nature

The Living Record of Science

《自然》百年科学经典

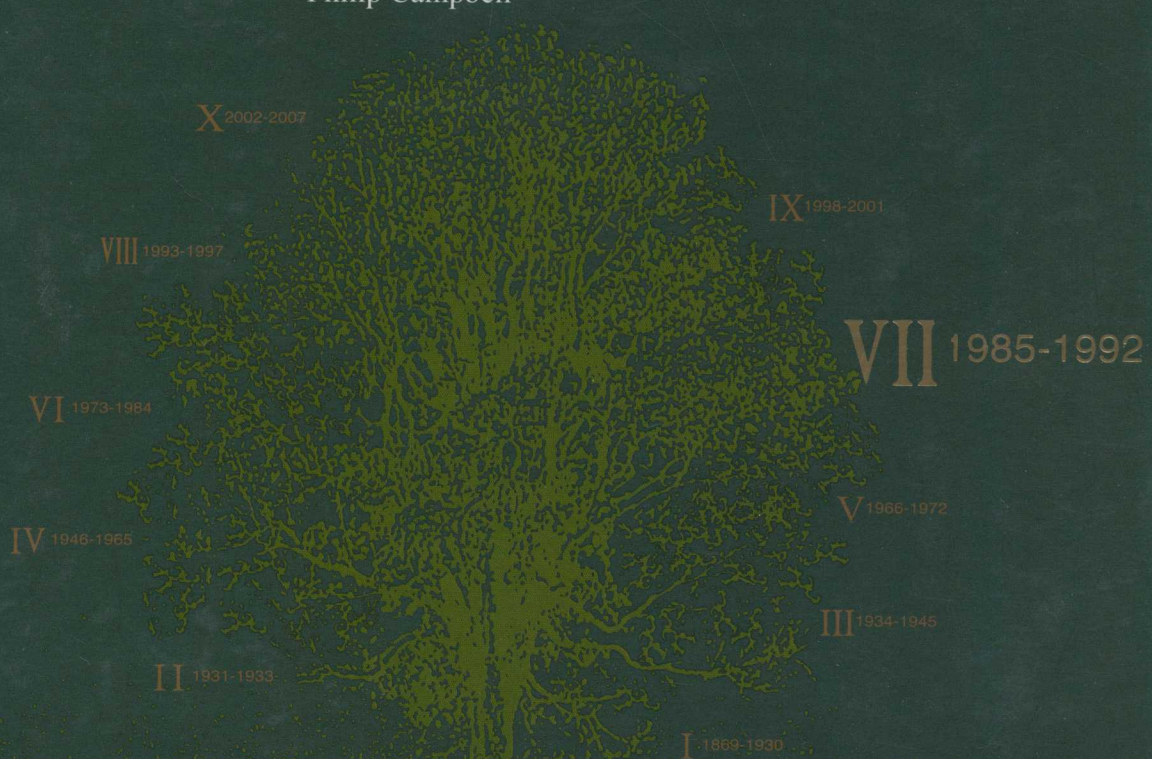
（英汉对照版）

第七卷

总顾问：李政道（Tsung-Dao Lee）

英方主编：Sir John Maddox
Philip Campbell

中方主编：路甬祥



外语教学与研究出版社 · 麦克米伦出版集团 · 自然出版集团

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS · MACMILLAN PUBLISHERS LTD · NATURE PUBLISHING GROUP

nature

The Living Record of Science

《自然》百年科学经典

（英汉对照版）

第七卷

总顾问：李政道（Tsung-Dao Lee）

英方主编：Sir John Maddox
Philip Campbell

中方主编：路甬祥

VII

1985-1992

外语教学与研究出版社 · 麦克米伦出版集团 · 自然出版集团

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS · MACMILLAN PUBLISHERS LTD · NATURE PUBLISHING GROUP

北京 BEIJING

图书在版编目(CIP)数据

《自然》百年科学经典. 第7卷, 1985～1992：英汉对照 /（英）马多克斯（Maddox, J.）,（英）坎贝尔（Campbell, P.）, 路甬祥主编. — 北京：外语教学与研究出版社, 2015.4
ISBN 978-7-5135-5920-1

Ⅰ. ①自… Ⅱ. ①马… ②坎… ③路… Ⅲ. ①自然科学—文集—英、汉 Ⅳ. ①N53

中国版本图书馆CIP数据核字（2015）第079511号

出 版 人：蔡剑峰
项目统筹：章思英 Charlotte Liu（加拿大）
项目负责：刘晓楠 黄小斌 Mary-Jane Newton（德）
责任编辑：黄小斌 王帅帅 张梦璇
装帧设计：孙莉明
出版发行：外语教学与研究出版社
社 址：北京市西三环北路19号（100089）
网 址：http://www.fltrp.com
印 刷：北京华联印刷有限公司
开 本：787×1092 1/16
印 张：63
版 次：2015年5月第1版 2015年5月第1次印刷
书 号：ISBN 978-7-5135-5920-1
定 价：568.00元

* * *

购书咨询：(010)88819929 电子邮箱：club@fltrp.com
外研书店：http://www.fltrpstore.com
凡印刷、装订质量问题，请联系我社印制部
联系电话：(010)61207896 电子邮箱：zhijian@fltrp.com
凡侵权、盗版书籍线索，请联系我社法律事务部
举报电话：(010)88817519 电子邮箱：banquan@fltrp.com
法律顾问：立方律师事务所 刘旭东律师
中咨律师事务所 殷 斌律师
物料号：259200001

《自然》百年科学经典（英汉对照版）

编译委员会

本卷翻译工作组稿人（以姓氏笔画为序）

王晓蕾　王耀杨　刘　明　刘晓楠　关秀清　李　琦　何　铭
郭红锋　蔡　迪　蔡则怡　王帅帅　黄小斌

本卷翻译人员（以姓氏笔画为序）

马　颖　王耀杨　牛慧冲　刘皓芳　齐红艳　许向科　李　梅
李　琦　肖　莉　张锦彬　张新彦　陈　林　金世超　周平博
周志华　周晓明　郑建全　侯彦婕　姜　薇　蔡则怡

本卷校对人员（以姓氏笔画为序）

丁　然　于同旭　马　昊　马　亮　王　敏　王　雷　王晓蕾
王海纳　元旭津　公　晗　邓铭瑞　石宇鹏　石爱洁　史未卿
史明澍　闫振丽　任峪铭　刘俊杰　许晨雨　苏怡汀　李　芳
李　妍　李　昆　李　萌　李　梅　李　静　李永新　李兆升
李盎然　朱　玥　杨　茜　吴兆军　张　敏　张　晴　张香香
张欣园　张琦玮　张瑶楠　陈　雄　罗小青　赵晓非　郑　征
宗伟凯　侯鉴璇　钱　磊　徐　玲　徐　悦　高坛光　郭晓博
夏　秋　黄　瑢　黄　璞　黄雪嫚　梁　瑜　葛聆泓　谢周丽
雷文欣　魏洪淼　Eric Leher（澳）

Contents
目录

Volume VII

(1985-1992)

Complete Nucleotide Sequence of the AIDS Virus, HTLV-III

L. Ratner *et al.*

Editor's Note

The early 1980s was a period of intense AIDS-related research and discovery. The AIDS-causing virus, then called human T-cell leukaemia virus III (HTLV-III), was identified and isolated, as were its modes of transmission. Here US biologist Robert Gallo and colleagues describe the complete nucleotide sequences of the AIDS virus, a milestone in AIDS research. Overall, the sequence resembles that of other RNA-encoded "retroviruses" containing the three hallmark genes *gag*, *pol* and *env*. But it also contains anomalies, not least the immense heterogeneity between clones. The genome has enabled researchers to define key viral genes and proteins, has shed light on the origins, nature and spread of HIV, and continues to influence diagnostic and drug development.

The complete nucleotide sequence of two human T-cell leukaemia type III (HTLV-III) proviral DNAs each have four long open reading frames, the first two corresponding to the *gag* and *pol* genes. The fourth open reading frame encodes two functional polypeptides, a large precursor of the major envelope glycoprotein and a smaller protein derived from the 3′-terminus long open reading frame analogous to the long open reading frame (*lor*) product of HTLV-I and -II.

HUMAN T-cell leukaemia (lymphotropic) viruses HTLV-I, -II and -III, a family of exogenous retroviruses, are associated with T-cell disorders including adult T-cell leukaemia lymphoma (ATLL) and the acquired immune deficiency syndrome (AIDS)[1-14]. These viruses share a number of biological and structural features: tropism for $OKT4^+$ lymphocytes[10,15-21], the ability to produce giant multinucleated cells in culture[10,17,22-24], immunological cross-reactivity of some virally encoded proteins[12,13], distant nucleic acid sequence similarities[8,9,25-28] and the preference for magnesium of the viral reverse transcriptase[29-31]. Moreover, the genome of HTLV-I and -II as well as the related bovine leukaemia virus (BLV) is somewhat longer than that of other retroviruses[8,9,25,32-34]. A sequence of 1,600–1,800 nucleotides is interposed between the 3′ end of the *env* gene and the long terminal repeat (LTR) sequences of these viruses[25,32-37], the 3′ portion of which is a long open reading frame (*lor*)[32,35,37]. Another feature that distinguishes HTLV-I, -II, -III and BLV from other retroviruses is the marked increase in the rate of transcription initiated within the viral LTR sequences in infected compared with uninfected cells[38,39]. This phenomenon, *trans*-acting transcriptional regulation, is not observed for other retroviruses[40] and is probably mediated by the *lor* gene product of these retroviruses[38].

艾滋病病毒HTLV-III的完整核苷酸序列

拉特纳等

编者按

20世纪80年代早期是艾滋病相关研究和发现的密集时期。人们鉴定并分离出了引发艾滋病的病毒——后被称为T细胞白血病病毒III（HTLV-III）并弄清了它的传播模式。在本文中，美国生物学家罗伯特·加洛及其同事描述了艾滋病病毒的完整核苷酸序列，这是艾滋病研究领域的一块里程碑。这个序列和其他RNA编码的“逆转录病毒”一样，都具有*gag*、*pol*和*env*这三个标志基因。但它也有一些不同，这些不同不仅仅局限于克隆间巨大的遗传异质性。基因组使得研究人员能够定义关键的病毒基因和蛋白，解释艾滋病的起源、性质和传播，进而影响诊断和药物的发展。

人类两种T细胞白血病III（HTLV-III）的前病毒DNA的完整序列各有四个长的开放阅读框，前两个与*gag*和*pol*基因相关。第四个开放阅读框编码两个功能多肽，主要包膜糖蛋白的大前体和来源于3′末端长开放阅读框的小蛋白，类似于HTLV-I和-II长开放阅读框（*lor*）的产物。

人类T细胞白血病（淋巴性）病毒HTLV-I,-II和-III,属于外源逆转录病毒家族，它们与T细胞紊乱有关，包括成人T淋巴细胞白血病（ATLL）和获得性免疫综合征（AIDS）[1-14]。这些病毒具有许多共同的生物学和结构特征：对$OKT4^+$淋巴细胞的嗜性[10,15-21]、在培养基中产生巨型多核细胞的能力[10,17,22-24]、与一些病毒编码蛋白的免疫学交叉反应[12,13]、远亲核苷酸序列之间的相似性[8,9,25-28]和病毒逆转录酶对镁的偏好[29-31]。此外，HTLV-I和-II与相关的牛白血病病毒（BLV）一样比其他的逆转录病毒的基因组都略长[8,9,25,32-34]。在*env*基因的3′末端和这些病毒的长末端重复序列（LTR）之间得到一段1,600~1,800个核苷酸的序列[25,32-37]，该序列3′部分是一个长的开放阅读框（*lor*）[32,35,37]。HTLV-I、-II、-III和BLV区别于其他逆转录病毒的另一个特点是，被感染的细胞中LTR序列区域起始转录的速率比未被感染的细胞有显著增长[38,39]。这种反式转录调控的现象在其他逆转录病毒中[40]没有观察到，可能是由这些逆转录病毒的*lor*基因的产物所介导的[38]。

Despite the similarity between HTLV-III and the other members of the HTLV-BLV family of viruses, the biology and pathology of HTLV-III differ substantially. Infection with HTLV-III results often in profound immunosuppression (AIDS), consequent on the depletion of the OKT4$^+$ cell population[10-14,41-43]. This effect is mirrored by a pronounced cytopathic rather than transforming effect of HTLV-III infection upon the OKT4$^+$ cells in lymphocyte cultures *in vitro*[10,11,20]. In contrast, infection with HTLV-I results in a low incidence of T-cell leukaemia lymphoma (an OKT4$^+$-cell malignancy)[1-6]. There is evidence also for some degree of immunodeficiency in HTLV-I patients[6,44]. Infection of primary lymphocytes in culture by HTLV-I and -II results in *in vitro* transformation of predominantly OKT4$^+$ cells[45,46]. A cytopathic effect of HTLV-I infection on lymphocytes is apparent, but the effect is not as pronounced as that in HTLV-III[16,17,45-48]. HTLV-III differs also from HTLV-I and -II in the extent of infectious virion production *in vivo* and *in vitro*. High titres of cell-free infectious virions can be obtained from AIDS patient semen and saliva and from the supernatant of cultures infected with HTLV-III[10,49-51]. Very few, if any, cell-free infectious virions can be recovered from ATLL patients or from cultures infected with HTLV-I or -II[52].

To investigate the biological activity of these viruses *in vitro* and *in vivo* and to provide information useful for the development of diagnostic and therapeutic reagents for AIDS, we have determined the complete nucleotide sequence of the HTLV-III provirus.

Genomic Structure of HTLV-III

Several closely related clones of HTLV-III DNA were obtained from the H9 cell line infected with HTLV-III present in the blood pooled from several American AIDS patients[10]. The complete primary nucleotide sequence of three unintegrated viral clones of 8.9, 5.3 and 3.6 kilobases (kb) in length[27] was determined with the partial sequence from an integrated proviral clone[53] (Fig. 1).

The HTLV-III provirus is 9,749 base pairs (bp) long. The overall structure of the provirus resembles that of other retroviruses. The sequences that encode viral proteins are flanked by LTR sequences. The LTR itself is flanked by inverted repeated sequences two nucleotides long (Fig. 1). Four long open reading frames are identified in the viral DNA (Fig. 2).

Long Terminal Repeat

A detailed analysis of the HTLV-III LTR is presented elsewhere[54]; it is 634 nucleotides long with U3, R and U5 regions of 453, 98 and 83 nucleotides, respectively. The boundaries of these regions of the LTR were defined by localization of the 5′-cap site by S_1 nuclease mapping and by measurement of the length of the strong stop DNA transcript, as well as by determination of the 3′-terminus of the viral RNA by sequence analysis of cDNA clones. A TATAA sequence typical of eukaryotic promoters, as well as the consensus sequence for polyadenylation, are indicated in Fig. 1. A transfer RNA binding site complementary to the 3′ end of $tRNA^{Lys}$ is located 3′ to the 5′ LTR. DNA sequence homologies to the LTR of HTLV-I[32], -II[55] and BLV[56] are indicated in Fig. 3.

尽管 HTLV-III 与 HTLV-BLV 病毒家族其他成员之间有相似性，但是 HTLV-III 与它们的生物学和病理学特征有着根本上的不同。HTLV-III 的感染常常导致完全的免疫抑制（AIDS），随后就是 $OKT4^+$ 细胞数减少 [10-14,41-43]。这一效应是通过 HTLV-III 感染体外淋巴细胞培养物 $OKT4^+$ 造成的细胞病变反映出来的，而非转染效应 [10,11,20]。相反地，细胞感染 HTLV-I 后仅导致一种低发病率的 T 淋巴细胞白血病（一种 $OKT4^+$ 细胞恶性肿瘤）[1-6]，而且也有 HTLV-I 病人存在不同程度的免疫缺陷的例子 [6,44]。HTLV-I 和 -II 体外感染导致培养的原代淋巴细胞大部分变为 $OKT4^+$ 细胞 [45,46]。HTLV-I 感染淋巴细胞引起的细胞病变效应是明显的，但是这一效应不如 HTLV-III 引起的显著 [16,17,45-48]。HTLV-III 与 HTLV-I 和 -II 的不同之处还在于它们在体内和体外产生的具有感染能力的病毒颗粒的数量不同。从艾滋病患者的精液和唾液以及 HTLV-III 感染的细胞培养液上清中能够获得高效价的感染性病毒颗粒 [10,49-51]。从 ATLL 病人或者从 HTLV-I 和 -II 感染的细胞培养基中，很少能够获得游离的病毒粒子 [52]。

为了研究这些病毒在体内和体外的生物学活性并为艾滋病的诊断和治疗药物的开发提供有用的信息，我们完成了 HTLV-III 前病毒完整核苷酸序列的测序。

HTLV-III 的基因组结构

从几个美国艾滋病患者血液中提取 HTLV-III，然后感染 H9 细胞系，从中获得了几个密切相关的 HTLV-III DNA 克隆 [10]。利用一个完整前病毒克隆的部分序列 [53]（图 1），我们测定了长度 [27] 分别为 8,900、5,300 和 3,600 个碱基的三个不完整病毒克隆的全部初级序列。

HTLV-III 前病毒长 9,749 个碱基对。前病毒的总体结构类似于其他的逆转录病毒。编码病毒蛋白的序列两侧为 LTR 序列。LTR 自身的两侧是 2 个核苷酸长度的反向重复序列（图 1）。 在病毒 DNA 中发现了四个长的开放阅读框（图 2）。

长末端重复序列

HTLV-III LTR 的详细分析已经被报道过 [54]；该序列长 634 个核苷酸，有 U3、R 和 U5 三个区域，长度分别是 453、98 和 83 个核苷酸。这些区域在 LTR 中的边界是通过 S_1 核酸酶谱和测量强制终止 DNA 转录产物的长度定位 5′端帽子位点，以及通过 cDNA 克隆的序列分析来确定病毒 RNA 的 3′末端，这些方法来确定的。在图 1 中标明了 TATAA 序列，它是真核启动子的典型序列，保守的多腺嘌呤序列也被标出。与 $tRNA^{Lys}$ 的 3′末端互补的转运 RNA 结合位点位于 5′LTR 的 3′末端。在图 3 中标出了与 HTLV-I[32]、-II[55] 和 BLV[56] 的 LTR 同源的序列。


CLONE
核苷酸位置 氨基酸残基

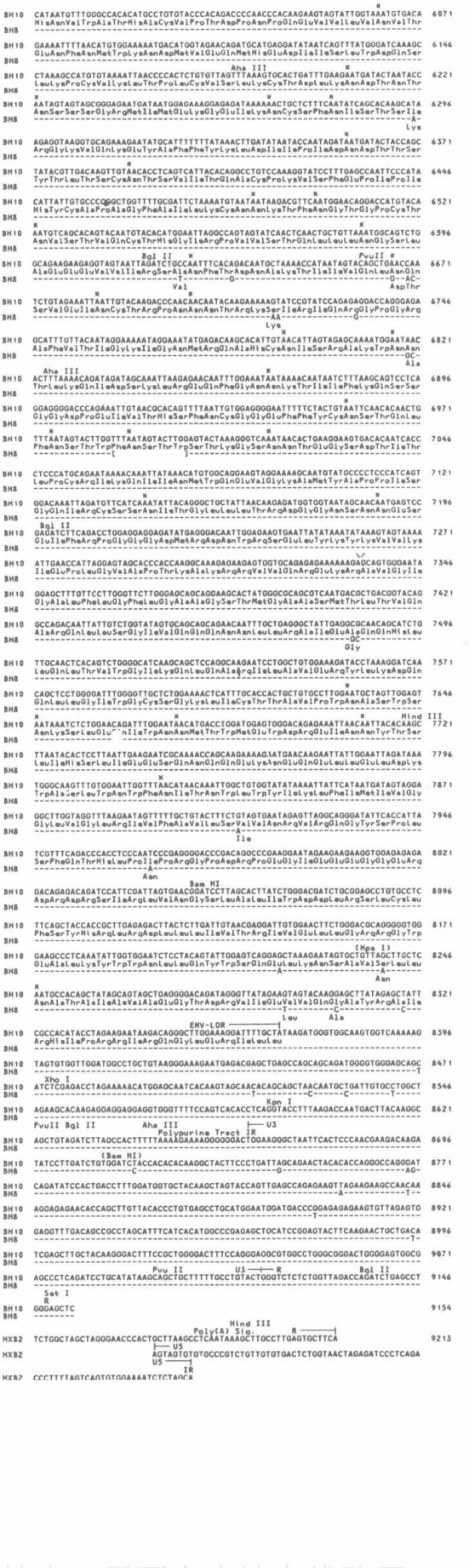


Fig. 1. Nucleotide sequence of HTLV-III. The complete nucleotide sequence of clone BH10 is shown together with the predicted amino acid sequence of the four largest open reading frames. The position of sequences encoding *gag* protein p17, the N-terminus of *gag* p24 and the C-terminus of *gag* p15 (which overlaps with the N-terminus of the *pol* protein) are indicated. The open reading frames for *pol*, *sor* and *env-lor* are indicated. The sequence of the remaining 182 bp of the HTLV-III provirus not present in clone BH10 (including a portion of R, U5, the tRNA primer binding site and a portion of the leader sequence) was derived from clone HXB2. The boundaries of R, U5 and U3, the positions of the polypurine tract, inverted repeated sequences (IR) and the transcriptional initiation (TATA) and termination (AATAA) signals are shown. The sequences of BH8 and BH5 are illustrated also; nucleotide and predicted amino acid differences compared with BH10 are listed and dashes are shown for identical sequences. Restriction enzyme sites are listed above the nucleotide sequence and sites present in clone BH8 but not BH10 are in parentheses. Deletions are also noted ([ ]) at nucleotides 251, 254, 5,671 and 6,987–7,001. The nucleotide positions (to the right of each line) start with the transcriptional initiation site and end with the viral RNA transcriptional termination site. The amino acid residues are numbered (to the right of each line) for the four largest open reading frames starting after the preceding termination codon in each case except *gag* which is enumerated from the first methionine codon. A proposed peptide cleavage site (v) and possible asparagine-linked glycosylation sites [77] are shown (*) for the *env-lor* open reading frame. The sequences in the LTR derived from clones BH8 and BH10 listed in the beginning of the figure are derived from the 3′ portion of each clone and are assumed to be identical to those present in the 5′ LTR of the integrated copies of these viral genomes. Clone HXB2 was derived from a recombinant phage library of *Xba*I-digested DNA from HTLV-III-infected H9 cell cloned in λJ1 (ref. 53). Clones BH10, BH8 and BH5 were derived from a library of *Sst*I-digested DNA from the Hirt supernatant fraction of HTLV-III-infected H9 cells cloned in λgtWes·λB (ref. 27). Both libraries were screened with a cDNA probe synthesized from virion RNA using oligo(dT) as a primer[26]. Clones BH8, BH5 and a portion of HXB2 were sequenced as described previously[91]. Clone BH10 was sequenced by the method of Sanger[92] modified by the use of oligonucleotides complementary to the M13 insert sequence as primers and using Klenow fragment of DNA polymerase I or reverse transcriptase as the polymerase.