

虽然人们可能无法取得前几年那么大的成就，但当最终弄懂了人类常见病背后的发病机制时，就会发现，今天的工作还是非常有意义的。

原文检索：Peter Donnelly. (2008) Progress and challenges in genome-wide association studies in humans, *Nature* 456 (7223): 728-731.



## 小词典

### 1. 荟萃分析 (meta-analyse)

荟萃分析，又称“Meta分析”。Meta意指较晚出现的更为综合的事物，而且通常用于命名一个新的相关的并对原始学科进行评论的学问，不但包括数据结合，而且包括结果的流行病学探索和评价，并以原始研究的发现取代个体作为分析实体。荟萃分析产生的主要的理由是：对于多个单独进行的研究而言，许多观察组样本过小，难以产生任何明确意见。在过去的十年中，医学研究领域中有有关荟萃分析的已发表的论文数量急剧地增加。那么，究竟什么是荟萃分析？Huque给出了一个有用的定义：“一种对被分析家认为可组合的多个各自独立的临床实验结果进行组合或整合性统计学分析。”荟萃分析的过程必须有统计学的参与，目前有相当多的统计学软件中已经包含了相应的分析模块，专门的Meta软件也已经问世。

## 四、全球联手，消灭疟疾 ——疟疾基因流行病学的研究

对疟疾基因组内的遗传变异现象进行大规模研究，将有望攻克疟疾这一顽症。但是，在疟疾肆虐的地区——发展中国家，进行这样的研究会受到科技发展水平和伦理道德的制约，同时还存在可操作性方面的问题。疟疾基因流行病学计划（Malaria Genomic Epidemiology Network, Malaria GEN）就是在这样的大背景下出现的。该计划集合了全球21个国家的众多科学家。人类希望集全球之力一起攻克疟疾。

疟疾是由单细胞寄生性的疟原虫（*Plasmodium*）造成的一种疾病。疟原虫是一种原生动物，生活史非常复杂（图1）。

自然界中有100多种疟原虫，其中5种能感染人类。而这5种疟原虫中，恶性疟原虫（*Plasmodium falciparum*）、间日疟原虫（*Plasmodium vivax*）、卵形疟原虫（*Plasmodium ovale*）以及部分诺尔斯（氏）疟原虫（*Plasmodium knowlesi*）能引起严重的疟疾，它们是大部分致死性疟疾的元凶。疟疾通过按蚊（*Anopheles*）的叮咬在人群之间传播。按

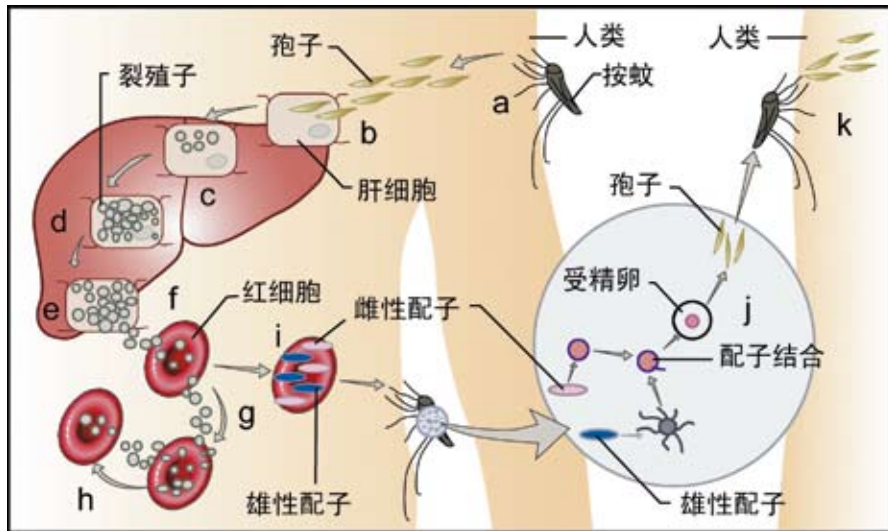


图1 疟原虫的生活史。a. 按蚊叮人。b. 疟原虫孢子随按蚊唾液流进人体血液，并随血液进入肝脏，入侵肝细胞。c. 肝细胞内，孢子体变成裂殖体。d. 裂殖体遗传物质大量复制，产生数以千计的裂殖子。e. 肝细胞破裂，释放出裂殖子。f. 裂殖子进入血液，侵入红细胞，变成滋养体。g. 滋养体吞食血红蛋白，逐渐生长，直至充满整个红细胞。接着，滋养体变成裂殖体，产生新的裂殖子。随后，红细胞破裂，释放出裂殖子。h. 裂殖子进而入侵其它红细胞。i-j-k. 也有些滋养体在红细胞中形成配子，在按蚊叮人时，如果吸入了含有疟原虫配子的红细胞，疟原虫配子就进入了按蚊的胃中，在其中形成孢子体，准备感染其他人。

蚊属于蚊子的一种，自然界中有大约400多种按蚊，其中40多种能传播疟疾。冈比亚按蚊（*Anopheles gambiae*）（图2）是最主要的传播媒介。

每年，疟疾会造成全球100万儿童死亡，超过5亿人发病。这种疾病大流行的背后实际上存在一个三种真核生物共进化（co-evolution）的问题（表1）。

深入理解疟原虫、按蚊和人这三个物种间生物进化竞赛的详细情况是攻克疟疾的关键。为了对抗抗疟疾药物，疟原虫也在不断进化。在长期寄生于人体的过程中，它们进化出了一套精妙的遗传机制来逃避人体免疫系统的“攻击”，这就是为何难以获得疟疾疫苗的主要原因。同样，按蚊也在不断地进化以对抗灭蚊剂。不过它们为何能够不被疟原虫感染倒是值得研究的问题，这也许能够给人们控制疟疾提供一丝线索。此外，疟原虫还是一个促进人类进化的强有力的因子。研究人员发现，某些疫区的人群对疟原虫表现出天然的抵抗力，如果能发现人体内具有抑制疟原虫繁殖的因子，弄清楚保护性免疫的分子机制，就有可能开发出预防疟疾的疫苗。



图2 冈比亚按蚊正在叮咬人类皮肤。  
图片来源：CDC

表1 疟疾相关的三种真核生物基因组

比较层面	疟原虫 ( <i>Plasmodium falciparum</i> )	按蚊	人类
基因组类型	单倍体（在蚊体内为二倍体），由14对染色体、线粒体基因组和apicoplast基因组组成	二倍体，2对常染色体，1对X、Y染色体和线粒体染色体	二倍体，22对常染色体，1对X、Y染色体和线粒体染色体
基因组大小 (碱基对)	$2.3 \times 10^7$	$2.8 \times 10^8$	$3.2 \times 10^9$
预计编码基因数	5,420	12,945	21,714
基因组变异数据库	约有 $10^5$ 个SNP收录在dbSNP数据库中 ( <a href="http://www.ncbi.nlm.nih.gov/projects/SNP">http://www.ncbi.nlm.nih.gov/projects/SNP</a> )。	约有 $10^6$ 个SNP收录在dbSNP数据库中。	约有 $6 \times 10^6$ 个SNP收录在dbSNP数据库中， $3 \times 10^6$ 个SNP收录在国际HapMap计划数据库中 ( <a href="http://www.hapmap.org">http://www.hapmap.org</a> )； $3 \times 10^4$ 个SNP收录在基因组变异数据库中 ( <a href="http://projects.tcag.ca/variation">http://projects.tcag.ca/variation</a> )。
进化选择事例	对抗疟疾药物产生抗药性。例如，因DHFR基因和DHPS基因位点突变造成对叶酸拮抗剂产生抗药性；因CRT基因及MDR基因突变对氯奎产生抗药性。	对灭蚊剂产生抗药性。例如，因kdr等位基因（该基因编码钠离子通道）突变造成对DDT和除虫菊素类灭蚊药产生抗药性。	对疟疾具有抵抗力。原因是1. HBB基因突变（形成异常血红蛋白HbS和HbC）；2. 携带HBA1或者HBA2基因（分别编码不同的血红蛋白 $\alpha$ 链而形成地中海贫血）；3. 缺乏G6PD基因。

注：CRT：抗氯奎转运基因；DDT：二氯二苯三氯乙烷，即滴滴涕；DHFR：二氢叶酸还原酶基因；DHPS：二氢叶酸合成酶基因；G6PD：葡萄糖-6-磷酸脱氢酶基因；HBB：血红蛋白 $\beta$ 链基因；kdr：敲除抵抗基因；MDR：多重抗药基因。

现在，由于有了新的全基因组关联分析（GWA）技术，人们得以在全基因组水平来研究人类抵抗疟原虫的遗传机制。在一个典型的GWA分析中，研究人员可以通过50多万个SNP对成千上万人的基因型进行鉴定。他们希望通过GWA分析和基于基因组内常见变异模式的统计推理方法，能够在不进行全基因组测序的情况下，找到与人体对疟疾具有天然抵抗力相关的所有的DNA变异位点。

使用GWA方法可以解决一个重要的问题——为何在同一个社区里几乎每一个孩子都会反复感染恶性疟原虫（导致大部分疟疾死亡的罪魁祸首），但其中只有部分孩子会发展成致命的凶险

型疟疾呢？流行病学数据显示，只有少部分恶性疟原虫感染患者会发展成重症疟疾，而这些重症疟疾患者的致病因素中有25%是遗传因素（图3）。

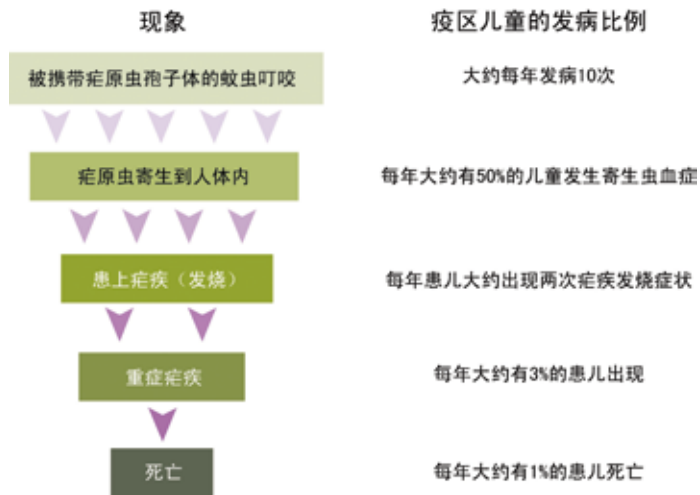


图3 疟疾疫区的疟疾发病状况。在疟疾流行区，感染恶性疟原虫的患儿可能出现的临床症状也不尽相同。一般情况下，他们感染疟疾后只会出现周期性发热等症状，但最终会被治愈。少部分患儿会发展成重症疟疾（重度贫血或脑型疟）这种危及生命的急症。反复感染后，人体会对疟原虫产生临床免疫力，即再次感染后不会再出现临床症状。上述数据只是一个参考概念，比如蚊虫叮咬频率等会受到很多环境因素的影响。除了环境因素之外，人类遗传因素也会起到一定的作用。据估计，影响力大约有25%。不过，人们到目前为止知道的有效遗传因素（如表1中所列出的HbS等）只占到了很少一部分，这也就意味着还有很多遗传背景因素等待人们去发现。这些遗传因素可以在疟疾发病的各个阶段起作用。例如，携带HbS基因的个体患重症疟疾的风险比不带该基因的个体低10倍，可能是因为HbS血红蛋白能控制血液中疟原虫的数量。

进行一个典型的GWA分析需要两个实验组，一组来自医院重症疟疾患者，另一个对照组则由随机抽取的普通人组成。比较这两组人群中某些SNP的出现频率，就有可能发现不同的遗传变异（SNP）对个体患上重症疟疾的影响力。因为个体是否发展为重症疟疾会受到很多遗传因素、环境因素以及这些因素在疾病的哪个阶段起作用等等因素的影响，这里面任何一个因素的作用可能都很小，因此只有调查大量的个体，通过统计学的方法才能发现这些因素的微小作用力。

原则上，人们也可以使用相同的方法来研究疟原虫抗药性或按蚊的抗药性等问题。不过，至少从目前来看，这还是不可行的，因为人们现在对疟原虫和按蚊基因组内的变异情况基本上还处于一无所知的阶段。使用GWA分析法研究恶性疟原虫还存在一个问题——蚊虫叮咬一次可能会使患者被好几种基因型不同的疟原虫感染。在非洲，还有这样一种情况，由于疟疾肆虐，恶性疟原虫基因组的连锁不平衡（linkage disequilibrium）水平非常低，而连锁不平衡现象正是GWA分析的基础。连锁不平衡指的是在某一群体中，不同位点上某两个等位基因出现在同一条单倍型上的频率与预期的随机频率之间存在明显差异的现象。在进行GWA分析时，如果基因连锁不平衡水平越低，就需要分析基因组中更多的位点才能得到有意义的结果。最近，单分子DNA大规模平行测序（massively parallel sequencing）技术的进步将有望解决这些问题。通过对恶性疟原虫在基因组中多个位点进行基因分型分析，以及对导致同一种感染的不同寄生虫基因型进行区分，可帮助研究人员取得有意义的研究结果。

下文将介绍全球疟疾研究协作体如何在生物学层面和致病性方面研究基因变异对人患疟疾的影响。本文主要研究的是人类基因组，因为如上所述，人们对人类基因变异情况的了解以及针对人类的研究方法要比对疟原虫和按蚊的多得多。而且，本文还突出了可行性问题，即为什

么用GWA方法研究疟疾比研究其它疾病要困难得多。与此同时，本文不仅介绍了研究人员如何在全球各地建立好几个大规模研究项目来共同解决这一个难题，以及从不同研究人员那里收集数据、标准化数据的操作流程，还将介绍研究人员制定的有关如何处理数据、标本所有权、知识产权以及伦理方面问题的政策。

## 1. GWA分析法用于研究疟疾所面临的问题

使用遗传分析法来研究为什么有些人对疟原虫具有天然免疫这一课题还存在着好几方面的困难。例如，临床研究的可行性问题、伦理道德问题，或是由于受试者来源广泛，个体之间的巨大差异导致的统计学方面的问题等等。

要收集到足够多的重症疟疾病例本身就是一件非常困难的事情。因为大部分的疟疾患者都来自贫民区，区内医疗卫生服务水平低下，根本没有系统的医疗记录。有大量的重症疟疾儿童患者在送达医院的数小时内死亡。因此，要获得准确的信息必须在患者入院的当时进行调查取样。这同时也意味着要保证医疗救治的质量，尤其在医疗资源匮乏的地区，科研人员的作用显得更为重要。同时，由于大量疟疾患儿也存在严重的贫血情况，所以就无法大量抽取他们的血液作为研究样品，而只能使用全基因组扩增技术获得足量的DNA样品。但是，这样就会降低基因分型的效率，从而影响了统计分析的准确性。此时只能靠扩大样本量来提高准确率，但这就进一步增加了工作的难度。

另外，在非洲，要设计一个优秀的适合用于GWA分析法的SNP基因分型研究策略也是非常具有挑战性的工作。因为在非洲人群里连锁不平衡的水平非常低（比欧洲人群里的低得多），所以需要更多的SNP位点进行测序才能进行准确的基因分型。以HapMap (<http://www.hapmap.org>)的数据为基础进行预测发现，要对非洲人进行GWA分析，需要检测150万个SNP位点，而对欧洲人则只需要检测60万个SNP位点，但这只是针对检测常见SNP位点而言。现在，人们还无法估计要检测所有变异情况共需要检测多少个SNP位点，除非能对非洲人种群进行大规模基因组测序，以求得到一份完整的非洲人常见基因变异图谱。

不同非洲种群间存在的人种差异对于GWA研究所需要的统计操作来说也是一大难题。许多非洲部落由好几个不同的人种组成，这些不同人种间哪怕微小的差异也会造成假阳性结果。因此，必须设计一个周密的研究计划才能保证研究工作的顺利进展并得出有价值的研究成果。

另一个问题则来自非洲人群间的遗传差异，这和在同一种群中进行研究是不一样的。在非洲不同地区人群间进行GWA分析并不能总是得出相同或相似的结果。比如，在研究哪种基因变异会导致哪种疾病的时候，基因组单倍体型结构（haplotype structure）的差异就会导致不同结果的出现。尤其是当基因位于基因组中“刚刚”经受过遗传选择的部位时更是如此。不同的人群可能携带的抗疟原虫基因也会有所不同，比如携带血红蛋白S或血红蛋白C都可以免遭疟原虫感染。出现这两种不同的血红蛋白是因为编码血红蛋白β链的HBB基因中出现了两种SNP突变造成的。而在非洲人群中，这两种SNP的分布情况并不相同。

不过，这种人群间的遗传差异也并非没有意义。例如，它们可以帮助人们发现只在某个特定种群中发生进化的遗传因素，或者在研究基因与环境相互作用的过程中发挥作用。更为重要的是，不同种群间连锁不平衡方式的差异能帮助人们通过邻近多态性发现具有决定作用的因果变异（causal variant）。这一发现非常重要，因为很多与疾病相关的SNP都并非致病的主要原

因，而仅仅显示出与决定性基因变异之间具有的相关性（由连锁不平衡现象导致）。所以，在非洲各地进行GWA研究可以得到很多非常有价值的信息，从而有助于找到决定性基因变异。

## 2. 建立一个全球合作机制

过去，有几家科研机构各自独立地就人体对疟原虫产生自然免疫现象背后的遗传机制进行了研究，不过他们研究的样本量都太小，研究的范围也比较窄。如果能有这样一种机制，而让这些分布在全球各地的研究机构之间共享数据，以此来扩大他们的样本量，那么大家获得成功的可能性就会大幅度提高。鉴于此，2003年，比尔与梅林达·盖茨基金会（Bill & Melinda Gates Foundation）与英国医学研究理事会（UK Medical Research Council）一起资助了一个项目。后来，该项目发展成了全球性的合作项目——疟疾基因流行病学计划（Malaria GEN）。该项目旨在开发一个网络软件，凭借该软件，全球的临床工作者和科研工作者都可以与他人共享各自的研究数据。同时，比尔与梅林达·盖茨基金会还资助了于2004年1月在加纳首都阿克拉召开的一场研讨会。来自非洲的10个医疗、科研团体参与了该次研讨会。会上，他们就与数据共享有关的伦理及知识产权问题进行了讨论。

Malaria GEN成立于2005年，由比尔与梅林达·盖茨基金会和惠康基金会（Wellcome Trust）共同资助。这两个基金会共同成立Malaria GEN项目就是为了在疟疾流行区进行大规模的流行病学调查，同时结合人类基因组变异研究来揭示人类对疟疾产生保护性免疫背后的机制。不过，要达成这一目标首先必须完成5个项目，即建立一个全球性的疟疾基因流行病学数据共享网络；搜集不同疟疾病情患者的DNA样品和临床数据资料；弄清疟疾疫区人群的遗传变异情况；找到与保护人体免受重症疟疾相关的遗传变异位点以及发现这种自我保护背后的免疫学机制。

参与Malaria GEN项目的研究人员大部分都是在疟疾疫区长期进行研究工作的临床工作者、流行病学工作者以及免疫学工作者，他们也都把自己多年来的研究数据和研究样品提交给了Malaria GEN项目。还有一些人则承担了高通量遗传变异分析（high-throughput analysis of genomic variation）、统计遗传



学 (statistical genetics) 或生命医学伦理 (biomedical ethics) 方面的技术支持工作。Malaria GEN项目主要由分布在全世界15个疟疾流行国家和6个非疟疾流行国家的Malaria GEN项目合作研究所组成, 更多信息请参见 [www.Malaria GEN.net/resource/1](http://www.Malaria GEN.net/resource/1)。这些研究所的疟疾研究基地大部分都位于撒哈拉以南非洲地区, 例如布基纳法索、喀麦隆、冈比亚、加纳、肯尼亚、马拉维、马里、尼日利亚、塞内加尔、苏丹和坦桑尼亚。还有一些研究基地位于巴布亚新几内亚、斯里兰卡、泰国和越南。

由于Malaria GEN计划是一个非常复杂的项目, 因此所有参与该计划的人员于2005年7月在英国牛津召开的Malaria GEN成立大会上一致同意, 分四个阶段完成Malaria GEN计划。

**第一阶段:** 建立一套标准化的操作规程和制度, 以统一DNA样品和表型数据的质量和规格 (知识框1)。更加重要的是还要建立一套规范以保证大家在共享数据的同时又能拥有各自的知识产权。

**第二阶段:** 建立一个核心的科学实验流程和策略, 该步骤中的一部分项目也被称为Consortial Project (知识框1, 表2)。

**第三阶段:** 帮助在疟疾流行区工作的科研人员开展临床及流行病学研究工作, 以完成第

二阶段制定的目标任务。参与的科研人员需要向Consortial Project项目组提交科研计划书, 该计划书必须与实施项目所在地研究院的科研方向相一致, 所用到的实验设备也必须是当地能提供的。然后由一个由来自喀麦隆、冈比亚、加纳、意大利、肯尼亚、马拉威、马里、斯里兰卡、苏丹、坦桑尼亚和英国的科学家组成的评审小组对这些计划书在科研设计和计划的可操作性方面进行审议, 如果评审通过就会给予相应的资金资助。

**第四阶段:** 数据管理和进行统计及遗传学分析。就数据分析工作也专门成立了一个国际协作组织。通过一套公开的选拔程序, 在每一个参与Malaria GEN项目的合作研究机构都会选出一名数据处理人员, 他们与Malaria GEN项目研究人员一起工作, 负责该团队的数据处理工作。所有的数据处理人员都会受到数据处理、统计遗传学和计算机操作方面的培训, 而且在整个工作过程中都提供全程的技术支持服务。负责培训和技术支持工作的是英国牛津大学 (University of Oxford) 和韦尔科姆基金会桑格学院 (Wellcome Trust Sanger Institute) 的统计学、遗传学和计算机科学专家组成的Malaria GEN资源中心。该中心还会定期在英国及各个研究所所在地举办数据分析研讨会, 可以在会上对数据分析人员进行系统培训, 数据分析人员也可以分享各自的心得体会, 还能在专家的亲自指导下进行数据分析操作。

## 知识框1 Malaria GEN项目介绍

Malaria GEN参与者为数据和研究样品共享制定了一套共同认可、遵守的原则和工作流程。其中重要的工作就是制定了4个Consortial Project项目, 每一个Consortial Project项目都有各自的行动计划和目标 (表2)。每一个Malaria GEN计划的参与者都可以选择将自己的数据提交给哪一个Consortial Project项目。

### 定义

**Consortial Project:** 指由项目管理委员会认可, 由Malaria GEN核心基金会负责实施, 所有的研究人员参与进行的一项项目。各个研究人员的数据和专业知识可以通过该项目得到共享。

**资源贡献研究者:** 将资料、数据、样品以及个人专业知识提交给Consortial Project项目的研究人员。

**研究者个人分析:** 指由资源贡献研究者就其本人获得的样品进行的分析。

### 重要原则

提交给Consortial Project项目的样品或临床数据的所有权属于资源贡献研究者。

资源贡献研究者可以在他同意进行的实验完成之后的任何时候要求返还他提交的DNA样品，也有权将该样品用于其它用途。

资源贡献研究者有权使用Consortial Project项目的所有基因分型数据资料，也有权将这些资料用于他们自己的科研工作之中。

资源贡献研究者负责保证他所提供的样品必须获得患者的知情同意，同时不违反当地的伦理道德并获得Malaria GEN伦理委员会的认可。

负责处理样品和数据的实验室要保证样品的安全和完好，并保证数据的机密性。

Malaria GEN出版的著作必须反映出所有对该项目有贡献的资源贡献研究者的贡献。

通过GWA研究得到的个体数据必须通过一个独立的数据系统对科学界开放。

只有在科学发现转变成能解决大众医疗问题的时候，比如生产出治疗疟疾的药品，才能谋求知识产权保护。

了解更多信息请登陆[www.Malaria.GEN.net/resource/1](http://www.Malaria.GEN.net/resource/1)。

表2 Malaria GEN Consortial Project项目介绍

项目目标	项目开展地点	项目介绍
1.分析与人类对重症疟疾易感或免疫现象相关的基因变异情况	布基纳法索、喀麦隆、冈比亚、加纳、肯尼亚、马拉维、马里、尼日利亚、巴布亚新几内亚、坦桑尼亚、越南	搜集了1万名以上的重症疟疾病例和对照组人群进行遗传联合研究； 在2010年之前要在3个研究地点完成对超过1万个人的GWA研究。
2.分析决定抗体对疟疾反应性的人类遗传因素	布基纳法索、肯尼亚、马里、塞内加尔、斯里兰卡、苏丹、坦桑尼亚	正在开发能鉴定疟疾抗原的特异性抗体的标准化试验方法； 该方法可以在探索基因与抗体反应性关联的研究中发挥作用。
3.分析疟疾流行区人群中基因组变异情况	布基纳法索、喀麦隆、冈比亚、加纳、肯尼亚、马里、尼日利亚、巴布亚新几内亚、苏丹、坦桑尼亚、越南	建立了一个DNA数据库以便在疟疾疫区进行人类基因组变异研究； 针对每一个位点至少需要对90名个人进行检测，同时要获得同意，才会公开发布该序列。
4.对与疟疾相关的数量性状进行遗传连锁研究	加纳、塞内加尔、泰国	在从事疟疾感染水平以及其它数量性状遗传连锁研究的科技工作者间建立了一个数据共享组织。



## 3. 数据共享

共享数据听起来简单，但做起来就很困难了。

首先，将这么多研究团体的数据搜集起来就是很困难的一件事。

其次，还存在数据归属权，即知识产权的问题。如何管理、发布数据以及用户如何使用数据等问题。

接下来将介绍Malaria GEN如何解决这些问题。

### 3.1 如何标准化处理数据以及如何搜集整合数据

将来自世界各地的研究数据搜集、整合起来并进行标准化处理是Malaria GEN项目的主要任务。比如，要完成其中心项目Consortial Project，首先就需要对重症疟疾有一个标准化的临床定义。重症疟疾包含好几种临床症状，也可以被称作几种亚型，即脑型疟疾（cerebral malaria，主要症状是昏迷）、重症贫血（profound anaemia）和呼吸窘迫（respiratory distress）。有一些人体遗传因素与保护患者不发展成重症疟疾有关，而另一些遗传因素则刚好与发生重症疟疾相关。因此，对重症疟疾下一个标准化的诊断标准就显得非常重要，因为有一些亚型是能够通过定量来判断的，例如贫血（检测血红蛋白浓度），但有一些亚型却无法通过定量判断，比如呼吸窘迫。2005年11月，在喀麦隆雅温得举办的非洲儿童重症疟疾大会（Severe Malaria in African Children network）上，各国研究人员和Malaria GEN项目的专家就对重症疟疾下了一个定义，详见<http://www.Malaria GEN.net/resource/1>。

在大型研究体系中，记录临床数据和流行病学数据时，会出现一个基因位点之间存在

变异（site-to-site variation）所带来的问题。因此，进行数据标准化处理就显得尤其重要。

Malaria GEN资源中心为此特意开发了一款网络版软件。研究人员可以在Malaria GEN网页上申请一个个人账号，将自己的标准化研究数据上传至该账号。由于该账号设有密码，所以其他人是无法进入的。网络上还有用于检查数据完整性和将数据转移到Consortial Project项目相关数据库的软件。在传递数据时，往往需要数据上传者对数据进行相应的修改以保持与上传的数据具有一致性，网络上的软件会协助其完成这一过程并会全程记录。

参与Malaria GEN项目的科研人员还需要对Consortial Project 2项目（该项目主要研究不同人群中与免疫反应和环境因素相关的遗传决定因素）的免疫试验数据进行标准化处理（表2）。在Consortial Project 2项目的第一阶段，主要工作是在中心基准实验室进行抗体测定工作，以保证来自不同地点的数据资料之间可以直接进行比对研究。不过，最终的目标是希望开发出一种实用的检测方法和标准化的检测试剂，以便在世界各地的合作研究机构中进行这些只能在中心基准实验室里开展的工作。

### 3.2 分享数据的同时明晰所有权

Malaria GEN的宗旨在于数据分享。因此，在Malaria GEN成立大会上就数据分享规则以及数据所属权问题进行了讨论并达成了一致（知识框1）。每一个Consortial Project项

目的工作成果都会被写成论文发表，所有参与其中的科研人员都是文章的作者。不过，他们也鼓励每一个参与Consortial Project项目的科研人员就自己的样品和数据进行分析、研究，

将研究成果写成论文单独发表。

研究人员在为每一个Consortial Project项目建立数据库的时候都有一个很重要的问题需要考虑，就是如何保护参与该项目的科研人员的知识产权。因为Malaria GEN数据库没有收录资源贡献研究者的个人信息，也不与任何合作研究所的数据库有链接。但是，Malaria GEN有一条重要原则：资源贡献研究者可以分

析他们自己提交的样品，并可以将他们的研究数据和本地的表型数据结合起来进行分析。这样，就建立了一套标准的操作流程，以保证由合作研究所设立的本地数据库可以收录Malaria GEN项目的数据，而且只要符合一定的伦理规范，他们是可以使用这些数据的。详见<http://www.Malaria GEN.net/resource/1>。

### 3.3 数据开放以及相关知识产权保护

由于GWA研究所带来的科学效益是具有积累效应的，因此，如果某一个研究的结果或数据能与更广泛的研究人员或研究机构共享，其价值将会得到很大的提升。

Malaria GEN就这方面的问题咨询了所有参与Malaria GEN项目的科研人员以及好几个Malaria GEN合作研究所伦理委员会的意见（<http://www.Malaria GEN.net/resource/1>）。广义来说，数据开放需要获得原始的知情同意以及相关伦理委员会的同意。在这个问题上，需要注意两个方面：其一，对研究数据的应用要避免引起伦理学方面的争议；其二，对于处于疟疾肆虐的发展中国家和富裕的发达国家的研究者来说，需要有一个有关发表研究数据结果的公平政策。因为通常而言，发展中国家的研究人员在遗传数据的分析水平和条件上显然要落后于发达国家。在平衡两者之间的力量后，Malaria GEN目前的政策是，在提供数据的研究者获得完整数据后9个月，才会在<http://www.ebi.ac.uk/ega>上公布GWA数据。此外，在疟疾流行国家还成立了工作小组，代表具有协作关系的研究部门以及伦理学问题监督部门行使其职能。有关数据应用的各种问题都会随时向该工作小组汇报，同

时，还会向其咨询有关数据公布政策的各种变化。

对于Malaria GEN而言，在数据发表前，制定出将研究数据作为知识产权予以保护的细则十分重要。如果一项科学发现可以为人类的健康带来好处，那么，相关人员应该全力以赴让那些最需要这一好处的人享受到它，而这一过程中，势必会牵涉到专利的问题。另一方面，如果在研究中，最大化地将相关数据公布于众，而结果又不能立即将研究发现应用于提高人们健康状况的话，就会引起不少争议。Malaria GEN目前的做法是，如果下列三项全部满足，那么就需要寻求知识产权的保护。

第一，该发现必须与医学应用直接相关；

第二，该发现具有立即得到知识产权保护的可能性；

第三，该发现在知识产权的保护下，可以进一步地得到发展（<http://www.Malaria GEN.net/resource/1>）。在这种情况下，那些非盈利的科研组织也有可能需要的情况下获得知识产权的保护。

此外，就其中所带来的经济效益而言，Malaria GEN也可以确保那些参与研究的人员获得相应的报酬。

### 3.4 科学研究中的伦理学问题

在多个国家研究者之间建立合作网络，会引起一系列的伦理学及社会学问题。在发展中

国家进行临床试验时，所涉及的伦理学标准问题相当复杂。此外，对个人基因组信息的详细

记录，也会引起许多新的社会问题。针对这些情况，Malaria GEN已经成立了一个由医学伦理学专家所组成的小组，小组成员与研究人员及合作科研院所一起工作，对不同地点的研究进行评估，以期建立一个属于Malaria GEN的最好的研究模式，规避或解决相应的伦理学问题。该小组还制定针对研究者及伦理学监督部门的培训方案，并在肯尼亚、泰国、越南等地举行研讨会。为帮助研究者克服各种具体的伦理学难题，以及弄清楚不同地区的习惯，小组成员还会亲临位于苏丹、冈比亚等地的研究点进行指导。

其中，最为重要的一项工作，就是找到与试验参与者进行有效沟通的方式。举一个例子，将一个患儿从她居住的村庄里带到一个繁忙的公立医院时，医生需要询问是否可以将给患儿检查过程中取到的血样的一部分用于研究项目。但这种情况下，医生会发现，很难将这一行为定义成是医学诊断所需，还是医学研究之用。诸如“研究”、“遗传学”、“实验

室”、“数据库”之类的字眼对于部分地区的患者及患者家属而言，就好像是没有任何意义的“天书”。所以，找到一个恰当的方式，让他们明白所需要面对的事情，是很重要的。在与研究者及伦理学审查部门的成员沟通交流后，Malaria GEN目前已经制定出一套向项目参与者取得知情同意过程中所需的沟通模版及指导手册（<http://www.Malaria GEN.net/resource/1>）。

此外，伦理学工作小组同时还致力于建立适合多文化背景下使用的咨询模式。Malaria GEN目前所面临的很多伦理学及社会问题都是因为参与研究项目的人员涉及多个国家，既包括富裕国家，也有落后的发展中国家，从而在研究项目进展的方方面面都会碰到因为不同文化所带来的问题。为克服这些困难，Malaria GEN所作出的努力及已经取得的成果都成为此类研究中有关数据分享，获得参与试验者的知情同意等方面的一个范例。

## 4. 未来展望

2008年9月，一项令人振奋的计划——全球疟疾行动计划（the Global Malaria Action Plan, <http://www.rbm.who.int/gmap>）公布了。该计划的目标是：消灭疟疾！该计划受到多个国际发展组织及政府的支持。研究人员预期在2010年，将疟疾病例减少一半；至2015年，消除因疟疾而引起的死亡。但是，人们也应该看到，这一计划的成功，还需要有效的杀虫剂以及抗疟疾药物的产生。此外，如果能够获得抗疟疾的疫苗，那么该计划的长期目标——消灭疟疾将更有可能实现。

基因组流行病学可以更加有效地监测疟原虫对抗疟疾药物抗药性以及按蚊对杀虫剂的抗药性的出现，以及为疟疾疫苗的研发提供更加深入的保护免疫机制，从而帮助人们实现消灭疟疾的目标。

如果基因组流行病学真的能像人们期望的那样发挥作用的话，那么还需要全球的研究人员联合起来，尽可能有效地获得相关的基因数据。那些在疟疾流行国家的研究小组，应该努力获得技术支持以及大规模基因变异分析的平台。而其他研究小组则应该携手对来自疟疾流行国家的基因数据进行分析，从而取得有意义的发现。Malaria GEN也正在努力摸索如何建立和维系科研领域里这样一种新型的全球合作模式。

原文检索：

The Malaria Genomic Epidemiology Network. (2008) A global network for investigating the genomic epidemiology of malaria, *Nature* 456 (7223): 732-738.

 筱玥/编译