and develop new antimalarial drugs to circumvent

this imminent global health threat. One source worth exploring as a starting point in this regard is higher plant-derived natural products.

Different classes of both microbial and plant-derived natural products have historically proven to be an indispensable source of lead compounds for the development of the antimalarial arsenal [4]. Three derivatives of microbial-produced natural product compounds remain clinically useful for malaria control. These are clindamycin (**1**) (Fig. 1), tetracycline (**2**) (Fig. 2), and doxycycline (**3**) (Fig. 3) [5, 6]. Clindamycin is a lincosamide antibiotic [7]. It is derived from the *Streptomyces*-produced natural product compound lincomycin (**4**) which belongs to the aminoglycoside class of natural products (Fig. 1) [7]. Tetracycline and doxycycline are 1st and 2nd generation semi-synthetic derivatives, respectively, belonging to the tetracycline class of compounds [8]. Tetracyclines were originally isolated from filamentous bacteria of the genus *Streptomyces* [8]. Each one of these three drugs, **1**–**3**, are co-administered with either artesunate or quinine and used as a second line treatment regimen for recurrent *P. falciparum* malaria. Furthermore, a combination of clindamycin and quinine is strongly recommended by the WHO for treating uncomplicated *P. falciparum* malaria in the first trimester of pregnancy [6].

Critical in the battle against malaria has been the contribution of plant-derived natural products. These phytochemicals are quinine (**5**), lapachol (**6**), and artemisinin (**7**) which belong to the quinoline alkaloid,
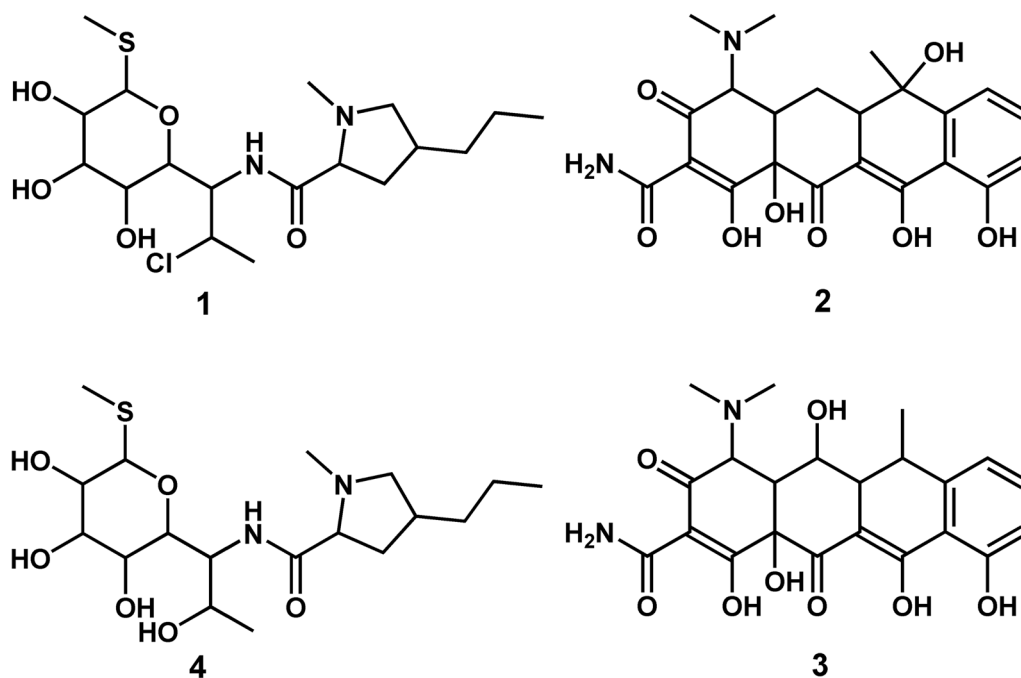
Fig. 1 2D illustration of bacterial (*Streptomyces*) derived antimalarial drugs. **1** and **4** are aminoglycosides, and **2** and **3** are tetracyclines

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 3 of 20
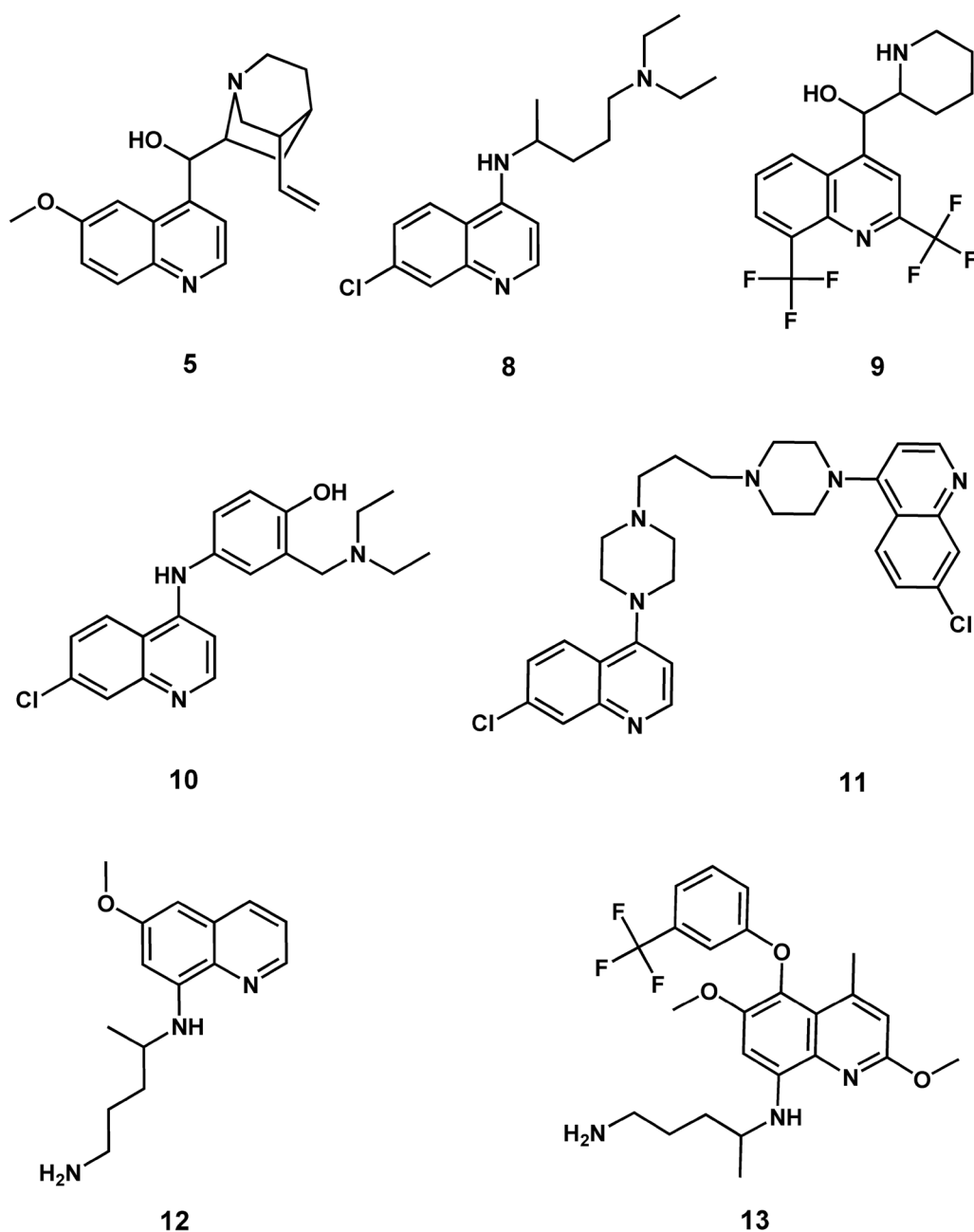
Fig. 2 2D illustration of plant-derived antimalarial drug compounds quinine (**5**) and its derivatives the 4-aminoquinolines **8**–**11** and 8-aminoquinolines **12** and **13**

naphthoquinone and sesquiterpene lactone classes of natural products, respectively (Figs. 2, 3) [4]. Quinine served as a template from which its derivatives, the aminoquinolines, including chloroquine (**8**), mefloquine (**9**), amodiaquine (**10**), piperaquine (**11**), and primaquine (**12**) were developed (Fig. 2) [4]. The most recent quinine derivative that has come to the fore is tafenoquine (**13**) (Fig. 2) [9]. For decades, chloroquine, an affordable and highly efficacious drug, was successfully used as the first line treatment drug for malaria. Mefloquine, amodiaquine and piperaquine are part of the ACT regimen [6]. Primaquine and tafenoquine are both used for preventing relapse of *P. vivax* and *P. ovale* with the former compound additionally used for blocking transmission of sexual *P. falciparum* parasites [6, 10].
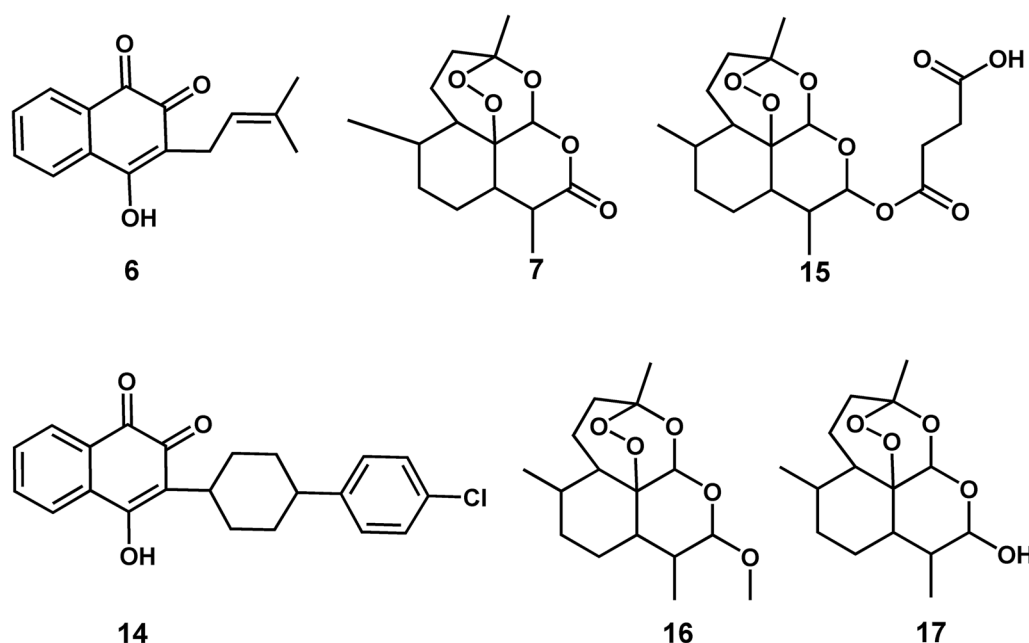
Fig. 3 2D illustration of plant-derived antimalarial drug compounds lapachol (**6**) and artemisinin (**7**) and their derivatives **14** and **15**–**17**, respectively

Lapachol served as a scaffold which inspired the development of atovaquone (14) [4] which currently, in combination with proguanil, is used as a casual prophylactic for malaria (Fig. 3) [6]. Three prolific semi-synthetic derivatives of artemisinin namely artesunate (15), artemether (16) and dihydroartemisinin (17) are the core components of the ACT regimen (Fig. 3) [6]. By and large, the contribution of natural products, particularly those which are plant-derived, has been profound in the fight against malaria. Against this background, it's only logical to continue investigating this immense source in search of valuable classes of natural product scaffolds to expedite development of the next generation of antimalarials.

In the past decade there has been an evolution in the natural product-based drug discovery field which allows for the targeted isolation of compounds, a paradigm shift from the classic 'blind' resource and time-consuming bioassay-guided approach. This transformation has been spurred on by the introduction of advanced hyphenated analytical techniques such as liquid chromatography coupled with mass spectrometry. Furthermore, there has been an advent of platforms such as GNPS [11] and SIRIUS [12] for high throughput spectral annotation and compound dereplication. Combined, these techniques and tools can facilitate the targeted isolation of structurally related phytochemicals, i.e., a natural "chemical series" of plant-derived analogues (analogous to a chemical series of compounds produced in a synthetic drug discovery program). This "series" of natural analogues

can be subjected to biological evaluation with the added advantage of acquiring its structure-activity relationship (SAR) data. The prompt establishment of the SAR for the targeted compound "series" guides precise development of their medicinal chemistry plans, to expedite their development as potential drugs in the discovery pipeline.

For decades, multitudes of phytochemicals of diverse structural classes have been screened in vitro for activity against asexual *P. falciparum* parasites. Given the immense structural variety and number of plant-derived natural products (ca. 133 881 plant-derived compounds reported in the dictionary of natural products by 2019 [13]), there is a need for a rational approach to prioritise classes of compounds for malaria drug discovery projects. This approach has been attempted in a previous study by Egieyeh and co-workers which primarily carried out a chemoinformatic-based analysis of a set of 1040 antiplasmodial natural compounds isolated from different sources including plants, microorganisms, and marine species [14]. Furthermore, this prior study briefly examined the relationship between chemical class and bioactivity. The outcome of the study was a list of specific natural compounds the authors recommended be prioritised for antimalarial drug discovery [14]. In accordance with this previous work, we here analyse an in-house data set of 2400 plant-derived natural compounds to ascertain which structural classes of natural compounds at the pathway, super class and class levels should reasonably be prioritised for antimalarial drug discovery. However, in

Moyo et al. Natural Products and Bioprospecting   (2023) 13:37

Page 5 of 20

variance to the study of Egieyeh et al., emphasis is herein placed on prioritising structural classes and not individual compounds. Furthermore, our work has delved more into the potency, additionally considering activity of the compounds against both drug-sensitive (D-S) and drug-resistant (D-R) intra-erythrocytic *P. falciparum* parasites, an aspect not previously undertaken by the earlier workers.

In this current study, we first evaluated the structural classes, namely pathways, super class, and class, based on four parameters, (i) potency, (ii) resistance index (RI), (iii) selectivity index (SI), and (iv) drug-likeness properties. We then used an unbiased scoring mechanism to rank structural classes based on their performance in two (potency and SI) of these four parameters. From the ranked list, the topmost pathway, super class, and class categories were subsequently identified. We envisage that findings from our study will be of value to the malaria drug discovery field and will potentially play a role in hastening the discovery of novel antimalarial chemotypes through target-based isolation.

2 Results, discussion and conclusion

2.1 Descriptive analysis: hit rate of natural product compounds in different pathways, super class and class categories

To identify the most quintessential structural class of plant-derived natural compounds for antimalaria drug discovery and development, we examined our previously manually curated database (Moyo et al. submitted). The database consists of a set of 2400 plant-derived natural compounds (representing 1.8% of known phytochemicals [13]) previously evaluated in vitro for their antiplasmodial activity against the intra-erythrocytic asexual *P. falciparum* parasites. These plant-derived compounds were identified and compiled from peer-reviewed literature sources from PubMed published over a course of 58 years between 1964 and 2021. Additional information available on the database included the pharmacological activity of the phytochemicals including their in vitro antiplasmodial potency against intra-erythrocytic asexual *P. falciparum* parasites, and cytotoxicity data.

To achieve the objective of our study we first assigned the 2400 natural product compounds to different structural classes, and to reduce subjectivity and ensure consistency used an online chemical ontology classification tool to classify all compounds/structures. Compounds in the database were assigned into three hierarchical classification categories of natural products, namely pathway (highest level), super class (middle level) and class (lowest level). The classification was carried out primarily using the online deep neural network-based structural classification ontology tool NPClassifier (Fig. 4) [15]. However,

NPClassifier in some cases apportioned compounds to multiple classes. Hence in such circumstances, another web-based application, ClassyFire [16], was then adopted as an "arbitrator", resolving this ambiguity, so assigning the compound into one specific class level. NPClassifier has been shown to outperform ClassyFire in accurately assigning secondary metabolites to their correct classes [15], hence its application for the primary classification.

Using NPClassifier, 97% (2349) of the compounds in the database were successfully classified into five different pathways, namely alkaloids, terpenes, polyketides, fatty acids, and shikimates and phenylpropanoids. Two hundred and twelve of the compounds multi-classified by NPClassifier were successfully reclassified at the class level using ClassyFire. Terpenes were the most represented pathway accounting for 40% of plant-derived compounds examined (Fig. 4). They were closely followed by shikimates and phenylpropanoids (30%), while fatty acids (2%) were the least represented. This frequency in the abundance of the pathways is consistent with their reported distribution in plants. This distribution partially matched that of antiplasmodial marine natural products reported elsewhere [17]. At the super class level triterpenoids and flavonoids were the two most prevalent whilst styrylpyrones and cyclic polyketides ranked as the least represented. Only phytochemicals with ≥ 10 compounds at a class level were considered for this study. At this low level of classification, amongst the three most represented classes were isoquinoline alkaloids (of super class tyrosine alkaloids), germacrane sesquiterpenoids (of super class triterpenoids), and corynanthe-type alkaloids (of super class tryptophan alkaloids). Aporphine alkaloids (of super class tyrosine alkaloids) and abietane diterpenoids (of super class diterpenoids) ranked amongst the least represented chemical entities at the class level (Fig. 4).

Having successfully classified the phytochemicals, the next step was to get a preliminary insight into the potency of natural products assigned to the different pathways, super class and class levels. This analysis was carried out by linking different compound classifications to their reported in vitro activity against both D-S and D-R intra-erythrocytic asexual *P. falciparum* parasites. The hit rates (HR) of the compounds per each pathway, super class and class were computed as previously reported by Moyo et al. (in review). The HR is the proportion of active compounds (defined as those with an $IC_{50} \leq 10$ μM) relative to the total number of compounds in that structural classification category. Considering the time and exorbitant financial resources invested in drug discovery projects, there is merit in working on structural classes with a high HR, the presumption being that this will ensure drugs are developed faster and more cost-effectively.

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 6 of 20
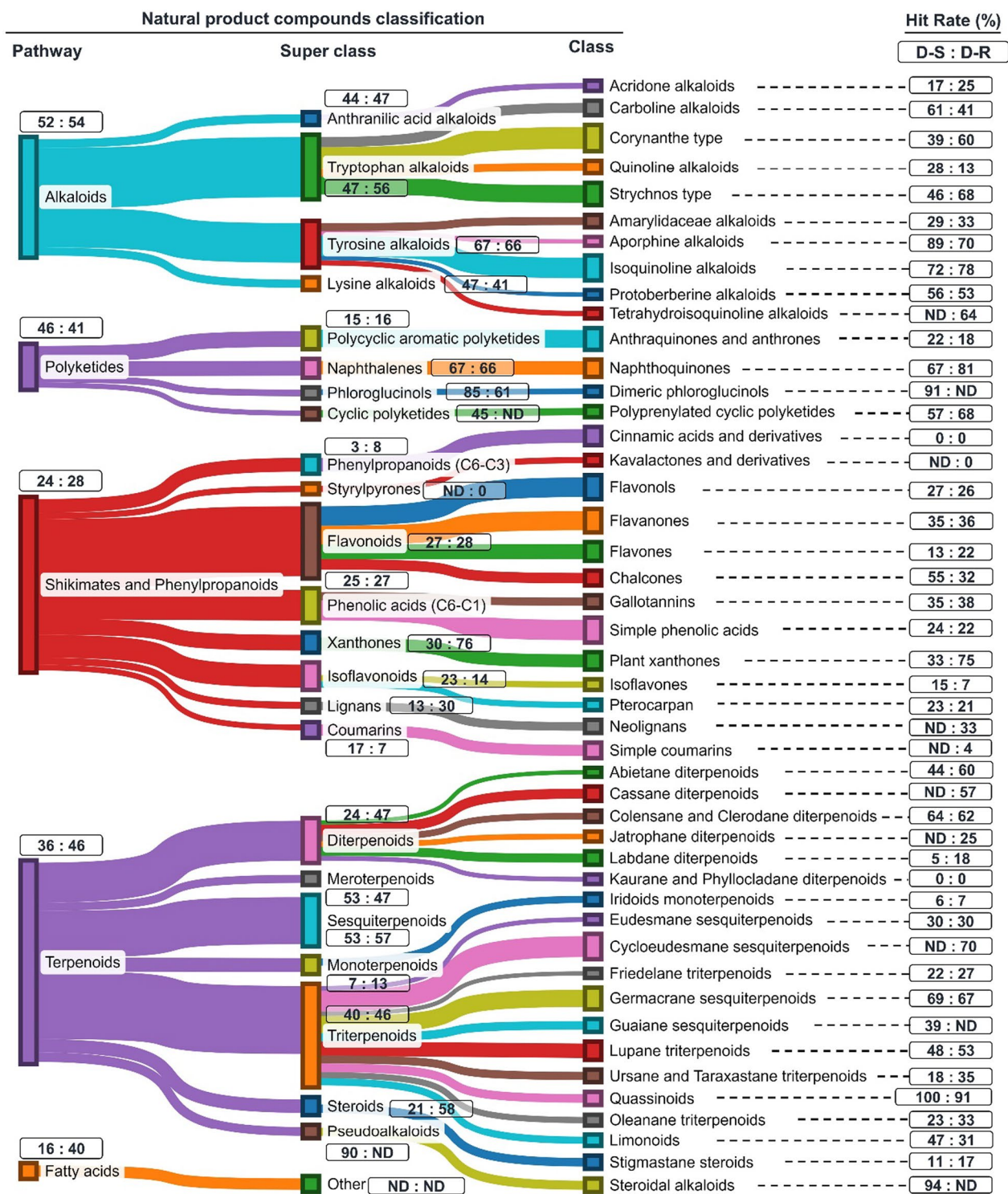
Fig. 4 Sankey chart illustration of pathways, super class and class of plant-derived natural compounds examined in this study. The size of nodes is proportional to the relative number of compounds across the different pathways and within the respective super classes and classes. Values in rectangles show hit rate (HR) against drug-sensitive (D-S) and drug-resistant (D-R) intra-erythrocytic asexual *P. falciparum* strains. *ND* not determined: This applies for compound pathways, super class and class of natural product compounds in which < 10 compounds per classification were evaluated for their antiplasmodial activity. The hit rate is the calculated % number of compounds with an $IC_{50} \leq 10$ µM for each structural classification. Classification of compounds into pathways, super class and class levels was primarily carried out using the online tool NPClassifier, and secondarily using ClassyFire, both automated online structural classification tools. Sankey chat was created using the online platform SankeyMATIC (https://sankeymatic.com/)

Moyo et al. Natural Products and Bioprospecting   (2023) 13:37

Page 7 of 20

Amongst the pathways, alkaloids had the highest proportion of active compounds and so achieved the topmost HR against both D-S (HR=52%, n (number of compounds)=281) and D-R (HR=54%, n=344) *P. falciparum* strains. Fatty acids had the lowest HR against D-S (HR=17%, n=24) while polyketides had the lowest HR against D-R (HR=28%, n=117) *P. falciparum* strains (Fig. 4). Interestingly, the fatty acids were more potent against D-R rather than to D-S asexual *P. falciparum* strains (difference of 23 between the HR).

At the super class level, pseudoalkaloids (HR=90%, n=20) and naphthalenes (HR=65%, n=34) had the highest ratio of active compounds against D-S and D-R *P. falciparum* strains, respectively. Phenylpropanoids (C6–C3) (HR=3.4%, n=29) and styrylpyrones (HR=0%, n=13) had the lowest ratio of active compounds against D-S and D-R *P. falciparum* parasites, respectively (Fig. 4). Xanthones, steroids, and diterpenoids had a remarkably high proportion of compounds active against D-R relative to D-S *P. falciparum* strains. The differences in the HR for activity against D-S and D-R parasites were 44, 36, and 22.7 for xanthones, steroids, and diterpenoids, respectively.

Quassinoids were revealed to be topmost at the class level, having the highest HR against both D-S (HR=100%, n=18) and D-R (HR=91%, n=54) intra-erythrocytic asexual *P. falciparum* parasites. Cinnamic acids and its derivatives (of super class phenylpropanoids) (HR=0%, n=29 and 23), and kaurane and phyllocladane diterpenoids (of super class diterpenoids) (HR=0%, n=10), had the lowest ratio of active compounds against both D-S and D-R asexual *P. falciparum* parasites. Kavalactones and derivatives (of super class styrylpyrones) also had a poor representation of active compounds (HR=0%, n=13) against D-R *P. falciparum* parasites. Consistent with its parent super class (xanthones), plant xanthones had a marked proportion of active compounds

against D-R compared to D-S parasites (a difference of 44 between the HR). Chalcones (of super class flavonoids), carboline alkaloids (of super class tryptophan alkaloids), and aporphine alkaloids (of super class tyrosine alkaloids) had a higher ratio of active compounds against D-S compared to D-R *P. falciparum* parasites (difference of ca. 19 between HR for all three classes of natural compounds) (Fig. 4).

In general, from this primary analysis we observe that chemical entities within the alkaloid pathway, and its associated super classes and classes have a high HR against both D-S and D-R *P. falciparum* parasites. In contrast, natural product compounds assigned to the shikimates and phenylpropanoids pathway and its related super class and class categories largely demonstrated low HR against the D-S and D-R *P. falciparum* parasite strains.

2.2 Antiplasmodial activity of natural products in different pathways, super classes and classes

Following the preliminary analysis of activity, we further examined the relevant pharmacological properties of compounds assigned to the different structural classification categories. We broadened our evaluation of potency by looking into the proportion of compounds per pathway, super class and class, classified as either highly active (HA) ($IC_{50} \leq 1$ μM), moderately active (MA) (10 μM $\geq IC_{50} > 1$ μM), or poorly active (PA) ($IC_{50} > 10$ μM). Furthermore, we investigated the resistance index (RI) and selectivity index (SI) for the reasons outlined above. A pathway, super class and class of priority for malaria drug discovery and development is taken to be one whose majority of compounds are classified as HA, while also having acceptable low RI (≤ 10) and high SI (≥ 10).

From this evaluation, most natural compounds in all five pathways were classified as PA (Table 1). The

Table 1 Antiplasmodial activity and cytotoxicity of higher plant-derived natural compound pathways

Pathway	Inhibition of parasite proliferation										Cytotoxicity	
	D-S			D-R								
	n	HA	MA	PA	n	HA	MA	PA	n	RI (≤10)	n	SI (≥10)
Alkaloids	281	22.1	29.9	50.2	344	23.3	31.1	45.6	153	98.0	232	50.0
Fatty acids	24	0.0	16.7	83.3	20	5.0	35.0	60.0			22	36.4
Polyketides	125	14.4	31.2	54.4	117	15.4	25.6	59.0	69	98.6	113	30.1
Shikimates[#]	389	3.6	20.1	76.3	507	4.3	23.7	72.0	176	96.6	209	25.8
Terpenoids	405	11.1	24.7	64.2	667	15.1	30.7	54.1	127	96.9	441	31.1

| 0 | 50 | 100 | **HM** |

n=number of compounds; HA, MA and PA values are expressed in % of total compounds evaluated; RI shows % number of compounds with RI ≤ 10; SI shows % number of compounds with SI ≥ 10. The heat map (HM) ranges from green (lowest value, 0%) to yellow (mid-range value, 50%) to red (highest value, 100%), visually illustrating the proportion of compounds classified as either HA, MA or PA. [#]Shikimates and Phenylpropanoids

alkaloids pathway had the greatest number of compounds classified as HA against both D-S (22%, n=281) and D-R (23%, n=344) asexual *P. falciparum* parasites (Table 1). Furthermore, many (50%) of the natural products in the alkaloid pathway (n=232) showed a good SI. Fatty acids and shikimates and phenylpropanoids had the fewest number of compounds classified as HA with most of the molecules (ca. 74%) in the latter pathway (n=209) demonstrating a poor SI. The emergence of alkaloids as the most prolific pathway is consistent with findings from Egieyeh et al. [14].

For the super class evaluation, most compounds in the naphthalenes were classified as HA against both D-S (43%, n=30) and D-R (43%, n=35) asexual *P. falciparum* parasites. Other super classes to have a considerable proportion (>25%) of their compounds classified as HA against D-S *P. falciparum* parasites were anthranilic acid alkaloids (n=18), lysine alkaloids (n=17), tyrosine alkaloids (n=83) and pseudoalkaloids (n=20). Against the D-R parasites, anthranilic acid alkaloids (n=23), tyrosine alkaloids (n=134) and tryptophan alkaloids (n=117) had a marked proportion (>25%) of their compounds

classified as HA. Additionally, the majority of compounds (>45%) from the anthranilic acid alkaloids (n=8), lysine alkaloids (n=14), tyrosine alkaloids (n=106) and pseudoalkaloids (n=20) showed acceptable SI (>10). Extensive historical investigations reflected in our database have shown that most compounds (>70%) of the super classes flavonoids, isoflavonoids, and phenolic acids (C6-C1) were PA against both D-S and D-R asexual *P. falciparum* parasites. Flavonoids, along with terpenoids, have previously been observed to exhibit generally low antiplasmodial activity [14]. Overall, most compounds showed an acceptable RI (< 10), see Table 2.

At the class level, the potency of quassinoids (of super class triterpenoids) was outstanding in having the majority of the compounds classified as HA against both D-S (77%, n=18) and D-R (63%, n=54) *P. falciparum* parasites. This level of potency was followed by naphthoquinones (of super class naphthalenes) whose majority of compounds emerged as HA against both D-S (44%, n=27) and D-R (57%, n=26) parasites. Isoquinoline alkaloids (of super class tyrosine alkaloids), steroidal alkaloids (of super class pseudoalkaloids), cycloeudesmane

Table 2 Antiplasmodial activity and cytotoxicity of higher plant-derived super classes of natural compounds

| | Super Class | Inhibition of parasite proliferation | | | | | | | | | Cytotoxicity | |
| | | D-S | | | D-R | | | | | | | |
		n	HA	MA	PA	n	HA	MA	PA	n	RI (≤10)	n	SI (≥10)
Alkaloids	Anthranilic acid alkaloids	18	33.3	11.1	55.6	23	26.1	21.7	52.2	10	100.0	8	75.0
	Lysine alkaloids	17	35.3	11.8	52.9	17	17.6	23.5	58.8	6	100.0	14	57.1
	Tryptophan alkaloids	128	14.8	32.0	53.1	117	25.6	29.9	44.4	60	98.3	75	46.7
	Tyrosine alkaloids	83	30.1	37.3	32.5	134	28.4	37.3	34.3	55	96.4	106	50.9
Polyketides	Cyclic polyketides	11	0.0	45.5	54.5							8	12.5
	Naphthalenes	30	43.3	23.3	33.3	35	42.9	22.9	34.3	25	96.0	29	41.4
	Phloroglucinols	14	0.0	85.7	14.3	13	23.1	38.5	38.5			19	31.6
	Polycyclic aromatic polyketides	33	3.0	12.1	84.8	37	0.0	16.2	83.8	27	100.0	32	21.9
Shikimates	Coumarins	12	0.0	16.7	100.0	29	0.0	6.9	93.1			5	40.0
	Flavonoids	147	5.4	21.8	72.8	190	4.2	23.7	72.1	71	97.2	89	30.3
	Isoflavonoids	48	0.0	22.9	77.1	72	0.0	13.9	86.1	33	97.0	17	0.0
	Lignans	15	0.0	13.3	86.7	43	11.6	18.6	69.8	7	100.0	14	42.9
	Styrylpyrones					13	0.0	0.0	100.0				
	Phenolic acids (C6-C1)	63	7.9	17.5	74.6	41	7.3	19.5	73.2	19	94.7	21	38.1
	Phenylpropanoids (C6-C3)	29	0.0	3.4	96.6	25	0.0	8.0	92.0	12	100.0	6	33.3
	Xanthones	33	3.0	27.3	69.7	51	9.8	64.7	25.5	11	90.9	29	20.7
Terpenoids	Diterpenoids	84	7.1	16.7	76.2	144	11.8	34.7	53.5	22	100.0	69	26.1
	Meroterpenoids	17	5.9	47.1	47.1	34	2.9	44.1	52.9			26	7.7
	Monoterpenoids	28	3.6	3.6	92.9	46	0.0	13.0	87.0			15	26.7
	Sesquiterpenoids	99	9.1	43.4	47.5	134	20.9	36.6	42.5	31	96.8	132	30.3
	Steroids	28	0.0	21.4	78.6	26	23.1	34.6	42.3			26	26.9
	Triterpenoids	143	19.6	20.3	60.1	281	17.8	28.1	54.1	64	95.3	170	35.9
	Pseudoalkaloids	20	30.0	60.0	10.0							20	75.0

| 0 | 50 | 100 | HM |

n = number of compounds; HA, MA and PA values are expressed as % of total compounds evaluated; RI shows % of compounds with RI ≤ 10; SI shows % of compounds with SI ≥ 10. The heat map (HM) ranges from green (lowest value, 0%) to yellow (mid-range value, 50%) to red (highest value, 100%), visually illustrating the proportion of compounds classified as either HA, MA or PA. #Shikimates and Phenylpropanoids

sesquiterpenoids (of super class triterpenoids), cassane diterpenoids (of super class diterpenoids), aporphine alkaloids (of super class tyrosine alkaloids), strychnos type alkaloids (of super class tryptophan alkaloids), and arylnaphthalene and aryltetralin lignans (both of super class triterpenoids) also fared well with > 30% of their compounds classified as HA either against D-S or D-R parasites. A considerable proportion of compounds (ranging from 44 to 84%) from all these mentioned classes also showed acceptable SI. The classes with the lowest proportion of plant-derived natural compounds classified as either HA or MA against *P. falciparum* parasites were kaurane and phyllocladane diterpenoids (of super class diterpenoids), kavalactones and derivatives (of super class styrylpyrones), simple coumarins (of super class coumarins), oleanane triterpenoids (of super class triterpenoids), labdane diterpenoids (of super class diterpenoids), isoflavones (of super class isiflavanoids), iridoids monoterpenoids (of super class monoterpenoids), flavones (of super class flavanoids), cinnamic acids and derivatives (of super class phenylpropanoids), anthraquinones and anthrones (of super class polycyclic and aromatic polyketides) and, interestingly, the quinoline alkaloids (of super class tryptophan alkaloids) (Table 3).

Overall, consistent with earlier observations, the alkaloid pathway and its associated sub-categories showed the most promising pharmacological properties including potency, RI and SI. Except for a few "pockets of brilliance", e.g., quassinoids, cycloeudesmane sesquiterpenoids, and steroidal alkaloids, by and large, most compounds assigned to the terpenoid and shikimates and phenylpropanoid pathways and related super classes and classes were classified as PA.

2.3 Drug-likeness assessment of compounds produced by different plant orders and families

Having assessed the pharmacological properties of plant-derived compounds assigned to different structural categories, we then investigated their drug-likeness for reasons outlined earlier (Moyo et al. submitted). Drug-likeness assesses the probability of a molecule to be bioavailable [18]. The current assessment was based on evaluating the in silico computed properties of compounds against criteria outlined by the Medicines for Malaria Venture (https://www.mmv.org/frontrunner-templates), Lipinski's Rule of 5 [19], Veber's rule [20] and Ghose filters [21]. To a large extent, most compounds in different pathways, super classes and classes complied with the set criteria. The most notable non-compliant structural classifications were the super classes fatty acids (rotatable bonds (RB) = 18) and steroids (molecular weight (MW) = 552, molecular refractory (MR) = 159), and the classes cycloeudesmane sesquiterpenoids (of

super class triterpenoids) (MW = 599), kaurane and phyllocladane diterpenoids (of super class diterpenoids) (MW = 768, hydrogen bond acceptors (HBA) = 16, hydrogen bond donors (HBD) = 10, MR = 181, and total polar surface area (TPSA) = 272), and gallotannins (of super class Phenolic acids (C6-C1)) (MW = 628, HBA = 17, MR = 10, and TPSA = 302). All structural classifications showed an acceptable PAINS score (Moyo et al. submitted). However, some classes demonstrated marked synthesis accessibility scores (≥ 6.5), most notable kaurane and phyllocladane diterpenoids (8.7), jatrophane diterpenoids (of super class diterpenoids) (7.2), cycloeudesmane sesquiterpenoids (of super class triterpenoids) (7.3), stigmastane steroids (of super class steroids) (6.7), limonoids (of super class cyclic polyketides) (6.6), polyprenylated cyclic polyketides (Hop meroterpenoids) (of super class triterpenoids) (6.9), and oleanane triterpenoids (of super class triterpenoids) (6.9), see Tables 4, 5, 6.

2.4 Overall ranking to identify compound structural classes for prioritisation in malaria drug discovery projects.

Having evaluated both pharmacological properties and drug-likeness we then employed an unbiased scoring system to rank different compound pathways, super classes and classes (Tables 7, 8, 9). For this process we decided against the inclusion of drug-likeness and RI as components as we noted that most compounds were compliant with these parameters.

Following implementation of this scoring system, the following results were realized: the top-ranked pathway was the alkaloids, top three ranked super classes were (i) pseudoalkaloids, (ii) naphthalenes and (iii) tyrosine alkaloids and top five ranked classes were (i) quassinoids, (ii) steroidal alkaloids (iii) cycloeudesmane sesquiterpenoids (iv) isoquinoline alkaloids and (v) naphthoquinones (Fig. 5).

Structurally, quassinoids are degraded triterpenes which are highly oxygenated [22, 23]. They are subdivided into 5 types based on the carbon skeleton structure, i.e., C-18, C-19, C-20, C-22, and C-25. In nature, they are mainly confined to the Simaroubaceae family being responsible for the bitter taste of species in this family [22, 23]. They are complex molecules which are difficult to synthesise. Nonetheless, innovative synthesis approaches for this class of compounds have been reported in several studies [24–28] and concisely reviewed by others [29, 30]. Biologically, quassinoids have demonstrated exceptional potency across different disease areas including cancer, HIV and, as noted in this study, malaria [22, 23]. Some compounds have shown single digit nanomolar activity against both D-S and D-R *P. falciparum* parasites, not sharing cross resistance with

Table 3 Antiplasmodial activity and cytotoxicity of higher plant-derived natural compound classes

	Class	Inhibition of parasite proliferation										Cytotoxicity	
		D-S				D-R							
		n	HA	MA	PA	n	HA	MA	PA	n	RI (≤10)	n	SI (≥10)
Alkaloids	Acridone alkaloids	12	0.0	16.7	83.3	12	0.0	25.0	75.0	5	100.0		
	Amaryllidaceae alkaloids	17	0.0	29.4	70.6	21	4.8	28.6	66.7	11	100.0	18	22.2
	Aporphine alkaloids	9	22.2	66.7	11.1	23	39.1	30.4	30.4			17	47.1
	Carboline alkaloids	21	19.0	42.9	38.1	31	22.6	19.4	58.1	12	100.0	16	68.8
	Isoquinoline alkaloids	43	34.9	37.2	27.9	59	30.5	47.5	22.0	27	96.3	50	68.0
	Protoberberine alkaloids	9	22.2	33.3	44.4	19	26.3	26.3	47.4	9	100.0	10	40.0
	Tetrahydroisoquinoline alkaloids					11	27.3	36.4	36.4			9	44.4
	Quinoline alkaloids	18	0.0	27.8	72.2	23	0.0	13.0	87.0	11	100.0	12	16.7
	Corynanthe type	46	10.9	28.3	60.9	45	17.8	42.2	40.0	23	95.7	30	40.0
	Strychnos type	37	16.2	29.7	54.1	19	47.4	21.1	31.6	16	100.0	13	46.2
Polyketides	Anthraquinones and anthrones	36	5.6	16.7	77.8	38	0.0	18.4	81.6	28	100.0	35	25.7
	Naphthoquinones	27	44.4	22.2	33.3	26	57.7	23.1	19.2	22	95.5	25	44.0
	Dimeric phloroglucinols	11	0.0	90.9	9.1							13	38.5
	Polyprenylated cyclic polyketides	14	0.0	57.1	42.9	19	0.0	68.4	31.6			15	6.7
Shikimates#	Chalcones	22	9.1	45.5	45.5	17	11.8	23.5	64.7	14	100.0	14	28.6
	Cinnamic acids and derivatives	29	0.0	0.0	100.0	23	0.0	0.0	100.0	8	100.0	5	20.0
	Kavalactones and derivatives					13	0.0	0.0	100.0				
	Flavanones	40	5.0	30.0	65.0	50	6.0	30.0	64.0	20	100.0	20	35.0
	Flavones	30	0.0	13.3	86.7	36	5.6	16.7	88.9	8	100.0	24	25.0
	Flavonols	41	12.2	14.6	73.2	58	0.0	25.9	74.1	23	95.7	21	23.8
	Gallotannins	17	17.6	17.6	64.7	13	15.4	23.1	61.5	11	100.0	10	50.0
	Isoflavones	13	0.0	15.4	84.6	29	0.0	6.9	93.1	12	100.0	7	0.0
	Neolignans					18	5.6	27.8	66.7				
	Plant xanthones	27	3.7	29.6	66.7	48	10.4	64.6	25.0	9	88.9	29	20.7
	Simple coumarins					23	0.0	4.3	95.7				
	Simple phenolic acids	42	4.8	19.0	76.2	27	3.7	18.5	77.8	7	85.7	10	30.0
	Pterocarpan	13	0.0	23.1	76.9	19	0.0	21.1	78.9	6	100.0	5	0.0
	Arylnaphthalene and aryltetralin lignans					11	36.4	18.2	45.5			7	71.4
Terpenoids	Cassane diterpenoids					21	33.3	23.8	42.9			11	54.5
	Colensane and Clerodane diterpenoids	14	21.4	42.9	35.7	26	11.5	50.0	38.5	10	100.0	15	0.0
	Cycloeudesmane sesquiterpenoids					44	54.5	15.9	29.5			13	84.6
	Eudesmane sesquiterpenoids	10	0.0	30.0	70.0	10	0.0	30.0	70.0	6	100.0	12	25.0
	Friedelane triterpenoids	9	11.1	11.1	77.8	11	18.2	9.1	72.7				
	Germacrane sesquiterpenoids	36	13.9	55.6	30.6	42	4.8	61.9	33.3	10	90.0	60	20.0
	Guaiane sesquiterpenoids	18	5.6	33.3	61.1							14	28.6
	Iridoids monoterpenoids	15	0.0	6.7	93.3	28	0.0	7.1	92.9				
	Jatrophane diterpenoids					12	8.3	16.7	75.0				
	Kaurane and Phyllocladane diterpenoids	10	0.0	0.0	100.0	10	0.0	0.0	100.0				
	Labdane diterpenoids	19	0.0	5.3	94.7	17	0.0	17.6	82.4				
	Limonoids	15	13.3	33.3	53.3	62	1.6	29.0	69.4	8	100.0	23	39.1
	Lupane triterpenoids	29	24.1	24.1	51.7	47	17.0	36.2	46.8	14	92.9	29	41.4
	Oleanane triterpenoids	13	0.0	23.1	100.0	42	2.4	31.0	66.7	9	100.0	31	19.4
	Quassinoids	18	77.8	22.2	0.0	54	63.0	27.8	9.3	12	83.3	35	62.9
	Steroidal alkaloids	19	31.6	63.2	5.3							20	75.0
	Abietane diterpenoids	9	0.0	44.4	55.6	15	6.7	53.3	40.0			12	25.0
	Stigmastane steroids	18	0.0	11.1	88.9	12	0.0	16.7	83.3			8	37.5
	Ursane and Taraxastane triterpenoids	17	0.0	17.6	82.4	17	0.0	35.3	76.5	5	100.0	17	29.4

0   50   100   **HM**

n = number of compounds; HA, MA and PA values are expressed in % of total compounds evaluated; RI shows % number of compounds with RI ≤ 10; SI shows % number of compounds with SI ≥ 10. The heat map (HM) ranges from green (lowest value, 0%) to yellow (mid-range value, 50%) to red (highest value, 100%), visually illustrating the proportion of compounds classified as either HA, MA or PA. #Shikimates and Phenylpropanoids

Table 4 Calculated mean physicochemical descriptors for natural products in different pathways

Pathway	Physicochemical Descriptors							PAINS	SA
	MW	RB	HBA	HBD	MR	TPSA	cLogP		
Alkaloids	378.6	2.8	4.2	1.3	113.7	57.2	3.1	0.2	4.3
Fatty acids	449.8	18.3	5.6	3.1	128.6	96.3	4.5	0.0	5.0
Polyketides	386.3	4.4	5.8	2.3	108.2	93.5	3.5	0.4	4.0
Shikimates[#]	385.1	4.3	7.2	3.3	101.9	113.6	2.4	0.2	4.0
Terpenoids	443.5	4.6	5.9	2.1	121.3	92.6	3.5	0.0	5.9

MW molecular weight, *RB* rotatable bonds, *HBA* hydrogen bond acceptors, *HBD* hydrogen bond donors, *MR* molar refractivity, *TPSA* total polar surface area, *cLogP* consensus LogP, *PAINS* pan-assay interference compounds, *SA* synthesis accessibility

Grey shaded figures are those which don't meet the set criteria; [#]Shikimates and Phenylpropanoids

Table 5 Calculated mean physicochemical descriptors for natural products in different super classes

	Super Class	Physicochemical Descriptors							PAINS	SA
		MW	RB	HBA	HBD	MR	TPSA	cLogP		
Alkaloids	Anthranilic acid alkaloids	279.2	1.4	3.3	0.7	82.4	51.3	2.5	0.0	2.5
	Lysine alkaloids	391.6	6.0	4.8	1.8	118.5	61.0	3.4	0.0	5.2
	Tryptophan alkaloids	407.4	2.3	3.5	1.6	124.9	56.5	2.9	0.3	5.0
	Tyrosine alkaloids	396.3	2.8	5.3	1.1	116.5	60.5	3.4	0.1	4.1
Polyketides	Cyclic polyketides	353.1	3.5	5.5	1.7	96.7	93.6	3.1	0.0	4.4
	Naphthalenes	418.0	3.2	5.6	2.4	122.8	82.2	3.9	0.5	4.2
	Phloroglucinols	422.6	7.4	6.0	3.3	121.2	106.9	4.5	0.0	4.3
	Polycyclic aromatic polyketides	359.0	2.5	6.1	2.5	95.2	104.4	2.5	1.0	3.2
Shikimates	Coumarins	319.8	3.6	6.3	1.8	84.4	90.8	2.1	0.1	3.7
	Flavonoids	422.7	4.0	8.1	4.1	111.1	132.0	2.3	0.3	4.3
	Isoflavonoids	327.0	2.5	5.3	1.6	89.6	74.2	2.9	0.0	3.6
	Lignans	373.2	6.0	6.0	1.5	100.8	76.2	3.1	0.0	4.1
	Pseudoalkaloids	456.0	4.8	3.7	1.7	135.7	53.4	4.7	0.0	6.1
	Phenolic acids (C6-C1)	359.8	4.5	8.3	4.2	88.0	140.4	1.2	0.4	3.6
	Phenylpropanoids (C6-C3)	345.3	6.4	6.9	3.6	88.7	115.3	1.3	0.4	3.7
	Xanthones	384.5	3.5	5.8	2.9	110.9	98.1	4.0	0.2	4.0
Terpenoids	Diterpenoids	430.6	5.5	5.9	1.7	118.1	89.5	3.7	0.1	5.6
	Meroterpenoids	483.7	7.6	5.1	2.2	142.4	88.9	5.4	0.5	5.8
	Monoterpenoids	318.8	4.8	6.6	3.1	78.7	103.3	0.7	0.0	4.8
	Sesquiterpenoids	374.4	3.7	5.9	1.3	98.6	87.8	2.3	0.0	5.5
	Steroids	552.3	8.0	5.5	2.3	159.4	81.9	5.7	0.0	7.2
	Triterpenoids	508.7	4.0	6.1	2.5	140.7	99.6	4.3	0.0	6.5
	Styrylpyrones	296.1	3.5	5.1	1.0	77.1	69.3	1.8	0.0	4.2

MW molecular weight, *RB* rotatable bonds, *HBA* hydrogen bond acceptors, *HBD* hydrogen bond donors, *MR* molar refractivity, *TPSA* total polar surface area, *cLogP* consensus LogP, *PAINS* pan-assay interference compounds, *SA* synthesis accessibility

Grey shaded figures are those which don't meet the set criteria; [#]Shikimates and Phenylpropanoids

chloroquine [31, 32]. It is therefore surprising that limited attention in the malaria field has been accorded this class of compounds. Almost half of the compounds analysed in this study were from two studies carried out by O'Neil and co-workers in 1986 and 1987 [31, 32]. One possible deterrent is the earlier discussed complexity in the synthesis of compounds in this quassinoids class. A work around for this could be to identify molecular features associated with potency and to use that to synthesise libraries of simple molecules whose structures are inspired by this class of compounds. This approach has gained impetus in the field with millions of virtual natural product-like compounds having been generated for drug discovery [33–35].

The steroidal alkaloid class of phytochemicals is limited in its distribution, amongst higher plants, being mainly

Table 6 Calculated mean physicochemical descriptors for natural products in different classes

Class		MW	RB	HBA	HBD	MR	TPSA	cLogP	PAINS	SA
		Physicochemical Descriptors							**PAINS**	**SA**
Alkaloids	Acridone alkaloids	287.2	1.6	3.8	0.8	82.8	56.2	2.5	0.0	2.5
	Amaryllidaceae alkaloids	315.8	1.4	5.5	1.2	85.6	64.3	1.3	0.0	4.5
	Aporphine alkaloids	322.8	1.8	4.3	0.7	94.7	51.8	3.0	0.0	3.1
	Carboline alkaloids	333.9	1.7	2.8	1.6	102.8	52.5	3.1	0.4	3.4
	Isoquinoline alkaloids	429.3	3.2	5.3	1.3	129.2	61.3	4.2	0.2	4.3
	Protoberberine alkaloids	351.8	2.8	4.8	0.9	99.3	57.2	2.6	0.3	3.3
	Tetrahydroisoquinoline alkaloids	564.5	3.5	7.5	0.9	168.2	70.3	4.7	0.0	6.2
	Quinoline alkaloids	273.2	3.4	3.6	0.3	80.1	44.6	3.1	0.0	2.9
	Corynanthe type	413.2	2.6	3.5	1.8	126.3	59.6	2.9	0.5	5.2
	Strychnos type	475.5	1.6	3.7	0.9	148.3	46.7	3.1	0.0	6.1
Polyketides	Anthraquinones and anthrones	363.6	2.5	6.2	2.6	96.5	106.4	2.5	1.0	3.3
	Naphthoquinones	460.1	3.1	5.9	2.5	136.5	85.5	4.3	0.7	4.9
	Dimeric phloroglucinols	462.4	7.8	7.1	3.6	130.3	125.3	4.4	0.0	4.5
	Polyprenylated cyclic polyketides (Hop me	574.8	8.6	6.1	2.6	168.2	108.1	6.3	0.8	6.9
Shikimates#	Chalcones	382.4	5.6	4.9	1.7	109.6	73.9	4.0	0.0	4.0
	Cinnamic acids and derivatives	477.1	8.7	11.1	6.1	115.5	181.7	0.2	0.6	4.9
	Kavalactones and derivatives	296.1	3.5	5.1	1.0	77.1	69.3	1.8	0.0	4.2
	Flavanones	405.2	3.3	6.4	3.3	110.3	105.1	3.3	0.0	4.1
	Flavones	431.1	3.3	8.2	4.0	115.0	136.7	2.6	0.2	4.2
	Flavonols	480.7	5.2	11.2	6.0	118.9	184.7	1.0	0.4	5.0
	Gallotannins	628.3	3.0	17.4	9.9	144.8	302.3	0.0	0.9	5.5
	Isoflavones	312.9	2.8	5.2	1.9	87.3	82.2	2.8	0.0	3.2
	Neolignans	368.9	7.5	5.5	2.0	102.5	75.8	3.3	0.0	4.0
	Plant xanthones	391.9	3.7	5.9	3.0	113.3	100.3	4.1	0.3	4.0
	Simple coumarins	312.4	3.8	5.8	1.4	84.1	81.8	2.4	0.0	3.6
	Simple phenolic acids	279.2	5.0	5.6	2.3	70.8	90.5	1.5	0.3	3.0
	Pterocarpan	322.2	1.8	5.2	1.2	86.5	63.2	3.0	0.1	3.9
	Arylnaphthalene and aryltetralin lignans	391.3	4.5	6.5	0.5	104.3	73.1	3.3	0.0	4.1
Terpenoids	Cassane diterpenoids	403.9	3.5	6.2	1.2	106.0	90.9	2.8	0.0	5.5
	Colensane and Clerodane diterpenoids	436.0	7.9	6.1	1.0	119.1	87.7	3.9	0.0	6.0
	Cycloeudesmane sesquiterpenoids	599.9	5.9	10.2	2.5	151.2	154.1	2.5	0.0	7.3
	Eudesmane sesquiterpenoids	249.1	0.7	3.0	1.3	70.0	50.2	2.3	0.0	4.3
	Friedelane triterpenoids	442.6	0.6	2.0	0.4	135.6	34.2	6.8	0.0	5.3
	Germacrane sesquiterpenoids	349.0	3.9	6.1	1.0	90.9	87.4	2.0	0.0	5.6
	Guaiane sesquiterpenoids	272.8	0.8	4.2	1.2	72.1	61.9	1.7	0.0	5.1
	Iridoids monoterpenoids	386.6	5.6	9.8	4.9	87.2	153.3	-1.1	0.0	5.6
	Jatrophane diterpenoids	663.9	11.0	12.2	1.2	169.9	172.8	3.4	0.0	7.2
	Kaurane and Phyllocladane diterpenoids	768.1	9.2	16.7	10.0	181.5	272.9	-0.8	0.0	8.7
	Labdane diterpenoids	361.0	5.4	3.7	1.4	105.0	60.2	4.3	0.1	5.0
	Limonoids	516.6	4.9	8.4	0.8	132.8	117.5	3.2	0.0	6.6
	Lupane triterpenoids	515.4	4.6	3.4	1.7	155.1	59.0	7.0	0.0	6.2
	Oleanane triterpenoids	580.6	3.4	7.4	4.7	159.1	130.0	4.1	0.0	6.9
	Quassinoids	463.1	2.9	9.4	3.7	111.5	148.8	0.4	0.0	6.3
	Steroidal alkaloids	470.9	5.0	3.8	1.6	140.5	50.4	5.0	0.0	6.3
	Abietane diterpenoids	300.2	1.1	2.3	0.6	90.1	39.2	4.2	0.9	4.1
	Stigmastane steroids	457.1	6.5	2.4	1.7	141.2	42.8	6.6	0.0	6.7
	Ursane and Taraxastane triterpenoids	471.6	1.6	3.4	2.3	140.3	63.7	5.7	0.0	6.4

Table 6 (continued)

MW molecular weight, *RB* rotatable bonds, *HBA* hydrogen bond acceptors, *HBD* hydrogen bond donors, *MR* molar refractivity, *TPSA* total polar surface area, *cLogP* consensus LogP, *PAINS* pan-assay interference compounds, *SA* synthesis accessibility

Grey shaded figures are those which don't meet the set criteria; #Shikimates and Phenylpropanoids

Table 7 Ranking compound pathways for antimalarial drug discovery from higher plants

Pathway	Points							Ranking
	D-S		D-R					
	HR	HA	HR	HA	SI	Total	Normalised	
Alkaloids	1	1	1	1	1	5	1	1
Fatty acids	5	5	4	4	2	20	4	4
Polyketides	2	2	3	2	4	13	2,6	2
Shikimates#	4	4	5	5	5	23	4,6	5
Terpenoids	3	3	2	3	3	14	2,8	3

1 [red] — 5 [green]   RSHM

HR Hit rate; ranking score heat map (RSHM) ranges from red (best ranking, lowest points) to green (lowest ranking, most points). #Shikimates and Phenylpropanoids

Table 8 Ranking compound super classes for antimalarial drug discovery from higher plants

| Super Class		Points							Ranking
		D-S		D-R					
		HR	HA	HR	HA	SI	Total	Normalised	
Alkaloids	Anthranilic acid alkaloids	10	3	8	3	1	25	5	4
	Lysine alkaloids	7	2	12	9	3	33	6,6	6
	Tryptophan alkaloids	8	7	7	4	5	31	6,2	5
	Tyrosine alkaloids	3	4	3	2	4	16	3,2	3
Polyketides	Cyclic polyketides	9	16			20	45	15	17
	Naphthalenes	4	1	2	1	7	15	3	2
	Phloroglucinols	2	16	4	5	12	39	7,8	7
	Polycyclic aromatic polyketides	19	14	16	16	18	83	16,6	19
Shikimates	Coumarins	18	16	20	16	8	78	15,6	18
	Flavonoids	13	12	14	14	13	66	13,2	15
	Isoflavonoids	16	16	17	16	22	87	17,4	22
	Lignans	20	16	13	11	6	66	13,2	15
	Pseudoalkaloids	1	5			1	7	2,3	1
	Phenolic acids (C6-C1)	14	9	15	13	9	60	12	12
	Phenylpropanoids (C6-C3)	22	16	19	16	11	84	16,8	20
	Xanthones	12	14	1	12	19	58	11,6	10
Terpenoids	Diterpenoids	15	10	10	10	17	62	12,4	14
	Meroterpenoids	5	11	9	15	21	61	12,2	13
	Monoterpenoids	21	13	18	16	16	84	16,8	20
	Sesquiterpenoids	6	8	6	7	14	41	8,2	8
	Steroids	17	16	5	5	15	58	11,6	10
	Triterpenoids	11	6	11	8	10	46	9,2	9
	Styrylpyrones			21	16		37	18,5	23

1 [red] — 23 [green]   RSHM

HR hit rate; ranking score heat map (RSHM) ranges from red (best ranking, lowest points) to green (lowest ranking, most points). #Shikimates and Phenylpropanoids

Table 9 Ranking compound classes for antimalarial drug discovery from higher plants

	Class	Points							Ranking
		D-S		D-R					
		HR	HA	HR	HA	SI	Total	Normalised	
Alkaloids	Acridone alkaloids	32	23	30	29		114	28,5	38
	Amaryllidaceae alkaloids	23	23	25	24	29	124	24,8	28
	Aporphine alkaloids	4	6	6	5	9	30	6	6
	Carboline alkaloids	9	9	18	11	4	51	10,2	11
	Isoquinoline alkaloids	5	3	3	8	5	24	4,8	4
	Protoberberine alkaloids	11	6	17	10	14	58	11,6	12
	Tetrahydroisoquinoline alkaloids			10	9	11	30	10	10
	Quinoline alkaloids	24	23	38	29	34	148	29,6	39
	Corynanthe type	17	16	12	13	14	72	14,4	16
	Strychnos type	15	11	7	4	10	47	9,4	9
Polyketides	Anthraquinones and anthrones	29	18	35	29	24	135	27	35
	Naphthoquinones	7	2	2	2	12	25	5	5
	Dimeric phloroglucinols	3	23			17	43	14,33	15
	Polyprenylated cyclic polyketides	10	23	7	29	35	104	20,8	25
Shikimates#	Chalcones	12	17	22	16	22	89	17,8	19
	Cinnamic acids and derivatives	38	23	42	29	31	163	32,6	44
	Kavalactones and derivatives			42	29		71	35,5	47
	Flavanones	20	20	20	21	19	100	20	24
	Flavones	34	23	32	22	25	136	27,2	36
	Flavonols	25	14	29	29	28	125	25	29
	Gallotannins	19	10	19	15	8	71	14,2	14
	Isoflavones	33	23	40	29	36	161	32,2	43
	Neolignans			23	22		45	22,5	27
	Plant xanthones	21	22	4	18	30	95	19	20
	Simple coumarins			41	29		70	35	46
	Simple phenolic acids	26	21	32	26	20	125	25	29
	Pterocarpan	27	23	34	29	36	149	29,8	40
	Arylnaphthalene and aryltetralin lignans			15	6	3	24	8	7
Terpenoids	Cassane diterpenoids			14	7	7	28	9,3	8
	Colensane and Clerodane diterpenoids	8	8	11	17	36	80	16	17
	Cycloeudesmane sesquiterpenoids			5	3	1	9	3	3
	Eudesmane sesquiterpenoids	22	23	27	29	25	126	25,2	33
	Friedelane triterpenoids	29	15	28	12		84	21	26
	Germacrane sesquiterpenoids	6	12	9	24	31	82	16,4	18
	Guaiane sesquiterpenoids	18	18			22	58	19,3	22
	Iridoids monoterpenoids	36	23	39	29		127	31,75	42
	Jatrophane diterpenoids			31	19		50	25	29
	Kaurane and Phyllocladane diterpenoids	38	23	42	29		132	33	45
	Labdane diterpenoids	37	23	36	29		125	31,25	41
	Limonoids	14	13	26	28	16	97	19,4	23
	Lupane triterpenoids	13	5	16	14	13	61	12,2	13
	Oleanane triterpenoids	27	23	23	27	33	133	26,6	34
	Quassinoids	1	1	1	1	6	10	2	1
	Steroidal alkaloids	2	4			2	8	2,7	2
	Abietane diterpenoids	16	23	12	20	25	96	19,2	21
	Stigmastane steroids	35	23	37	29	18	142	28,4	37
	Ursane and Taraxastane triterpenoids	31	23	21	29	21	125	25	29

1	47	RSHM

HR hit rate; ranking score heat map (RSHM) ranges from red (best ranking, lowest points) to green (lowest ranking, most points). #Shikimates and Phenylpropanoids

confined to a few plant families that include the Solanaceae, Buxaceae, Apocynaceae, and Liliaceae sensu lato [36, 37]. Structurally, steroidal alkaloids consist of a basic steroidal skeleton containing either one or two nitrogen atoms either in the rings or on an attached functional group [37]. From a medicinal chemistry perspective, this

Quassinoids (Terpenoids, Triterpenoids)

18: IC$_{50}$ = 0.001 µM   **19: IC$_{50}$ = 0.002 µM**
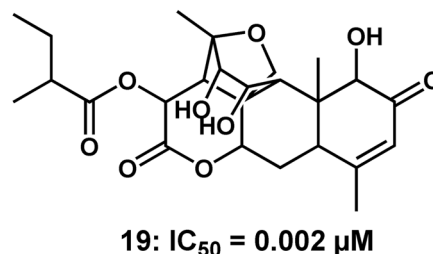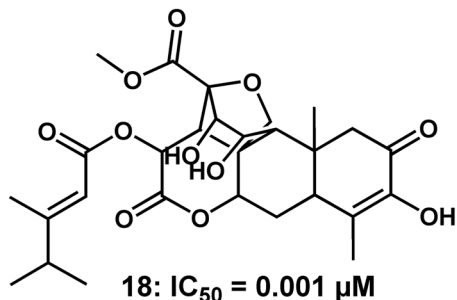
Steroidal alkaloids (Terpenoids, Pseudoalkaloids)

20: IC$_{50}$ = 0.180 µM   **21: IC$_{50}$ = 0.200 µM**
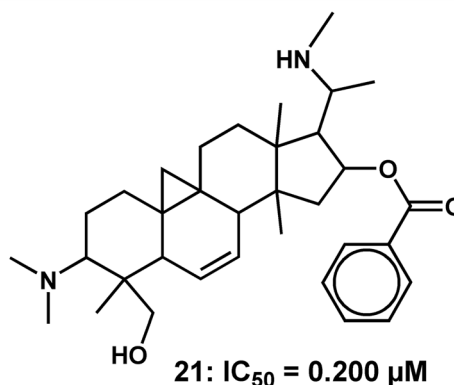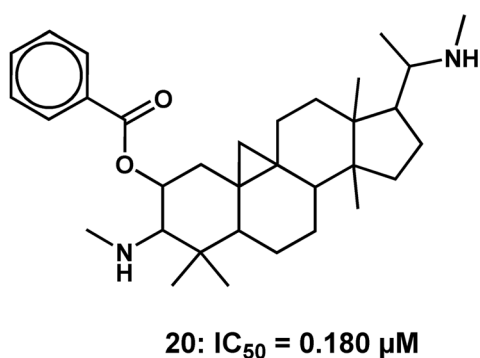
Cycloeudesmane sesquiterpenoids (Terpenoids, Sesquiterpenoids)

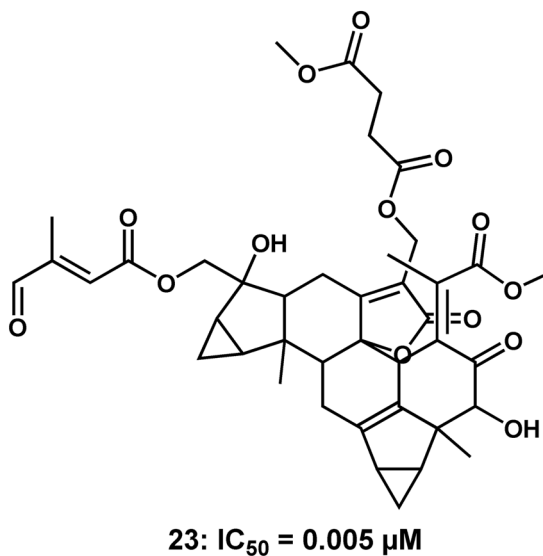**22: IC$_{50}$ = 0.001 µM**   **23: IC$_{50}$ = 0.005 µM**
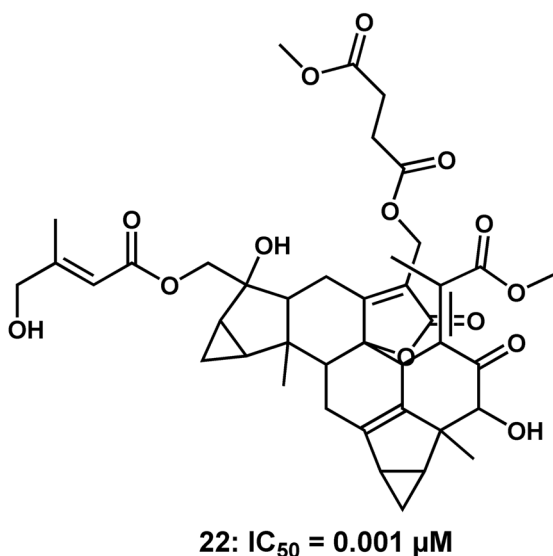
Fig. 5 2D illustration of representative HA natural product compound structures from the top five ranked compound classes. Indicated IC$_{50}$'s are measured against intra-erythrocytic asexual *P. falciparum* parasites

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 16 of 20

Isoquinoline alkaloids (Alkaloids, Tyrosine alkaloids)
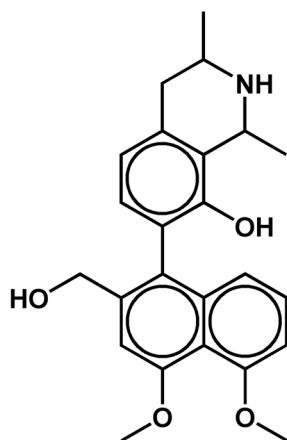
24: IC$_{50}$ = 0.005 µM
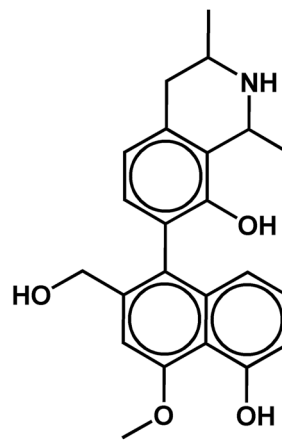
25: IC$_{50}$ = 0.008 µM

Naphthoquinones (Polyketides, Naphthalenes)
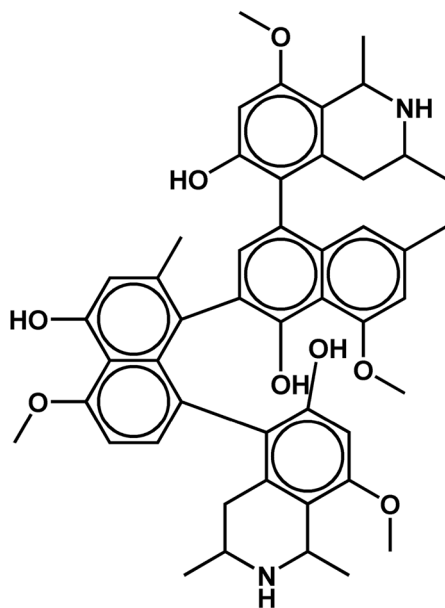
26: IC$_{50}$ = 0.004 µM
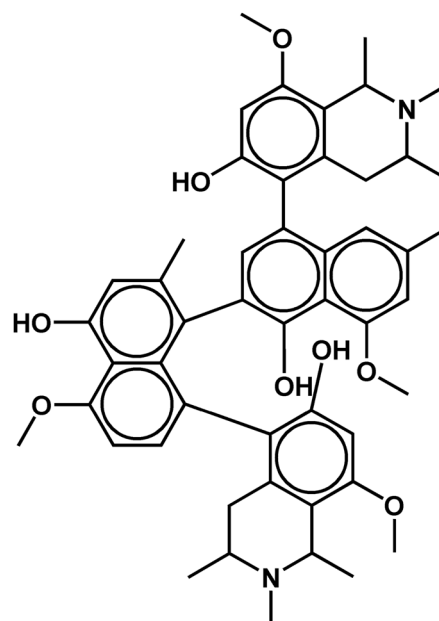
27: IC$_{50}$ = 0.006 µM

Fig. 5 continued

complex class of compounds presents a synthesis challenge mainly brought about by the chiral nature of the molecules. This challenge was already noted by the high synthesis accessibility score highlighted in this study. Nonetheless, despite their complexity, several steroidal alkaloids have been successfully synthesised [38–40]. This has allowed for structural modifications leading

to the development of analogues of parent compounds. Synthesis of analogues has resulted in the establishment of comprehensive SAR studies which have informed the progress of initial leads to clinical candidates [38–40].

The unique structural combination of the steroidal and alkaloid moieties endows this class of natural compounds with a unique set of physicochemical and biological

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 17 of 20
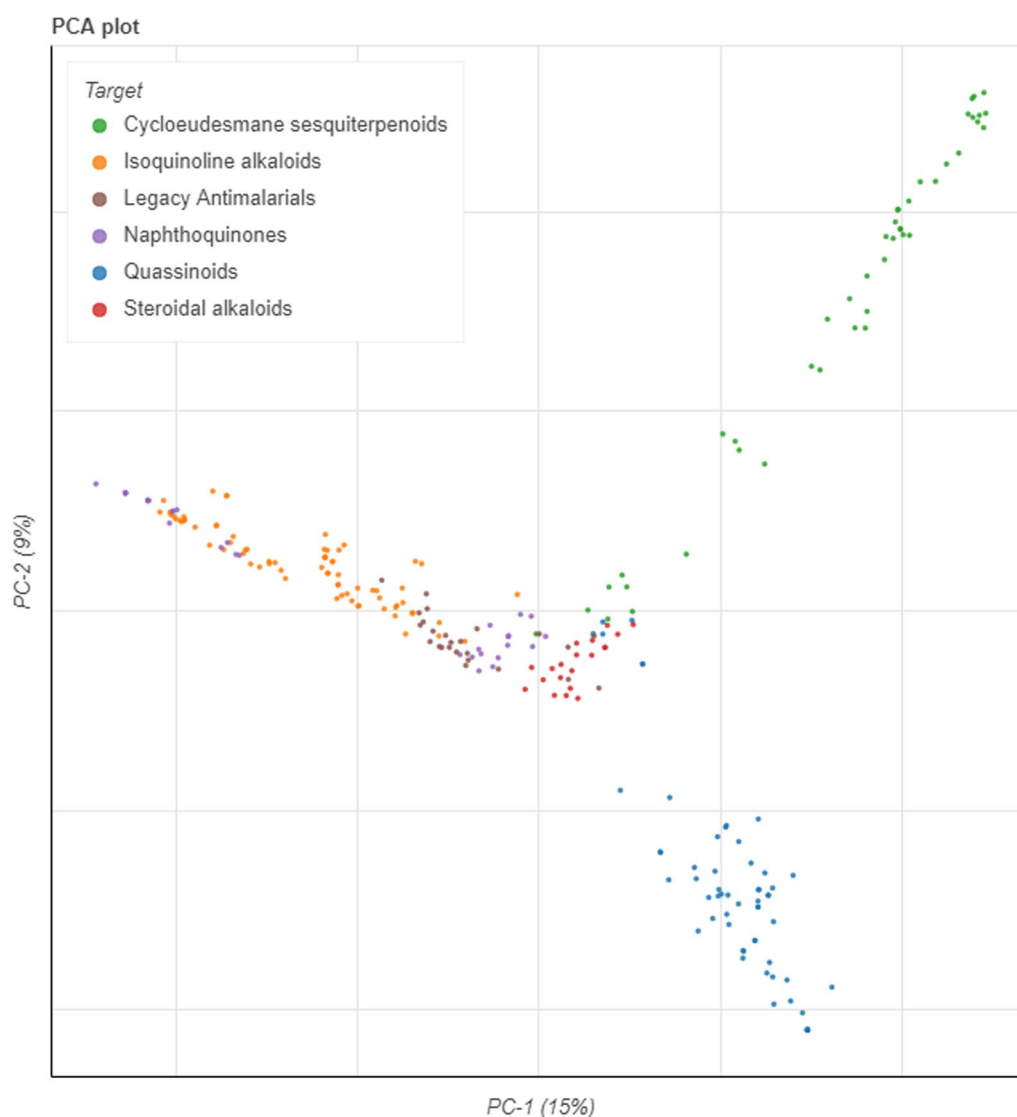
Fig. 6 Launched drug chemical space (generated using the principal component analysis (PCA) dimensional reduction method) of the 'legacy' antimalarials and natural product compounds from the top five ranked compound classes. The online Python library for chemical space visualization, ChemPlot, was used to launch the chemical space of the natural compounds and 'legacy' antimalarials

properties. Consequently, they have a wide spectrum of biological activity including anti-inflammatory, anti-microbial and anticholinergic [36]. They have received marked interest in the cancer field for their exceptional activity epitomised by the clinically approved prostate cancer therapeutic abiraterone acetate [37]. Another noteworthy antineoplastic agent of the steroidal alkaloids class is cyclopamine, a natural compound isolated from the plant *Veratrum californicum* (Melanthiaceae) [37]. While cyclopamine is still in clinical studies, two of its analogues, vismodegib and sonidegib have already been approved by the Food and Drug Agency for cancer treatment [37]. Overall, steroidal alkaloids, despite their complex structure, are evidently amenable to structural modifications which have successfully led to the development of drugs.

Despite their exceptional biological activity, steroidal alkaloids have received comparatively subdued attention for their antiprotozoal activity, particularly against *P. falciparum* parasites. Foremost in their examination has been a study carried out by Szabo et al. [41] in which 25 alkaloids, including steroidal alkaloids (as per classification of NPClassifier) were isolated from *Buxus sempervirens* (Buxaceae). Their study provided evidence of the potential of this class of compounds, with five compounds showing exceptional activity with IC_{50}
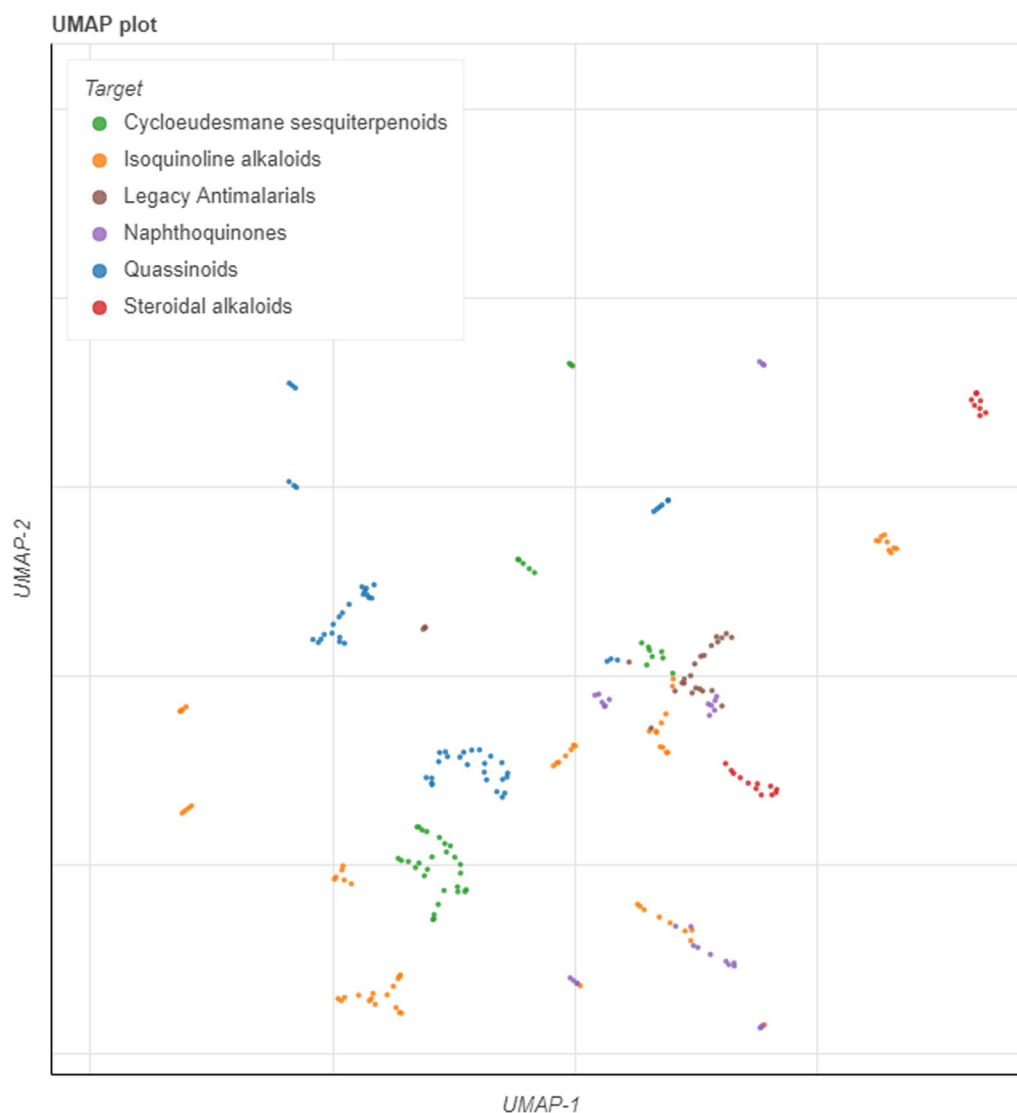
Fig. 7 Launched drug chemical space (generated using the uniform manifold approximation and projection (UMAP) dimensional reduction method) of the 'legacy' antimalarials and natural product compounds from the top five ranked classes. The online Python library for chemical space visualization, ChemPlot, was used to launch the chemical space of the natural compounds and 'legacy' antimalarials

values < 1 μM [41]. Moreover, the authors highlight that SAR studies are in progress and are anticipated to give insights on structural modifications that could improve the potency of the compounds and reduce their toxicity, hence improving their selectivity further [41].

Cycloeudesmane sesquiterpenoids belong to the same super class category as artemisinin, namely the sesquiterpenoids. Despite their prolific activity, this class of compounds is yet to receive marked attention in the antimalarial drug discovery field. All compounds assessed in this study were from a single study by Zhou et al. [42]. To our best knowledge, there are few antimalarial studies that include this class of compounds, making it an

interesting exploratory prospect. In contrast, the isoquinoline and naphthoquinone classes have received an abundance of antiplasmodial research attention [43–46]. These alkaloids have been isolated principally from species of African Ancistrocladaceae and Dioncophyllaceae, providing hope that they will be the source of the first plant-derived antimalarial drug to be discovered on the African continent.

Following identification of the top-ranked classes of natural product compounds, we have sought to structurally compare them with the legacy set of antimalarials, i.e., the old and currently used malaria therapeutics [47]. To assess the structural similarity between the

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 19 of 20

legacy antimalarial series and compounds assigned to the top five ranked classes, we launched the chemical space of the two groups of compounds (Figs. 6, 7). The chemical space of compounds is a multi-dimensional environment in which structurally similar molecules are grouped closely together [48] while structurally distinct compounds occupy a different space. The co-ordinates of the compounds within the 2D chemical space were plotted using four-dimensional reduction tools. In this study we chose two different dimensional reduction tools, namely the principal component analysis (PCA) and uniform manifold approximation and projection (UMAP) [48]. Encouragingly, compounds from the top five ranked classes largely clustered separately from the legacy antimalarials indicating structural differences between the two groups of compounds (Figs. 6, 7). Overlap in the chemical space was noted for the legacy antimalarials, and isoquinoline alkaloids in the PCA plot. This overlap could be explained by some structural similarities between the closely related isoquinoline alkaloids and quinoline derivatives which largely dominate the antimalarial legacy compounds. Nevertheless, most of the natural product classes occupied a different chemical space to legacy antimalarials, indicative of structural differences between these compound groups.

In conclusion, higher plant-derived natural products present a myriad of biologically exciting, structurally complex, and diverse molecules with great potential for development into antimalarial drugs. They offer a great avenue to find novel chemotypes that could circumvent the antimalarial drug resistance threat. Targeted isolation of compounds offers a great opportunity to isolate structurally related compounds of antiplasmodial interest that can be subjected to a fully-fledged drug discovery program, to expedite product delivery in a malaria drug discovery pipeline. Our study has provided insights to support this notion by cogently identifying structural classes of phytochemicals with acceptable in vitro antiplasmodial potency, RI, SI and drug-likeness properties. These compound classes (and super classes) should rationally be prioritized for target-based natural product drug discovery in the development of novel antimalarial chemotypes.

3 Methods
Detailed methods for this article are provided in our companion article Moyo et al. (submitted).

Author contributions
Conceptualisation of study, PM and VJM.; Collecting data sets, PM, LI, SMM, WR, and AWA.; Formal analysis, PM, AWA, WR, MW and NRC.; Writing—original draft, PM.; Writing—review and editing, PM, VJM, LI, SMM, AWA, MW, and NRC. Supervision and funds acquisition VJM.; All authors read and approved the final manuscript.

Availability of data and materials
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Competing interests
MW is a co-founder of Ometa Labs LLC.

Author details
[1]Department of Chemistry, Faculty of Natural and Agricultural Sciences, Biodiscovery Center, University of Pretoria, Private Bag X 20, Hatfield, Pretoria 0028, South Africa. [2]Department of Physical and Biological Sciences, Murang'a University of Technology Murang'a, Murang'a, Kenya. [3]Computer Science and Engineering, University of California Riverside, 900 University Ave, Riverside, CA 92521, USA. [4]Biodiversity Research and Monitoring Directorate, South African National Biodiversity Institute, Berea Road, P.O. Box 52099, Durban 4007, South Africa. [5]School of Chemistry and Physics, University of KwaZulu-Natal, Durban 4041, South Africa.

References
1. WHO, World malaria report 2022. 2022: World Health Organization.
2. Menard D, Dondorp A. Antimalarial drug resistance: a threat to malaria elimination. Cold Spring Harb Perspect Med. 2017;7(7): a025619.
3. Takala-Harrison S, et al. Independent emergence of artemisinin resistance mutations among *Plasmodium falciparum* in Southeast Asia. J Infect Dis. 2015;211(5):670–9.
4. Wells TNC. Natural products as starting points for future anti-malarial therapies: going back to our roots? Malar J. 2011;10(1):S3.
5. Gaillard T, et al. Antibiotics in malaria therapy: which antibiotics except tetracyclines and macrolides may be used against malaria? Malar J. 2016;15(1):556.
6. Organization WH. Guidelines for the treatment of malaria. 2015: World Health Organization.
7. Spížek J, Řezanka T. Lincomycin, clindamycin and their applications. Appl Microbiol Biotechnol. 2004;64(4):455–64.
8. Nelson ML, Levy SB. The history of the tetracyclines. Ann N Y Acad Sci. 2011;1241(1):17–32.
9. Peters W. The evolution of tafenoquine—antimalarial for a new millennium? J R Soc Med. 1999;92(7):345–52.
10. Watson JA, Nekkab N, White M. Tafenoquine for the prevention of *Plasmodium vivax* malaria relapse. Lancet Microbe. 2021;2(5):e175–6.
11. Wang M, et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat Biotechnol. 2016;34(8):828–37.
12. Dührkop K, et al. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods. 2019;16(4):299–302.
13. Chassagne F, et al. The landscape of natural product diversity and their pharmacological relevance from a focus on the Dictionary of Natural Products®. Phytochem Rev. 2019;18:601–22.

Moyo *et al. Natural Products and Bioprospecting*   (2023) 13:37

Page 20 of 20

14. Egieyeh SA, et al. Prioritization of anti-malarial hits from nature: chemo-informatic profiling of natural products with in vitro antiplasmodial activities and currently registered anti-malarial drugs. Malar J. 2016;15:1–23.

15. Kim HW, et al. NPClassifier: a deep neural network-based structural classification tool for natural products. J Nat Prod. 2021;84(11):2795–807.

16. Djoumbou Feunang Y, et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. J Cheminformat. 2016;8:1–20.

17. Hai Y, et al. Trends of antimalarial marine natural products: progresses, challenges and opportunities. Nat Prod Rep. 2022;39(5):969–90.

18. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7(1):42717.

19. Lipinski CA, et al. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. 1997;23(1):3–25.

20. Veber DF, et al. Molecular properties that influence the oral bioavailability of drug candidates. J Med Chem. 2002;45(12):2615–23.

21. Ghose AK, Viswanadhan VN, Wendoloski JJ. A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. J Comb Chem. 1999;1(1):55–68.

22. Chakraborty D, Pal A. Quassinoids: chemistry and novel detection techniques. Berlin: Springer; 2013. p. 3345–66.

23. Duan Z-K, et al. Quassinoids: phytochemistry and antitumor prospect. Phytochemistry. 2021;187: 112769.

24. Kraus GA, Taschner MJ. Model studies for the synthesis of quassinoids. 1. Construction of the BCE ring system. J Org Chem. 1980;45(6):1175–6.

25. Herscovici J, et al. Stereocontrolled routes to functionalized [1,8-bc] naphthopyran. A study on the total synthesis of quassinoids and tetrahydronaphthalene antibiotics. J Org Chem. 1993;58(15):3928–37.

26. Gross RS, Grieco PA, Collins JL. Synthetic studies on quassinoids: total synthesis of (±)-chaparrinone. J Am Chem Soc. 1990;112(25):9436–7.

27. Ziegler FE, et al. Practical routes to two functionalized decalones for the synthesis of quassinoids. J Org Chem. 1986;51(24):4573–9.

28. Thomas WP, Pronin SV. A concise enantioselective approach to quassinoids. J Am Chem Soc. 2021;144(1):118–22.

29. Pazur EJ, Wipf P. Recent syntheses and biological profiling of quassinoids. Org Biomol Chem. 2022;20(19):3870–89.

30. Kawada K, Kim M, Watt DS. Synthesis of quassinoids. A review. Org Prep Proc Int. 1989;21(5):521–618.

31. O'Neill MJ, et al. Plants as sources of antimalarial drugs: in vitro antimalarial activities of some quassinoids. Antimicrob Agents Chemother. 1986;30(1):101–4.

32. O'Neill MJ, et al. Plants as sources of antimalarial drugs, Part 4: Activity of *Brucea javanica* fruits against chloroquine-resistant *Plasmodium falciparum* in vitro and against *Plasmodium berghei* in vivo. J Nat Prod. 1987;50(1):41–8.

33. Tay DW, et al. 67 million natural product-like compound database generated via molecular language processing. Sci Data. 2023;10(1):296.

34. Li Y, et al. Designing natural product-like virtual libraries using deep molecule generative models. Macromolecules. 2018;3:5.

35. Yu MJ. Natural product-like virtual libraries: recursive atom-based enumeration. J Chem Inf Model. 2011;51(3):541–57.

36. Abd Karim HA, Ismail NH, Osman CP. Steroidal alkaloids from the apocynaceae family: their isolation and biological activity. Nat Prod Commun. 2022;17(11):1934578X221141265.

37. Xiang M-L, et al. Chemistry and bioactivities of natural steroidal alkaloids. Nat Prod Bioprospect. 2022;12(1):23.

38. Sharpe RJ, Johnson JS. A global and local desymmetrization approach to the synthesis of steroidal alkaloids: stereocontrolled total synthesis of paspaline. J Am Chem Soc. 2015;137(15):4968–71.

39. Tokuyama T, Daly J, Witkop B. Structure of batrachotoxin, a steroidal alkaloid from the Colombian arrow poison frog, *Phyllobates aurotaenia*, and partial synthesis of batrachotoxin and its analogs and homologs. J Am Chem Soc. 1969;91(14):3931–8.

40. Zha X, et al. Efficient synthesis of solasodine, *O*-acetylsolasodine, and soladulcidine as anticancer steroidal alkaloids. Chem Biodivers. 2007;4(1):25–31.

41. Szabó LU, et al. Antiprotozoal nor-triterpene alkaloids from Buxus sempervirens L. Antibiotics. 2021;10(6):696.

42. Zhou B, et al. Nanomolar antimalarial agents against chloroquine-resistant *Plasmodium falciparum* from medicinal plants and their structure–activity relationships. J Nat Prod. 2017;80(1):96–107.

43. Lombe BK, Feineis D, Bringmann G. Dimeric naphthylisoquinoline alkaloids: polyketide-derived axially chiral bioactive quateraryls. Nat Prod Rep. 2019;36(11):1513–45.

44. Bringmann G, et al. Highly selective antiplasmodial naphthylisoquinoline alkaloids from *Ancistrocladus tectorius*. Phytochemistry. 2013;91:220–8.

45. Bringmann G, et al. Habropetaline A, an antimalarial naphthylisoquinoline alkaloid from *Triphyophyllum peltatum*. Phytochemistry. 2003;62(3):345–9.

46. Moyo P, et al. Naphthylisoquinoline alkaloids, validated as hit multistage antiplasmodial natural products. Int J Parasitol Drugs Drug Resist. 2020;13:51–8.

47. Charman SA, et al. An in vitro toolbox to accelerate anti-malarial drug discovery and development. Malar J. 2020;19(1):1–27.

48. Cihan Sorkun M, et al. ChemPlot, a Python library for chemical space visualization. Chem Methods. 2022;2(7): e202200005.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.