

解读生命密码，了解那些「命中注定」



《麻省理工科技评论》10 年合辑（套装全 5 册）

[查看详情 >](#)

解读生命密码，了解那些「命中注定」

对于普通人来说，基因测序获得的数据就好比「天书」，一段段由 A、T、G、C 组成的序列就好像一段段暗码，玄而又玄又不知所云。因而专业的解读，帮助人们解析个人遗传信息与健康风险、身体特征以及血缘关系显得十分必要。

对于目前的测序市场，主要有科研级和消费级两种市场需求的基因检测。对于科研级来说，专业的测序为诊断、治疗提供了精准的遗传数据，但始终曲高和寡，不够「接地气」。对于一部分普通人来说，解读自己更多的是出于好奇，比如祖先从何而来，自己患有高血压、糖尿病的风险有多大，甚至包括自己可能会长多高、活多久。对于这一部分受众来说，精准的全基因组测序显得大材小用，因而最近几年，以 40 岁以下人群为主体的消费级基因测序开始蓬勃起来。

平民化的需求，要求基因的获取方式必须简单易得，因而与传统的基因检测从血液或细胞组织中提取遗传物质不同，消费级的基因测序多通过唾液获得，进而用来了解疾病风险。同样，与严肃的临床应用不同，消费级的基因测序只需要提取消费者关注的遗传信息，因而并不需要全基因组测序，更多的是只选择重要的外显子部分进行测序储存，而这样又无形中降低了测序成本。

尽管只是一种商业化、娱乐化的解读，但个人信息安全的问题仍不容小觑。正是由于如今测序技术的巨大进步，极少量的基因样本便足以追根溯源，了解样品提供者的亲缘关系，甚至可以推测个人特征，在一些情境下甚至可能导致基因歧视。随着商业化基因测序被广泛接纳，个人遗传信息安全的意识也必须全面加强。

如今智能手机拥有生活的每个角落，随着智能化、线上产品的发展，消费者可以

▲ 赞同



Q
评论

🎁
分享

iPhone 为例，消费者可以在应用商店中下载不同的软件，用来实现购物、运动、学习等目的。这些应用软件所带来的生活便利如今早已司空见惯，但你是否想过，也许某一天消费者也可以通过智能手机解读自己的遗传信息，定制私人生活呢？

相比传统的科研级基因测序给患者带来绝对精准严肃的医疗信息，对于消费级基因测序来说，消费者关心的可能只是何种运动、饮食更适合他，或一天是不是必须睡够 8 小时才健康这种五花八门而又「无关痛痒」的问题。而 Helix 正是立志要满足消费者的所有需求——创立 DNA 应用商店。它希望通过与应用程序开发人员、生物公司和提供基因检测服务的医疗机构合作，直接为普通消费者提供有关祖源、娱乐、家庭（遗传病携带者筛查）、健身、健康和营养等方面的基因检测产品和服务，使在线获取基因信息成为可能。

Helix 是 Illumina 在 2015 年成立的公司，旨在打造全球最大规模面向大众的基因信息服务平台，在创立之初，便获得了超过 1 亿美元的投资。基于「一次测序，多次查询」（Sequence Once, Query Oenn）的经营理念，消费者只需通过采集唾液的方式向该平台提供遗传信息样本，便可以根据自身需求，在手机上购买相关应用查询个人信息。由于其创新性的商业模式，为普通消费者提供了更为方便快捷、成本低廉的方式了解个人风险和遗传倾向，在公司创立之初，便被《麻省理工科技评论》评为 2016 年度十大突破性技术^[13]。

2017 年 7 月 25 日，Helix 的 DNA 应用商店（DNA App Store）正式上线，开启了消费级基因组学新纪元。这家位于旧金山的初创公司正式实现了他的商业梦想，成为「基因测序领域的苹果」。就像手机上的应用商店一样，通过 Helix 的应用平台，消费者可以轻松地探秘个人的遗传信息，而初始的基因测序，只需要 80 美元（有折扣时价格甚至低至 20 美元）。

Helix 首先对消费者的外显子组进行测序并且将这些遗传信息存储起来，根据消费者需求将这些信息提供给她 App 商店平台上的其他公司。通过该平台，消费者可以通过支付一定的费用，在祖源、娱乐、家庭遗传病风险、健身、营养及健康等多个方面，选择自己感兴趣的进行基因解读，甚至获得个性化定制服务。该平台真正做到让普通人可以接近自身遗传信息、解读自我，实现个人基因组一站式消费体验。

2018 年 1 月，Helix 宣布了与内华达州的合作项目 Healthy Nevada Project。根据协议，Helix 将为上万名内华达州居民提供外显子测序服务，一旦该项合作

成功，未来很可能复制到其他地区，潜在利润十分可观。

2018 年 3 月，Helix 宣布完成 B 轮融资的第一轮，并计划 B 轮融资 2 亿美元。Helix 称未来资金将主要用于开发基于下一代测序技术的创新型健康产品，医疗保健领域将会是下一个重点。

消费级基因检测领域正处于飞速上升的时期，据估计，截至 2019 年 6 月，约有 2,600 万人购买了消费级基因检测，而其中一半以上来自 2019 年。广阔的受众以及良好的市场接纳度为未来消费级基因检测市场提供了无限的可能，同时 Helix 低廉的价格、多元化的应用平台给了消费者更多的选择，未来的成绩值得期待。

然而对于 DNA 应用程序的商业化推广，部分专家持谨慎态度。来自马萨诸塞州综合医院的 Daniel MacArthur 认为，关于生活方式、营养及健康方面的基因测试往往缺乏科学根据，过度地消费基因检测带来的过高期待，很容易使普通民众对临床上的基因检测价值产生怀疑。

因而内部监管必不可少，对于普通公众来说，很难对一款 DNA 应用的科学价值做出判断，为了避免噱头式的骗局发生，Helix 必须加强对平台的监管，在每一款应用上架前都必须进行测试验证，保证程序质量。同时，美国食品及药品监督管理局也在随时关注基因测序的信息安全问题。

除了满足消费者的猎奇心理，消费级的基因检测也有一定的实用价值，目前来说，消费级基因测序在产前无创 DNA 测序的应用中最为成熟。使用产前 DNA 测序，不仅可以对胎儿进行常规遗传疾病的诊断，如唐氏综合征等，甚至可以预测未来的身体特征及智力等私密的个人信息。这项技术也对公众的道德提出挑战：是否应该在婴儿并没有选择权的情况下帮助婴儿选择 DNA 测序，甚至因此改变婴儿的命运。2013 年这项技术被《麻省理工科技评论》评为年度全球十大突破性技术，却引发了巨大的伦理争议^[14]。

产前诊断是指在怀孕中，为侦查生育缺陷，如神经管缺陷、染色体异常、遗传疾病等症状，而对胎儿进行的筛查或诊断性检查。产前诊断方法包括侵入性方法和非侵入性方法。目前侵入性方法仍旧是确诊唐氏综合征等染色体疾病的主要方式，并以羊膜腔穿刺最为常用，该种诊断方法存在一定的流产风险。

1997 年，香港地区的科学家卢煜明通过研究发现，孕妇的血液中含有数以万亿计的胎儿 DNA 碎片，其中来自胎儿的 DNA 高达 15%，这使得无创筛选成为可能。2010 年，卢煜明等人再次发表研究成果，证明孕妇体内游离的 DNA 片段可以用来重建胎儿的完整的基因组，并可用于无创性产前诊断。

只需要几毫升来自孕妇的外周血，研究者便可以收集其中胎儿的 DNA 碎片，并通过高通量测序技术结合生物信息分析，计算胎儿患染色体非整倍体疾病的风险。此技术能同时检测 21-三体、18-三体及 13-三体，还可发现其他染色体非整倍体及染色体缺失/重复。由于这种无创的筛查方式更加安全方便，一经提出便被广泛接受，成为有史以来普及最迅速的检测之一。

为了扩大自身在测序市场上的强大优势并布局临床应用，2013 年 Illumina 以 3.5 亿美元的价格成功收购了美国产前诊断公司 Verinata Health。Verinata 是一家位于加州的初创企业，其业务主要是利用孕妇的血液为她们提供产前诊断服务。但在此之前该公司几乎没有任何收入，而这项收购发生时距该技术面世仅有一年的时间，足以见得该技术的突破性以及广阔的应用前景。

但产前 DNA 测序技术一经提出，便面临着巨大的道德争议。完整的基因测序就像是打开了潘多拉的盒子，将一个还未谋面的胎儿的所有信息都展现在眼前，很难预判这种信息的获取会带来什么样的结果。就比如关于发色的信息，未来婴儿头发是直是弯，或是否会秃顶。显然，几乎没有人会因为这种情况终结一个胎儿的生命，但也并不意味着不会发生。

由此引发的人为「完美婴儿」的担忧并非多余，对于一些并不致残、致命的缺点，不同人的接受度不同。甚至对于一些目前定义为疾病的遗传缺陷，未来很有可能在婴儿发病前就已攻克，而在这之前胎儿并没有机会决定自己的命运。在未出生前，胎儿的所有优缺点都已被一览无余，与可以自主选择是否接收全基因组测序的成年人不同，胎儿无权选择是否接受基因检测，而讽刺的是，这些检查结果却可能决定他的一生。

同时无创产前基因检测也存在一定的技术局限：目前不能检测易位导致的染色体异常；不能检测基因遗传疾病；不能预测晚期妊娠并发症等。随着国内二孩政策的展开，越来越多的高龄产妇选择产前 DNA 检测为胎儿的健康保驾护航，但值得注意的是，无创 DNA 产前筛查属于筛查范畴，而羊水穿刺是诊断范畴，尽管二者具有相似的准确度，但两者有本质的区别，前者无法代替后者。

同时，过度强调基因的重要性会使人们远离探索自我的初衷，对基因风险的过度解读会使公众只关注可能带来的弊端，因而增加无谓的恐惧和焦虑等情绪。很显然，可能患病并不等同于一定患病，但很多人会将这种可能性当作确定性对待。

DNA 测序似乎进入了一个困局，它存在的价值是因为可以「预告」人生，但一旦预先知道了这样的可能，人生还会继续吗？研究者不禁陷入了迷茫，难道就因为公众对于可能性的误读就要因噎废食，将如此价格低廉又高效的方法束之高阁吗？

当然不会，既然这种可能性说起来虚无缥缈，那为什么我们不将其量化、精准化呢？基于这一想法，研究者提出了基于基因的风险评估体系，也被称为基因占卜，2018 年被《麻省理工科技评论》评为年度十大突破性技术之一^[15]。同时，2018 年也曾被预测为该技术的爆发元年。

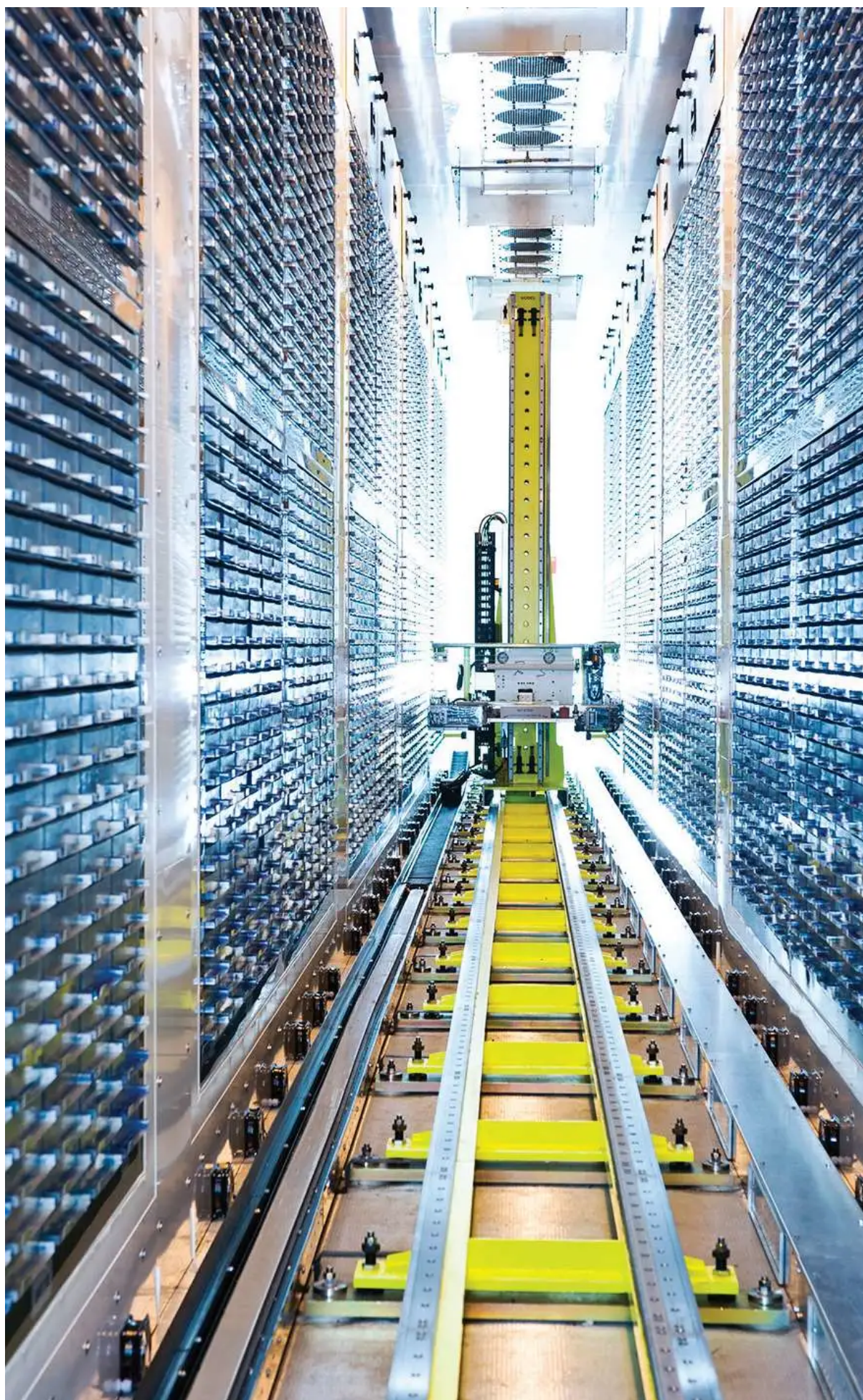
基因占卜是从海量的遗传数据中捕捉细微的遗传差别所带来的致病可能性。对于单个基因突变引起的遗传疾病，如镰状细胞性贫血（由单碱基突变导致），毫无疑问，一次简单的基因测序便可以确诊疾病。但是像这种突变和疾病之间一一对应的情况，并不存在于大多数疾病中，相反，大多数疾病的复杂诱因直到今天仍旧是未解之谜。

人类从不孤军奋战，而身体中的器官、细胞，甚至基因也是一样。对于人体中一种特性的控制，常常有多个基因共同参与，而对于不同人，每个基因又有着成千上万种不同的编码可能。来自得克萨斯大学的心脏病专家 Amit Khera 正在架构一个叫作「多基因评分」的体系。所谓「多」，指的是他的运算不只涉及单个基因，而是涵盖了数千个基因。

而构建这样的平台，首先需要有一个浩大的数据库作为支撑，例如最近的一次研究中，研究者为了探究失眠的原因，使用了 131 万余人的遗传信息，创下了纪录。在英国生物样本库（UK Biobank）里涵盖了约 50 万英国人的 DNA 和医疗记录，遗传学家通过对此类的大数据库进行深挖，解密这些人的基因组信息与疾病、个性甚至习惯之间的联系。

通过海量的研究信息捕捉细微的遗传变化带来的影响，听起来似乎有些矛盾，但如果将基因的评分比作一本账簿就好理解多了。为了建立多基因评分平台，凯拉使用了个体人类基因组中的 660 万个位置信息。每个位点都有一个 DNA 碱基，可能在我的体内是 G 而你的体内则是 A。在以往的研究中，遗传的随意性加大了

理解疾病的难度，但基于遗传大数据，凯拉现在可以查找在某特定位点是 G 的人，患心脏病的风险会有什么变化，也许它只会让风险增加 0.1%；其他位置的 G 也有可能让风险减少 0.2%。这些影响微不足道，但微小的遗传影响合计起来便会产生不容忽视的后果。



英国生物样本库中冻存着由 50 万名英国志愿者捐赠的 DNA 样品，轨道上的机器人正在此搜索样品

「在我所想象的未来，你可以在很小的时候就获得一份（身体状况）报告，」Khera 说道，「上面将提供你在 10 种疾病上的评分，比如你将有 90% 的可能性患有心脏病，50% 的可能性患有乳腺癌，至少 10% 的风险患有糖尿病。」

这种评分对普通人来说或许只是一种可能性，但对于一些制药公司的药物研发及临床试验却很有价值。相比于其他基于表观的检测方法，DNA 检测具有独一无二的预见性，并可以随时检测，同时结果并不会随年龄或环境发生变化。2017 年，来自加州大学圣地亚哥分校的脑科学研究者 Anders Dale 称，他计划推出一款阿尔茨海默症风险计算器。该风险计算器可以对用户是否会患病及何时患病进行预估。

尽管该项服务还未启动发行，但几家药物公司早已开始与其接洽。目前在阿尔茨海默症药物研发方面各大制药公司已经投入数十亿美元，但纷纷折戟。其中一个重要的原因就是已知的受试者均已发病，目前并没有办法得知未来谁会患病，因而也很难得知预防性药物是否真的有效。如果制药公司可以在阿尔茨海默症高危人群中试用药物，一切就变得简单多了。未来的药品将有可能贴有这样的标注：「建议多基因风险评分达 90 分及以上人群使用」。

然而，评分系统的这个概念并非新创，而是越来越丰富的数据使得它变得更加精准、更加完善。来自 Helix 的高级科学家 Sharon Briggs 猜测，未来几年内，来自 Helix 的 DNA 应用商店中的绝大多数产品都将采用风险评估体系。

不仅如此，拥有如此庞大的数据信息，研究者甚至可以反追疾病背后的「犯罪团伙」。10 年前，人类基因组计划完成之初，医学研究人员发起了首个现代基因搜索，他们仍旧希望通过几个主要的遗传元凶来解释糖尿病等常见病的病因。

「我猜测，糖尿病将和 12 个基因有关，而这些基因的真身将在两年内陆续揭晓。」现任美国国家卫生研究院院长的 Francis Collins 于 2006 年自信地宣布，他也是人类基因组测序的领军人物之一。

时间如期而至，但结果却不尽如人意。2009 年，柯林斯和其他的遗传学家不得不郁闷地开始谈论「缺失的遗传可能性」。那么导致疾病的基因究竟在哪里？答案是，无处不在。到 2014 年，研究者终于有了足够的遗传信息库去证明这个结果。随着参与糖尿病相关基因搜索研究的人数从 661 人变为 1 万人，最终升至 8 万多人，背后的一个个答案也蜂拥而至。现在我们已经知道在人类 DNA 中超过 400 个位点与 2 型糖尿病相关，每一个影响都微小且难以检测，而并非之前预测的 12 个基因。

不难看出无论是基因占卜还是确定疾病的幕后黑手，其命门都是数据库。那么海量的数据信息从何而来，又是谁手中握有如此大量的遗传数据呢？

医疗科研机构自然是其中之一，但最大的主力你也许并未想到——消费级测序公司。2018 年，全球共有 1,200 万人完成了基因测序，而通过 23andMe 和 Ancestry.com 公司的服务获取了遗传信息及家族谱树的消费者就有 1,000 万之多。以基因测序公司 23andme 为例，消费者只需要支付 99 美元（折扣时低至 69 美元）就可以获得自己的祖源信息，如果支付更多，那么消费者将会在原始数据的基础上获得多项解读。但测序、解读的费用并不是 23andme 从消费者身上获取的全部收益，制药公司也喜欢为这部分信息买单。单就 2015 年年初卖给 Genentech 用于帕金森研究的数据库一笔生意，23andme 就赚了 6,000 万美元。

一方出售，一方购买，看起来再平常不过。但对于普罗大众来说，使用或获取数据库信息时却远没有这么便捷。如今互联网已经让地球成为地球村，人们已经习惯了从互联网获取、分享信息，但基因信息的交流仍旧闭塞。在测序技术平民化的今天，无法获取大量信息的原因并不是技术限制，而是研究人员不愿意分享，由于隐私方面的法律法规，贸然将人们的基因组信息上传到互联网有可能触犯法律。

而全球基因组学与健康联盟（Global Alliance for Genomics and Health，GA4GH）要做的正是帮助人们打破这最后的限制，通过建立让 DNA 数据库互通的技术标准和平台，尽可能地连通每一个基因组信息，将基因测序带来的健康红利最大化。根据这一设想，研究者最终将构建一个包含数百万基因组的全球网络，即 DNA 的互联网。

这一想法的提出首次将遗传数据的应用从个体诊断推及到整个群体的参与，而这些共享的数据也必将给一些罕见病及未知疾病的诊疗带来深远的影响，因而 2015 年 DNA 互联网这一技术被《麻省理工科技评论》评为年度十大突破性技术 [16]。

如今看来，这是时代所需的必然产物，但在基因测序方兴未艾，尤其是个人消费级基因测序刚刚起步时提出这一构想，是相当具有前瞻性的。

全球基因组学与健康联盟成立于 2013 年，是一个由医疗机构、研究型大学和生命科学公司等组成的非营利性组织，最初由 50 名来自 8 个国家的科学家共同建

立，旨在搭建一个共享基因及临床数据的平台。建立之初，该组织将工作重点集中在基因组数据、临床数据、安全和隐私、监管及伦理四个方面。

早在 2014 年，该组织就发布了名为「灯塔」（Beacon）的基础 DNA 搜索引擎，他们将数据库比喻成「灯塔」，每一个愿意加入的团体或个人只需要安装一个服务器程序，而这个安装的过程则被称为「点亮灯塔」。系统中的用户可以对其他数据库进行简单的问询，而其他灯塔可以回复「是」或「否」的答案。在这个过程中数据库的数据并没有被移动，因而保证了遗传信息的相对安全与私密。虽然这个过程很简单，但首次连通了孤立的数据库，是遗传信息交流的一个重要开端。

随着近几年基因测序领域井喷式的发展，海量的遗传信息正在涌现。2017 年 10 月，全球基因组学与健康联盟公布了名为「GA4GH Connect」的未来五年战略规划，呼吁联盟的 500 多名成员开发新的数据共享标准，用于当前重大的国际基因组学数据计划。

一群人想着从科研机构入手，而另一群人则是直接瞄准了每个基因组的拥有人——普通消费者。遗传信息与普通可销毁的信息不同，一旦被外界获取，整个人就如同「裸体」一般示于人前，家庭关系、身体特征、缺陷等都被一览无余。很多人出于隐私或花费的考虑，并不愿意接受基因测序或共享遗传信息。而今，由哈佛大学遗传学专家 George Church 牵头创立的公司 Nebula Genomics 就很好地均衡了信息共享与隐私问题，他们承诺可以以低于 1,000 美元的价格帮助顾客完成基因组测序，并通过区块链保障信息的安全，甚至消费者可以亲自参与到自身遗传信息的决策中，依照自身的意愿通过区块链决定储存或交易遗传信息 [17]。

「这是一种应对基因组学挑战的新方法，这个方法综合考量了测序花费、遗传信息保护、数据管理以及基因组大数据的处理，基于区块链技术，让更多人真正地『拥有』自己的遗传信息。」丘奇说道。

区块链具有去中心化、公开安全及难以篡改的特点，以区块链为媒介完成消费者遗传信息的保存及分享，可以使普通人直接面向制药公司，选择是否交易个人遗传数据。在交易中，购买方的身份要求绝对透明，而数据提供者则可以保持匿名状态。事实上，生物技术公司和医药公司并不能下载消费者分享的数据，而只是「借用」部分信息。

这一构想不仅可以保障遗传信息所有者的隐私安全，同时可以将遗传数据转化成加密货币，实现真正的价值。但随着 DNA 测序及解读的飞速发展，未来很可能会因此滋生一些怪相，比如根据智商预测为孩子选择学校，或根据身体条件评估决定是否录用职员这类荒唐事，更严重的甚至可能引发由于遗传信息决定的社会分级。残酷的是，你的分级并非因为测序，而是从出生的那一刻就已经决定了，并无法更改。

但这些由于发展过快带来的脱节最终都将会被弥补，并且可以预见，未来科研相关的基因检测仍将是核心与重点，而消费级的基因检测将会成为热点。值得注意的是，尽管目前消费级基因检测的市场看起来已经繁花似锦，但这些也只是序曲，消费级基因检测的时代还远未到来，随着测序成本的降低、解读应用的完善以及个人信息安全保护机制的实现，未来 5 年内，消费级基因检测必将进入一个新的舞台。

评论



添加评论...