

学校编码: 10384

分类号 _____ 密级 _____

学号: 21620131152479

UDC _____

厦 门 大 学

硕 士 学 位 论 文

约氏疟原虫转录组的基因结构和亚种间的差异性剪接

Gene structure analysis and differentially splicing between
Plasmodium yoelii subspecies

刘建文

指导教师姓名: 李剑 助理教授

刘升发 教授

专 业 名 称: 动物学

论文提交日期: 2016 年 4 月

论文答辩时间: 2016 年 5 月

学位授予日期:

答辩委员会主席: _____

评 阅 人: _____

2016 年 5 月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为(苏新专教授、刘升发教授)课题(组)的研究成果,获得“约氏疟原虫甲氟喹抗性的基因定位及功能研究”(主持人为李剑助理教授,项目编号:81201324)和“我国疟原虫抗药性的分子流行病学及全基因组关联性研究”(主持人为潘卫庆教授,项目编号:81220108019)课题(组)经费或实验室的资助,在(苏新专教授、刘升发教授)实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

1. 经厦门大学保密委员会审查核定的保密学位论文，于 年 月 日解密，解密后适用上述授权。

2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月 日

目录

摘要.....	1
Abstract	2
第一章 前言	3
1.1 疟原虫及其生活史	3
1.2 全球疟疾现状	6
1.3 我国疟疾现状	8
1.4 疟疾的传播媒介—按蚊	10
1.5 疟疾防治	11
1.6 药物抗性	12
1.6.1 媒介对杀虫剂抗性.....	12
1.6.2 抗疟药物抗性.....	13
1.7 疟疾疫苗	19
1.8 本研究目的和意义	20
第二章 约氏疟原虫 NSM 虫株的生长表型和药物敏感性	21
2.1 引言	21
2.2 材料与方法	21
2.2.1 虫株的复苏与感染.....	21
2.2.2 蚊媒感染步骤.....	23
2.2.3 按蚊卵囊及唾液腺解剖步骤.....	23
2.2.4 约氏疟原虫 DNA 提取.....	25
2.2.5 微卫星标记扩增程序.....	26

2.3 结果	27
2.3.1 约氏疟原虫 NSM 虫株的生长表型	27
2.3.1.1 NSM 虫株原虫率变化	27
2.3.1.2 NSM 虫株配子体率变化	28
2.3.1.3 NSM 的卵囊和子孢子发育情况	29
2.3.1.4 约氏疟原虫 NSM 各生长时期的微卫星基因分型	30
2.3.2 约氏疟原虫 NSM 虫株的 MQ 抗药性	31
2.4 讨论	34
第三章 约氏疟原虫基因转录组的结构分析和亚种间差异性剪接	35
3.1 引言	35
3.2 实验方法	36
3.2.1 疟原虫与小鼠感染	36
3.2.2 RNA 样本提取	36
3.2.3 验证 cDNA 的 PCR 扩增反应程序	38
3.2.4 链特异性文库的构建及测序	39
3.2.5 读取序列的过滤和序列的基因组定位	39
3.2.6 序列统计和基因表达水平的评估	40
3.2.7 新转录基因和选择性转录的检测	40
3.2.8 单核苷酸多态性 SNPs 和插入或缺失的检测	41
3.2.9 SNPs 和选择性剪接的实验验证	41
3.3 结果	41
3.3.1 链特异性 RNA-seq	41
3.3.2 读取序列的染色体定位	42
3.3.3 基因表达水平的评估	44
3.3.4 基因选择性剪接和约氏疟原虫亚种间的差异性剪接内含子	49
3.3.5 在 YM 和 NSM 虫株间单核苷酸多态性的检测	54
3.3.6 反义链转录基因	54

3.3.7 选择性/差异性剪接内含子和 SNPs 的验证.....	55
3.4 讨论.....	58
3.5 结论.....	60
第四章 总结.....	61
参考文献.....	62
附录一. 与约氏疟原虫 YM 预测基因模式不匹配的新转录基因和新型基因 结构.....	69
攻读学位期间发表的论文.....	93
致谢.....	94

Contents

Abstract in Chinese	1
Abstract in English	2
Chapter 1: Introduction	3
1.1 <i>Plasmodium</i> life cycle	3
1.2 The current situation of world malaria	6
1.3 Malaria in China	8
1.4 Malaria vectors	10
1.5 Malaria control and treatment	11
1.6 Drug resistant	12
1.6.1 Insecticide-resistant in <i>Anopheles</i>	12
1.6.2 Drug-resistant in <i>Plasmodium</i>	13
1.7 Vaccine	19
1.8 Study aim.....	20
Chapter 2: Growth phenotype in blood-stage, oocyst and sporozoite development, and mefloquine response of <i>Plasmodium yoelii</i> spp NSM	21
2.1 Introduction	21
2.2 Materials and methods.....	21
2.2.1 Parasite maintain and infection	21
2.2.2 Parasite maintain and infection	23
2.2.3 Dissection of mosquitoes	23
2.2.4 DNA extraction	25
2.2.5 PCR cycling program.....	26

2.3 Results	27
2.3.1 Growth phenotype of <i>P. yoelii</i> spp NSM.....	27
2.3.1.1 Parasitemia of the NSM parasite	27
2.3.1.2 Gametocytemia.....	28
2.3.1.3 Oocyst and sporozoite development in mosquitoes	29
2.3.1.4 Microsatellites genotyping on different development stages of NSM	30
2.3.2 Drug response	31
2.4 Discussion.....	34
Chapter 3: Gene structure analysis and differentially splicing between <i>P. yoelii</i>	
subspecies.....	35
3.1 Introduction	35
3.2 Materials and methods.....	36
3.2.1 Parasites and Infection of mice	36
3.2.2 RNA sample collection and preparation	36
3.2.3 Detection of cDNA and PCR cycling programs	38
3.2.4 Strand-specific library preparation and sequencing.....	39
3.2.5 Sequence reads cleanup and mapping to genome.....	39
3.2.6 Read counts and estimation of gene expression level.....	40
3.2.7 Detection of new transcripts and events of alternative transcripts	40
3.2.8 Detection of SNPs and indels	41
3.2.9 Experimental confirmation of selected SNPs and AS events	41
3.3 Results	41
3.3.1 Directional RNA-seq	41
3.3.2 Mapping sequence reads to chromosomes.....	42
3.3.3 Estimates of gene expression levels.....	44
3.3.4 Alternatively and differentially spliced introns between two <i>P. yoelii</i> subspecies ..	49
3.3.5 Detection of SNPs between YM and NSM parasites.....	54

3.3.6 Genes with antisense transcripts	54
3.3.7 Confirmation of selected SNPs and alternatively/differentially spliced introns.....	55
3.4 Discussion	58
3.5 Conclusion.....	60
Chapter 4: Conclusion	61
References	62
Additional file 1. New genes or genes with structures that do not match the predicted gene models in the <i>P. yoelii</i> YM genome sequence	69
Publications	93
Acknowledgements.....	94

摘要

啮齿类动物约氏疟原虫 (*Plasmodium yoelii*) 是研究宿主-寄生虫相互关系和疟原虫病理分子基础的一种重要的动物模型。尽管约氏疟原虫 YM 虫株的基本基因组信息已被测定, 在最近也有许多啮齿类动物疟原虫的 RNA-seq 序列信息已见报道, 但不同疟原虫品系或者亚种之间可能存在基因编码区及 mRNA 转录结构上的差异。构建多种疟原虫品系和亚种的 cDNA 文库将有助于我们对疟原虫遗传模式和转录组结构的认识。

本研究, 我们用约氏疟原虫 NSM 虫株有或无 MQ 药物压力的红内期原虫样本构建了两个 cDNA 文库并对其进行深度测序, 获得转录组序列再与最近公布的 YM 虫株 cDNA 序列相比对, 以检测编码区单核苷酸多态性 (SNPs)、内含子-外显子结构差异、约氏疟原虫亚种间的差异性剪接以及在 MQ 药物压力下的基因表达变化。

对约氏疟原虫 NSM 虫株的转录组进行深度测序, 我们获得了大约有 56,000,000 个 100bp 双端测序序列, 其测序覆盖度约为 225 倍。在与 YM 虫株转录组序列的比较检测中发现了 5' 和 3' 非翻译区的内含子、内含子-外显子边界改变、选择性剪接转录、正义链转录和反义链转录的重叠以及新预测的基因转录。在序列数据比对中发现了差异性剪接内含子, 比如一个亚种中的剪接内含子在另一个亚种转录组中可能并不是剪接内含子。同时, 将 NSM 的转录组序列与 YM 的转录组序列的比对也鉴定出在两个虫株间的大约 8400 个单核苷酸多态性 (SNPs)。

在约氏疟原虫两个亚种间非翻译区内含子和差异性剪接内含子的发现提示, 这些内含子对基因表达的调控和对疟原虫进化的影响有着重要的作用。

关键词: 疟疾; RNA 深度测序; 基因表达; 单核苷酸多态性; 选择性剪接

Abstract

The rodent malaria parasite *Plasmodium yoelii* is an important animal model for studying host-parasite interaction and molecular basis of malaria pathogenesis. Although a draft genome of *P. yoelii yoelii* YM is available, and RNA sequencing (RNA-seq) data for several rodent malaria species (RMP) were reported recently, variations in coding regions and structure of mRNA transcript are likely present between different parasite strains or subspecies. Sequencing of cDNA libraries from additional parasite strains/subspecies will help improve the gene models and genome annotation.

In this study, two directional cDNA libraries from mixed blood stages of a subspecies of *P. yoelii* (*P. y. nigeriensis* NSM) with or without mefloquine (MQ) treatment were sequenced, and the sequence reads were compared to the genome and cDNA sequences of *P. y. yoelii* YM in public databases to investigate single nucleotide polymorphisms (SNPs) in coding regions, variations in intron–exon structure and differential splicing between *P. yoelii* subspecies, and variations in gene expression under MQ pressure.

Approximately 56 million of 100 bp paired-end reads were obtained, providing an average of ~225-fold coverage for the coding regions. Comparison of the sequence reads to the YM genome revealed introns in 5' and 3' untranslated regions (UTRs), altered intron/exon boundaries, alternative splicing, overlapping sense-antisense reads, and potentially new transcripts. Interestingly, comparison of the NSM RNA-seq reads obtained here with those of YM discovered differentially spliced introns; e.g., spliced introns in one subspecies but not the other. Alignment of the NSM cDNA sequences to the YM genome sequence also identified ~84,000 SNPs between the two parasites.

The discoveries of UTR introns and differentially spliced introns between *P. yoelii* subspecies raise interesting questions on the potential role of these introns in regulating gene expression and evolution of malaria parasites.

Keywords: Malaria, RNA-seq, Gene expression, Single nucleotide polymorphism (SNPs), Alternative splicing,

第一章 前言

1.1 疟原虫及其生活史

疟疾是由疟原虫 (*Plasmodium*) 引起的经按蚊传播的传染性寄生虫病。疟原虫属 (*Plasmodium*) 隶属顶复合器门 (*Apicomplexa*), 孢子纲 (*Sporozoa*)。大部分脊椎动物都可作为疟原虫中间宿主, 如鸟类、啮齿动物、蜥蜴等。目前, 以人类为宿主的疟原虫有五种, 即恶性疟原虫 (*Plasmodium falciparum*)、三日疟原虫 (*Plasmodium malariae*)、卵形疟原虫 (*Plasmodium ovale*)、间日疟原虫 (*Plasmodium vivax*) 以及诺氏疟原虫 (*Plasmodium knowlesi*)。其中非洲流行疟疾主要是恶性疟原虫, 它也是造成患者死亡率最高的疟原虫; 流行最广泛的疟原虫是间日疟原虫。原本在长尾猴和短尾猴中传染的诺氏猴疟原虫近年也被报道能通过自然界的传疟按蚊感染人类^[1, 2], 其可以通过传疟按蚊造成猴与人、人与人之间的相互传染^[3, 4], 据报道我国最早在 1978 年有解剖感染诺氏疟病死恒河猴意外手术割伤而感染人的诺氏疟病例^[5], 2005 年有见报道称复查镜检并 PCR 验证 1998 年感染的疟疾病例血涂片, 检测到我国第一例由诺氏疟原虫自然感染的患者^[6], 在 2014 年我国疟疾疫情中亦有报道 1 例诺氏疟原虫感染病例^[7]。

疟疾由感染疟原虫的雌性按蚊叮咬人体, 并将其体内寄生的疟原虫子孢子传入人体内而引起, 临床表现为周期性冷热发作, 脾肿大, 贫血, 脑、肝、肾、心、肠、胃等受损引起的各种综合症。

根据疟原虫所处的环境和自身形态的不同, 疟原虫的生命周期可以大致分为三个阶段 (图1-1): 包括在人体内发育的肝细胞内阶段和红细胞内阶段以及按蚊体内的有性期阶段:

(1). 红细胞外期(exo-erythrocytic cycle, 红外期)即肝细胞内阶段 (图1-1 A): 当唾液腺中带有成熟子孢子(sporozoite)的雌性按蚊刺吸人血时, 子孢子随唾液进入人体, 约经30分钟后随血流侵入肝细胞, 裂体增殖形成红细胞外期裂殖体。成熟的红细胞外期裂殖体内含数以万计的裂殖子^[8]。裂殖子胀破肝细胞后释出, 一部分裂殖子被巨噬细胞吞

噬，其余部分侵入红细胞，开始红细胞内期的发育。各种疟原虫完成红细胞外期所需时间不同：间日疟原虫约8天，恶性疟原虫约6天，三日疟原虫为11~12天，卵形疟原虫为9天。

目前一般认为间日疟原虫和卵形疟原虫子孢子在遗传学上有不同的两种类型，即速发型子孢子(tachysporozoites, TS)和迟发型子孢子(bradysporozoites, BS)。它们的区别主要在于子孢子进入肝细胞后，速发型子孢子继续发育完成红细胞外期的裂体增殖，而迟发型子孢子需经过一段时间(数月至年余)休眠期之后，才完成红细胞外期的裂体增殖。

(2). 红细胞内期(erythrocytic cycle, 红内期; 图 1-1 B): 红细胞外期(肝细胞内期)的裂殖子从宿主(脊椎动物)肝细胞释放出来，进入血流后很快侵入宿主红细胞。侵入红细胞的裂殖子经两个时期：环状体(ring)时期、滋养体(trophozoite)时期，而形成包含一定数量裂殖子的裂殖体(schizont)。包含裂殖子的红细胞破裂后，裂殖子释出血液中，重复红细胞内期的裂体增殖过程。完成一代红细胞内期裂体增殖，因虫株不同各种疟原虫完成裂体生殖的时间也有差异，间日疟原虫约需 48 小时，恶性疟原虫约需 36~48 小时，三日疟原虫约需 72 小时，卵形疟原虫约需 48 小时。

疟原虫在经过几代红细胞内裂体增殖后，部分裂殖子侵入红细胞后不再进行裂体增殖而是发育成雌、雄配子体(female/male gametocyte)。恶性疟原虫的配子体主要在肝、脾、骨髓等器官的血窦或微血管里发育，成熟后才出现于外周血液中(约在无性体出现后7~10天才见于外周血液中)。配子体的进一步发育需在蚊胃中进行，否则在人体内经30~60天即衰老变性而被清除。

(3). 有性期阶段(sexual stages, 蚊期发育阶段, 有性期阶段; 图 1-1 C): 当雌性按蚊叮咬疟原虫终宿主时，雌、雄配子随血液进入蚊胃继续发育。雄配子体核分裂形成细丝状的雄配子(male gamete)，雄配子在蚊胃中自由游动至雌配子(female gamete)，受精形成合子(zygote)。二倍体的合子减数分裂后形成动合子(ookinete)。动合子穿过胃壁上皮细胞或其间隙，在蚊胃的基底膜下形成圆球形的卵囊(oocyst)。卵囊内的细胞核和胞质反复分裂，以孢子增殖的方式形成数以万计的子孢子(sporozoite)。子孢子随卵囊破裂释出或由囊壁钻出，经血淋巴集中于按蚊的唾液腺，发育为成熟子孢子。在最适条件下，不同种疟原虫在按蚊体内发育成熟所需时间有所不同：间日疟原虫约为9~10天，恶性疟原虫约为12~15天，卵形疟原虫约为16天，三日疟原虫约为25~28天。

Life Cycle

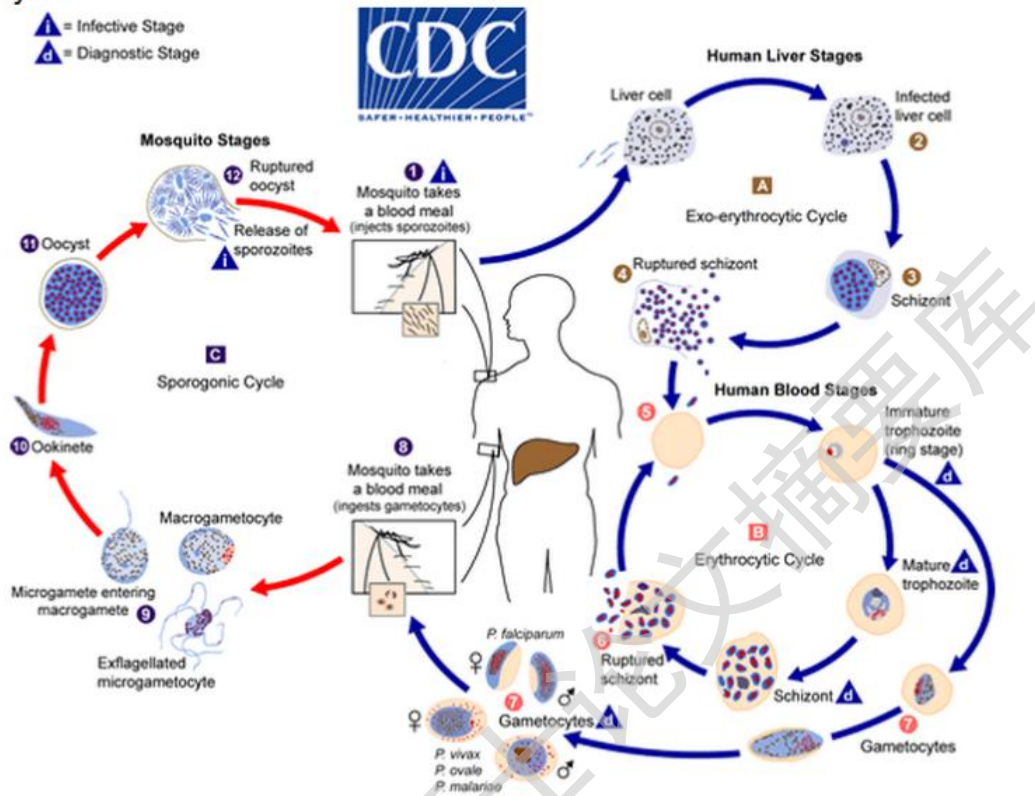


图 1-1. 疟原虫生活史

Figure 1-1. The life cycle of malaria parasite.

(图片来源 <http://www.cdc.gov/malaria/about/biology/index.html>)

A、肝细胞时期，亦称红细胞外期或称肝细胞阶段，子孢子入侵肝细胞，无性生殖形成裂殖体，裂殖体破裂释放裂殖子；B、红细胞时期，即红细胞内期，裂殖子侵入红细胞，此时的疟原虫为环状体时期，而后分为两个部分分化，一部分环状体在红细胞内成长为成熟的滋养体进而无性生殖分裂为裂殖子，此时有多个裂殖子的红细胞被称为裂殖体，成熟裂殖体破裂会释放出大量裂殖子进一步侵入新的红细胞；另一部分环状体经多次裂体生殖后转向有性生殖发育为配子体，进而分化为雌性配子体和雄性配子体，在按蚊吸食血液时进入按蚊体内，开始疟原虫有性生殖阶段；C、蚊期，即孢子生殖即有性生殖阶段，进入按蚊体内的雌、雄配子体形成一个动合子，动合子形成卵囊，成熟的卵囊释放子孢子，在按蚊吸食人血时子孢子经唾液腺进入人体，开始疟原虫下一周期生活史。

疟疾的流行分布非常广泛，其界线大致处于北纬 60 度和南纬 40 度之间，这个范围之外的地区不发生疟疾的原因主要是气候条件的限制，疟疾几乎不能在月份平均气温在 15.6℃ 以下的地区，而南北半球月平均 15.6℃ 等温线相当于北纬 60 度和南纬 40 度线。但在这条带状地域内，大部分地区也没有疟疾的传播，海拔高度是一个限制性条件，1000 米以上的地区疟疾就很少见，且愈高愈少。这可能与疟疾的传播媒介按蚊的生活习性有关，按蚊在低温和高海拔地域其生活力大大降低。

1.2 全球疟疾现状

据世界卫生组织 (World Health Organization, WHO) 发布的《2015 年世界疟疾报告》^[9]，全球疟疾病例正逐渐下降，从 2000 年的 2.62 亿，降到 2015 年的 2.14 亿，大约减少了 18%。2015 年度报告显示，疟疾感染病例主要发生在非洲区域（约占 88%），接下来是东南亚区域（约占 10%）和地中海东部地区（约占 2%）。考虑到人口增长的因素，疟疾发病率在 2000 年到 2015 年间下降了 37%。总的来说，在 2000 年有持续性传染的 106 个国家中，已经有 57 个国家减少了疟疾发病率 >75%，有 18 个国家被估计减少了疟疾发病率在 50-75%（图 1-2）。

全球疟疾死亡人数从 2000 年的 83.9 万人下降到了 2015 年的 43.8 万人，从死亡率的情况看减少了 48%。在 2015 年主要的疟疾死亡病例发生在非洲（90%），其次是东南亚地区（7%）以及地中海东部区域（2%）。在 2000 年到 2015 年之间全球疟疾致死率共减少了 60%。

据统计，全球五岁以下儿童的患疟死亡人数从 2000 年的 72.3 万降至 2015 年的 30.6 万。儿童死亡人数的减少多数发生在非洲地区，在那里估计死亡人数从 2000 年的 69.4 万人下降到了 2015 年的 29.2 万人。疟疾已不再是撒哈拉以南非洲地区儿童首要的死亡原因，在 2015 年，最高死因排名疟疾是第四位，占撒哈拉以南非洲地区儿童死亡人数的 10%。然而，疟疾依然是一个主要的儿童杀手，尤其在撒哈拉以南非洲地区，每两分钟就有一个儿童丧失生命。感染疟原虫的儿童比例从 2000 年开始在非洲疫区已经减半。在 2-10 岁儿童中的感染流行估计从 2000 年的 33% 已经下降到了 2015 年的 16%，这一改变有四分之三发生在 2005 年之后。

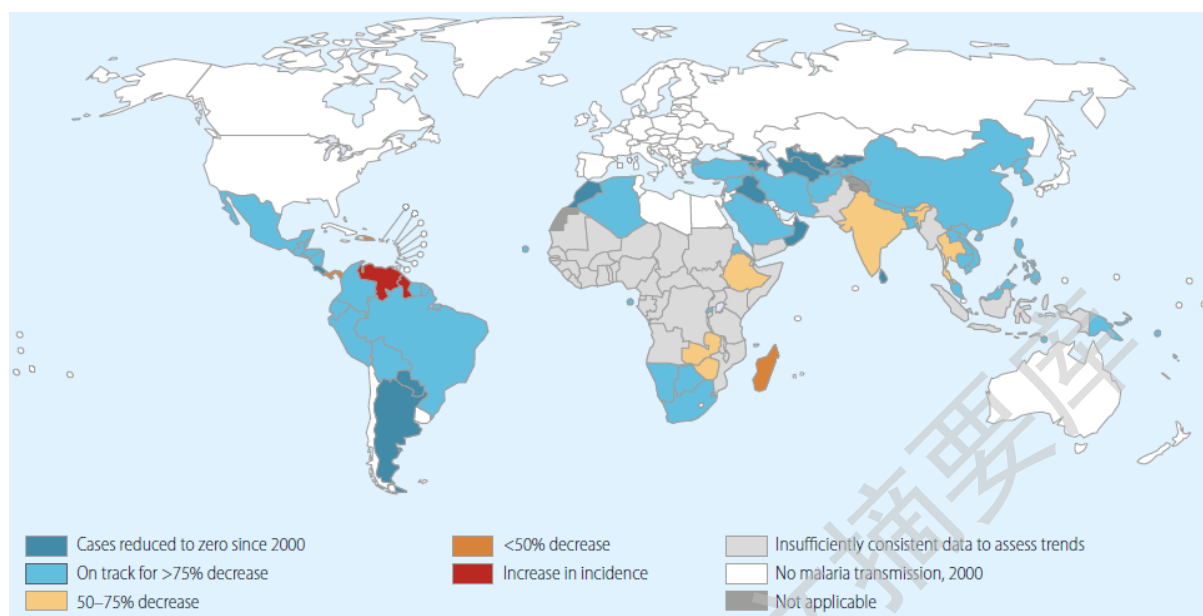


图 1-2. 2000-2015 年世界各国疟疾发病率变化

Figure 1-2. Projected changes in malaria incidence rates, by country 2000-2015 (World malaria report 2014)^[10].

在 2000 年到 2015 年间,按 2000 年度的发病率和死亡率计算已经累计在全球减少 12 亿人患病和减少 620 万人死亡。在撒哈拉以南的非洲,在 2001 年至 2015 年间,由于疟疾控制干预措施而减少了 70% 的疟疾发病,避免了 6.63 亿的疟疾病例。因为疟疾干预措施的施行而避免感染的 6.63 亿人口中,69% 得益于杀虫剂药浸蚊帐的使用,21% 得益于青蒿素复方疗法的推广,另有 10% 是由于室内滞留喷雾的使用。

越来越多的国家正步入消除疟疾的行列。在 2000 年仅仅有 13 个国家消除疟疾到少于 1000 个病例,但是在 2015 年有 33 个国家估计达到了这个标准。同时,在 2014 年,已有 16 个国家报告零感染本土病例,另外有 3 个国家和地区报告少于 10 个本土病例。欧洲地区在 2015 年首次报告零例本土病例^[9]。

然而,随着全球气候变暖,气候的变化影响了疟原虫和按蚊的生态环境,使疟疾传播季节延长;流动人口增加以及带虫者传入又使得一些原来已经消灭疟疾的地区重新发病;按蚊生活习性的改变及疟原虫耐药性的增强也都会使疟疾发病率升高,这些因素增加了今后疟疾防治的难度。

Degree papers are in the “[Xiamen University Electronic Theses and Dissertations Database](#)”.

Fulltexts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.