

海明威的猫

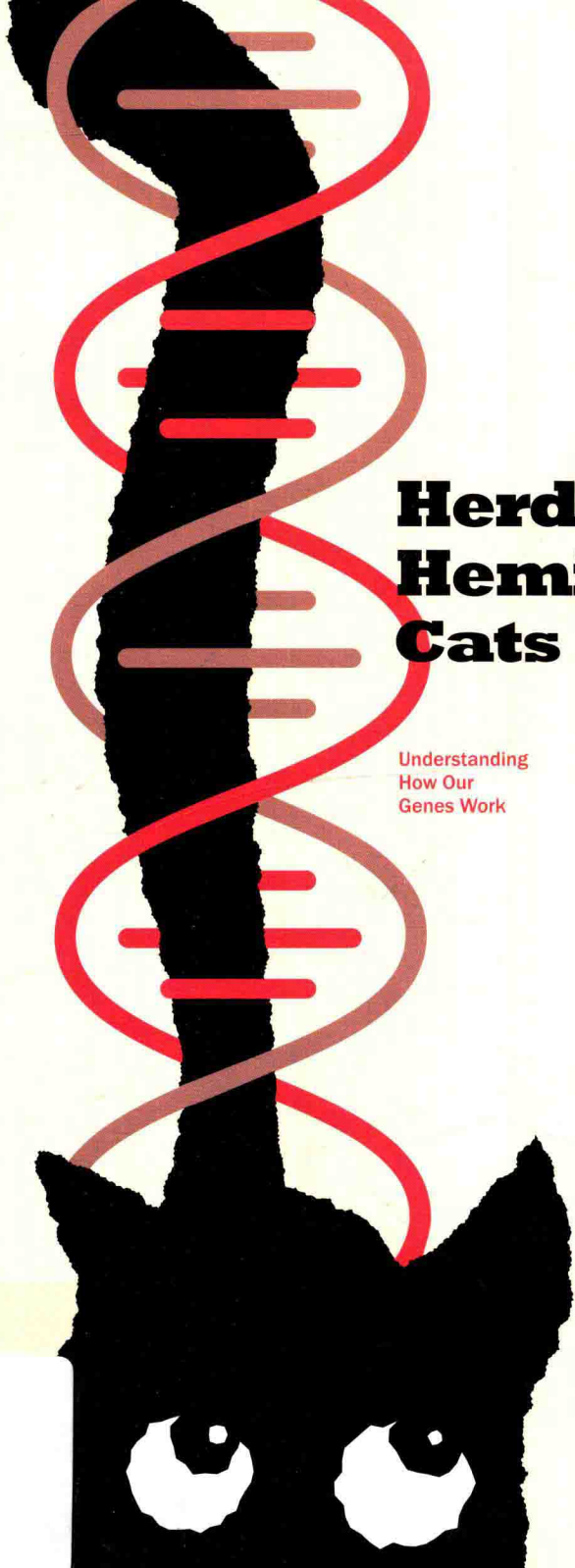
Herding Hemingway's Cats

Understanding
How Our
Genes Work

基因是
如何运作的

「英」凯特·阿尼 (Kat Arney) —— 著
葛文聪 —— 译

中信出版集团



**Herding
Hemingway's
Cats**

Understanding
How Our
Genes Work

海明威的猫
基因是如何运作的

「英」凯特·阿尼 (Kat Arney) —— 著

葛文聪 —— 译

中信出版集团 · 北京



图书在版编目 (CIP) 数据

海明威的猫：基因是如何运作的 / (英) 凯特·阿尼著；葛文聪译. -- 北京：中信出版社，2017.12

书名原文：Herding Hemingway's Cats:

Understanding How Our Genes Work

ISBN 978-7-5086-7793-4

I. ①海… II. ①凯… ②葛… III. ①基因—普及读物 IV. ①Q343.1-49

中国版本图书馆 CIP 数据核字 (2017) 第 146892 号

HERDING HEMINGWAY'S CATS: UNDERSTANDING
HOW OUR GENES WORK

By

KAT ARNEY

Copyright: © KAT ARNEY, 2016

This edition arranged with BLOOMSBURY PUBLISHING PLC
through Big Apple Agency, Inc., Labuan, Malaysia.

Simplified Chinese edition copyright:

2017 Beijing Paper Jump Cultural Development Company Ltd

All rights reserved.

本书仅限中国大陆地区发行销售

海明威的猫：基因是如何运作的

著 者：[英] 凯特·阿尼

译 者：葛文聪

出版发行：中信出版集团股份有限公司

(北京市朝阳区惠新东街甲 4 号富盛大厦 2 座 邮编 100029)

承印者：北京通州皇家印刷厂

开 本：880mm×1230mm 1/32

印 张：10.75 字 数：230 千字

版 次：2017 年 12 月第 1 版

印 次：2017 年 12 月第 1 次印刷

京权图字：01-2017-6074

广告经营许可证：京朝工商广字第 8087 号

书 号：ISBN 978-7-5086-7793-4

定 价：48.00 元

版权所有·侵权必究

如有印刷、装订问题，本公司负责调换。

服务热线：400-600-8099

投稿邮箱：author@citicpub.com

献给挚爱的爸爸、妈妈

感谢你们的养育



海 明 威 的 猫

Herding Hemingway's Cats:
Understanding How Our Genes Work

序言

碱基决定一切

话要从一张猫的照片说起。有一次，我在伦敦的英国皇家学会（Royal Society）参加一个科学会议。当时，我坐在会场的后排。忽然，一只爪子格外大、样子很可爱的猫吸引了我的视线。“这是一只海明威猫。”演讲者一边说着，一边指着他背后的大屏幕上显示的动物，“这种猫有6趾，是多指畸形动物。据说，欧内斯特·海明威喜欢这种猫，直到今天，它们依然生活在佛罗里达州的海明威故居里。再看这里……”他按了一下电脑，换了一张幻灯片，这张幻灯片上尽是畸形人手的照片，“……这些是畸形六指患儿的手。导致这类畸形的是同一种遗传错误。”

这些六趾猫和六指人的照片可能会很自然地让人以为这种畸形是单个基因错误所导致的。然而，事实并非如此。导致这种畸形的实际上是DNA中的一段错误区域，这个区域发挥着调控开关的功能，在人胚胎或猫胚胎的发育过程中，它通常会在合适的时间、合

适的位置开启 1 个基因，指导人指或猫趾的形成。此外，从分子层面看，这个调控开关与其所调控的基因相距甚远。了解了海明威猫及其被损害的调控开关，我便思考下面的问题：基因是如何发挥作用的？作为一名科普作家和科学节目主持人，我怎样才能通过我的工作向大众解释基因发挥作用的原理？



我与现代遗传学的第一次真正接触是在我的中学时代，这还得感谢那位令人生畏的迈尔斯副校长。他常常怒气冲冲地穿过学校的走廊，冷冰冰的目光里透着苛刻和不满，他还随意扣留学生。除了担任副校长，他还兼任我们的生物老师。一天，轮到他主持学校的例会。我们老老实实成群结队地走进会议大厅，盘腿坐在地板上，尽量避开他的目光。他穿着黑色学袍走到台上，长袍在身后像个斗篷一样飘动着，他手里紧紧攥着的好像是一本杂志，不过那肯定是某种科学期刊。他站在台上，情绪激动不已，愤愤地朝我们挥动那本杂志，仿佛那是从男士卫生间的水箱后面搜出的一本色情读物。“看看这个吧！”他一边大声地说着，一边拍打着布满了字母 A、C、T、G 的一页纸，那些字母好像在无限地交替排列着。“这就像电话簿一样！全都是字母。字母，字母，字母。”他停顿了一下，喘了口气接着说，“这就是现在的生物学。”

迈尔斯先生当时可能以为 ACTG 是一个由 4 个字母组成的词。然而近年来，基因领域的术语已经颇为常见。基因就是你的 DNA 中决定你的眼睛蓝、肚子大或者头发卷的那些成分。报纸上说，基因控制着我们患癌症、心脏病、酗酒及老年痴呆症等疾病的风险。花 1000 美元就可以在数日内测出你自己的基因组序列，测序结果会被妥善地存储到一个优盘里交给你。遗传知识有力量拯救我们——有些药物是针对肿瘤中的错误基因进行治疗的，基因治疗领域的一些新近的研究成果能够使盲人重见光明。我们可以在地球上看似无尽的生命变体中追溯显而易见的基因发展轨迹——归根结底，进化只是遗传物质随时间发展的结果。

今天，我们愉快地谈论着“在我们的基因中”的一切，从歌唱天赋到危及生命的癌症都是我们谈论的话题。我的母亲对家族史非常痴迷，认为我或者我的姐妹在性格上的方方面面几乎都能追溯到我们早已作古的祖先那里，他们有可能是虔诚的传教士，也有可能是郁郁寡欢的酒鬼。（这两种性格特质我竟然都具备。谢谢你，妈妈。）科普作家兴致盎然地描述基因组——一个生物体所携带的一套完整的 DNA 序列——将其比喻为生命的蓝图，类似于计算机的程序或建筑师的设计图。DNA 的双螺旋结构已经成为一种文化符号，甚至连一些广告也将其用作科学的代名词。虽然遗传学术语已经渗入到公众的意识中，但是真正了解基因及其功能的人却少之又少。

大多数生物学教材把基因定义为特定的 DNA “字母”链——叫作碱基的化学物质，所有的生命体都包含 4 种碱基：A（adenine，腺嘌呤）、C（cytosine，胞嘧啶）、T（thymine，胸腺嘧啶）和 G（guanine，鸟嘌呤）。这些字母的具体排列顺序编码成为指令，指

导身体里的细胞制造出各种各样的分子，这就像不同的食谱指导人做蛋糕、馅饼或者炖菜一样。

不幸的是，我们不能简单地把我们的基因组看成具有生命的张三或者李四。随着技术的发展，科学家们发现了更多隐藏在我们的细胞内部的秘密，基因开始变得深不可测。在你的身体里，几乎每一个细胞中都包裹着 2 米多长的 DNA 长链，这个长链上布满了成千上万的基因，而这些基因需要在合适的时间、合适的位置被开启或关闭。根据我们现在的了解，基因组可不像一套装订整齐的食谱，而是像一个动态的螺旋状的图书馆，里面有很多生物学读者和作者在忙忙碌碌。书上的内容被不停地复制，不停地修改，有时甚至会被彻底损坏。每一本书上都写满了注释，还有成千上万页的内容看上去简直就是废话。面对这个图书馆的编目系统，即使是世界上最能干的图书管理员也会感到崩溃。然而，就是在这样的混沌中，我们创造着生命。

最初，当一个幸运的精子进入一个卵子内，基因活动所演绎的复杂的生物芭蕾舞便创造出生命，使我们从一个细胞发育成一个婴儿。复杂的结构由此形成了：弯弯曲曲的血管，布满褶皱的大脑，形似鸟儿的小小骨架。在这个过程中，每一种细胞都发挥了自身的专长。皮肤细胞发展成结实而平整的皮肤，把肉体包裹起来；脑细胞之间通过微小的突触传递脑电波。这一切都是由基因来控制的。然而基因是如何做到这些的？基因究竟是怎样发挥功能的？

为了寻找答案，我以面谈和电话沟通的形式访问了近年来世界各地工作在遗传学前沿的一些研究者。我的访问名单并不是通过某

种特别高明的方法来筛选制定的。在这些受访者中，有一些是我以前做科研工作时就认识的，有一些是别人推荐给我的，还有一些是因为他们发表的言论或著作引起了我的兴趣，于是我决定去会一会他们。他们的故事并不会都在本书中呈现，但每一个与我交谈过的人都以某种方式影响了我的思想。我很感谢他们花时间慷慨地与我分享了他们的学术成果，还陪我喝了咖啡。真是喝了很多咖啡。

除了与每个人谈论他们的研究之外，在每次采访结束时，我都会提出同一个问题：“你觉得什么是怪异的？”科学的历史告诉我们，今日无法解释的怪异现象是明日革命性的突破。在梳理眼前知识的同时，我也想展望一下未来的路径。当我开始提笔写这本书的时候，我以为自己知之甚广。毕竟我花了许多年来研究遗传学，在实验室里工作，潜心研究基因转录和翻译的过程。然而，当我开始与这些人交谈，阅读他们的研究成果，继而思考这些内容，我先前的想法就像任性的小猫一样飞身离我而去。出乎意料的是，关于我们的基因，可以确定的内容远远少于许多人预想的数量，而未解之谜的数量则大大超出了许多人的想象。

在此书中，我尝试把我从各处汲取的理念汇集在一起。这绝非一本艰深的遗传学教科书，相反，这本书中充满了故事，讲述你、你的基因以及基因组是如何发挥作用创造出生命的传奇。希望你会喜欢这本书。



有几个问题需要注意：在书中的不同语境中，我使用了“读取”和“转录”两个词。前者是指科学家获取 DNA 中的字母（碱基）顺序的行为——这项技术被称为“DNA 测序”；后者是指细胞内部的运转，当一个基因被开启，这个基因上的信息被复制，形成一个叫作 RNA 的信息，这个过程被称作“转录”，我在文中会讲述基因被读取或转录生成 RNA 的过程。希望大家能够清晰地分辨不同的语境所指向的内容。

我倾向于使用“字母”来描述 A、C、T、G 四种碱基——在 RNA 和 DNA 中的这 4 个碱基构成了生命的字母表。在 RNA 中，字母 T 被一个叫作 U 的非常相似的碱基所取代。为了简化表达，不论是表达 DNA 还是 RNA 中的这个碱基，我在全书中都使用 T。另外，我也避免使用基因术语的标准名称，防止满篇都是神秘费解的首字母缩略词。

如果你想了解更多，现在有足够多的优秀书籍更加深入地讨论生命的机制，我在本书的注释和参考书目中也提到了我最喜欢的一些书。在最后的的部分，你可以找到一些链接、研究论文、更深入的阅读材料以及实用的专业术语词汇表。

c 序言 / 碱基决定一切

001 第一章 / 贵不在有，而在于用

045 第二章 / 细说“垃圾”

029 第三章 / 浅谈法则

073 第四章 / 调控开关

051 第五章 / 组合的秘密

063 第六章 / 有拇指的猫

075 第七章 / 有脆的鱼

085 第八章 / 老鼠、人类和地鼠，噢，天啊！

099 第九章 / 派对之城

111 第十章 / 滋养我们的基因组

133 第十一章 / 剪切和粘贴

149 第十二章 / 大自然的红笔

161 第十三章 / 不断增长的环

171 第十四章 / 基因的沉寂

189 第十五章 / 复活之夜

197 第十六章 / 跳来跳去

209 第十七章 / 开启一罐多变的虫子

223 第十八章 / 每个人都是个突变体

239 第十九章 / 打开黑匣子

255 第二十章 / 错在父母

271 第二十一章 / 遇见米老鼠

289 第二十二章 / 寻找 21 世纪基因

303 词汇表

307 推荐阅读

326 致谢

第一章

贵不在有，而在于用

没有上帝，没有自由意志，只有 DNA。

卡里·史蒂芬森 (Kari Stefansson)

基因解码公司 (DeCODE Genetics) 创始人



埃里克·米斯卡（Eric Miska）有点儿沮丧。“这非常复杂。”他叹息道，“我们获得了这些数据，但是却无法有效地理解它们。以前，在大卫·莱特曼的电视节目中有一个‘探求真相’的环节，观众要投票判定某件事是否有趣。我们目前的处境就与之类似。”

他带我来到了剑桥大学格登研究所（Gurdon Institute in Cambridge），走进了他的办公室。他那高大的身躯落座在一把办公椅上，双脚翘到桌子上摆来摆去。我俩之间无须拘谨于礼数——埃里克和我好像回到了以前在格登研究所一起读研究生的同窗时光。现在，他是剑桥大学遗传学专业的赫歇尔·史密斯（Herchel Smith）讲席教授，而且就在今天，他刚刚发现自己登上了名人录。然而，即使这样的好消息也不能提振他那低落的情绪。他觉得自己已经衰老，性情乖戾。显然，他正在思考把一生的时光都投入到又苦又累的科学工作中是否值得。“我正在撰写一本书，讲的是基因是如何发挥作用的，”我兴奋地对他说，“我想，和你聊聊应该是件好事。”“听上去不错啊。”他用僵硬的德国口音回答道，“等你找到了答案，就告诉我吧。”我心想，别这

样，埃里克，至少你可以试着表现出热情的语气啊。在他身后有一个大大的鱼缸，里面有很多观赏鱼吐着气泡游来游去。12月的阳光渐渐在窗上褪去，昏暗的办公室里，一盏灯投下一束奇异的光晕，笼罩着我们的对话。

埃里克的情绪低落是可以理解的。由于DNA测序技术的成本大大降低，而效率大大提高，世界各地的实验室里都安装了带有很多指示灯的巨型机器，它们大量地输出数不尽的遗传密码。20世纪90年代，我们俩还是学生，那时DNA序列只能一小段一小段地读取，每次只能读取几百个字母。读取一段完整的DNA序列需要花几个月的时间费力地把一小段一小段信息顺次拼凑起来，形成完整的序列信息。今天，读取一套完整的人类基因组图谱在几天之内就可以完成。然而，这并没有使我们的基因内容变得更加清晰，相反，情况实际上变得非常复杂。

问题在于，我们并非是戴着厚厚的眼镜、留着络腮胡子、具有超能力的天外来客所设计创造出来的。我们的基因组是在数千万年的进化过程中逐渐被拼凑在一起的，而且由于缺乏更富有诗意的拼凑方法，我们的基因组里面充满了垃圾。



1977年是令人瞩目的一年，至少在我眼里是这样，因为这一

年发生了遗传学历史上的三件大事。^[1]第一，电影《星球大战》上映了。这部影片让我们了解了许多遗传学领域的关键概念，如先天遗传与后天环境之间的博弈、近亲繁殖的种种风险以及克隆技术的潜力。第二，性手枪乐队推出了专辑，这张专辑印证了生物学领域的一个重要理念：你无须完美，只需能够跨越障碍创造出生命力即可。第三，DNA 分析技术诞生了，科学家们得以借此“阅读生命之书”，解码人类基因中的信息。

20 世纪 80 年代早期，在消费电子领域，Betamax 录像系统和 VHS 录像系统展开了角逐。同一时期，两种不同的 DNA 读取方法也在争夺主导地位。最后，由英国科学家弗雷德·桑格（Fred Sanger）开发的更简易的技术战胜了由美国研究者开发的技术。弗雷德的方法在其后的 30 年里对遗传学产生了革命性的影响，直到最近，这种方法才被现代的“新一代”技术所取代——高速、高分辨率的“蓝光光盘”取代了弗雷德厚重的“VHS 磁带”。

大多数人都知道 DNA 是双螺旋结构——这个看似旋转梯子的结构是詹姆斯·沃森（James Watson）和弗朗西斯·克里克（Francis Crick）在 1953 年发现的，在研究过程中，他们还得到了罗莎琳德·富兰克林（Rosalind Franklin）和莫里斯·威尔金斯（Maurice Wilkins）的大力帮助。这个梯子的两侧是由糖分子紧密结合形成的长链，中间的一级级阶梯是由叫作碱基的 4 个化学“字母”构成的，这 4 种碱基通常用其首字母表示，即 A（腺嘌呤）、G（鸟嘌呤）、C（胞嘧啶）和 T（胸腺嘧啶）。它们是构成生命的字

[1] 我也是在这一年出生的，这是否算得上是遗传学历史上的一件大事尚未可知。