

# nature

The Living Record of Science

《自然》百年科学经典

英汉对照版 (平装本)

第七卷 (上)

总顾问: 李政道 (Tsung-Dao Lee)

英方主编: Sir John Maddox      中方主编: 路甬祥  
Sir Philip Campbell

X 2002-2007

IX 1996-2001

VIII 1993-1997

VII 1985-1992

VI 1973-1984

V 1966-1972

IV 1946-1965

III 1934-1945

II 1931-1933

I 1869-1930

外语教学与研究出版社 · 麦克米伦教育 · 自然科研

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS · MACMILLAN EDUCATION · NATURE/RESEARCH

# nature

The Living Record of Science

《自然》百年科学经典



英汉对照版 (平装本)

第七卷 (上)

总顾问: 李政道 (Tsung-Dao Lee)

英方主编: Sir John Maddox      中方主编: 路甬祥  
Sir Philip Campbell



外语教学与研究出版社 · 麦克米伦教育 · 自然科研

FOREIGN LANGUAGE TEACHING AND RESEARCH PRESS · MACMILLAN EDUCATION · NATURE RESEARCH

北京 BEIJING

Original English Text © Macmillan Publishers Limited  
Chinese Translation © Foreign Language Teaching and Research Press

This edition is published under arrangement with Macmillan Publishers (China) Limited. It is for sale in the People's Republic of China only, excluding Hong Kong SAR, Macao SAR and Taiwan Province, and may not be bought for export therefrom.

### 图书在版编目(CIP)数据

《自然》百年科学经典·第七卷·上, 1985-1992: 英汉对照 / (英) 约翰·马多克斯 (Sir John Maddox), (英) 菲利普·坎贝尔 (Sir Philip Campbell), 路甬祥主编. -- 北京: 外语教学与研究出版社, 2017.8

ISBN 978-7-5135-9426-4

I. ①自… II. ①约… ②菲… 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 (2017) 第 217343 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 Chris Balderston (美)  
责任编辑 黄小斌  
执行编辑 王丽霞  
装帧设计 彩奇风  
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 32  
版 次 2017 年 10 月第 1 版 2017 年 10 月第 1 次印刷  
书 号 ISBN 978-7-5135-9426-4  
定 价 168.00 元

购书咨询: (010) 88819926 电子邮箱: [club@fltrp.com](mailto:club@fltrp.com)

外研书店: <https://waiyants.tmall.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](mailto: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](mailto: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: 294260001

# 《自然》百年科学经典（英汉对照版）

总顾问：李政道（Tsung-Dao Lee）

英方主编：Sir John Maddox

中方主编：路甬祥

Sir Philip Campbell

## 编审委员会

英方编委：

Philip Ball

Vikram Savkar

David Swinbanks

中方编委（以姓氏笔画为序）：

许智宏

赵忠贤

滕吉文

## 本卷审稿专家（以姓氏笔画为序）

于 贵	王 琛	方向东	田立德	吕雪梅	朱道本	刘 新
刘冬生	刘佳佳	安俊岭	孙 军	孙 松	李 彦	李吉均
李兴中	李道季	肖景发	吴 琳	沈 杰	张忠杰	陈 文
陈尔强	周 江	胡文平	胡松年	胡卓伟	闻新宇	姚建年
秦志海	聂军胜	顾镇南	徐仁新	郭建栋	董 为	蒋世仰
韩汝珊						



## 编译委员会

### 本卷翻译工作组稿人 (以姓氏笔画为序)

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

### 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	王帅帅
王晓蕾	王海纳	王德孚	元旭津	公 晗	邓铭瑞	石宇鹏
石爱洁	史未卿	史明澍	朱 玥	任峻铭	刘俊杰	闫 妍
闫振丽	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	杨 茜	吴兆军
张 敏	张 晴	张欣园	张香香	张梦旋	张琦玮	张瑶楠
陈 雄	罗小青	郑 征	宗伟凯	赵晓非	侯鉴璇	姜 薇
夏 秋	钱 磊	徐 玲	徐 悦	高坛光	郭晓博	黄 瑢
黄 璞	黄雪嫒	梁 瑜	葛聆飒	谢周丽	雷文欣	魏洪淼

Eric Leher (澳)

# Contents

## 目录

Complete Nucleotide Sequence of the AIDS Virus, HTLV-III .....	2
艾滋病病毒 HTLV-III 的完整核苷酸序列 .....	3
Does the Ocean-Atmosphere System Have More than One Stable Mode of Operation? .....	26
海洋-大气系统是否存在不止一种稳定运行模态? .....	27
Large Losses of Total Ozone in Antarctica Reveal Seasonal $\text{ClO}_x/\text{NO}_x$ Interaction .....	52
南极臭氧总量大量损失揭示季节性 $\text{ClO}_x/\text{NO}_x$ 的相互作用 .....	53
A 150,000-year Climatic Record from Antarctic Ice .....	68
南极冰芯记录的过去 15 万年以来的气候变化 .....	69
$\text{C}_{60}$ : Buckminsterfullerene .....	96
$\text{C}_{60}$ : 巴克敏斯特富勒烯 .....	97
2.5-Myr <i>Australopithecus boisei</i> from West of Lake Turkana, Kenya .....	106
肯尼亚图尔卡纳湖以西 250 万年前的南方古猿鲍氏种 .....	107
Forty Years of Genetic Recombination in Bacteria: A Fortieth Anniversary Reminiscence .....	128
细菌遗传重组四十年——四十周年纪念 .....	129
Mitochondrial DNA and Human Evolution .....	140
线粒体 DNA 与人类进化 .....	141
Multi-channel Seismic Imaging of a Crustal Magma Chamber along the East Pacific Rise .....	166
沿东太平洋海隆的地壳岩浆房多道地震成像 .....	167
Transmission Dynamics of HIV Infection .....	190
HIV 的传播动力学 .....	191
Interstellar Diamonds in Meteorites .....	222
陨石中的星际金刚石 .....	223
Oceanic Phytoplankton, Atmospheric Sulphur, Cloud Albedo and Climate .....	236
海洋浮游植物、大气硫、云反照率与气候 .....	237

Structure of the Repressor–Operator Complex of Bacteriophage 434 .....	270
噬菌体 434 阻遏蛋白–操纵基因复合物的结构 .....	271
Iron Deficiency Limits Phytoplankton Growth in the North-East Pacific Subarctic .....	302
亚北极区东北太平洋海域铁缺乏对浮游植物生长的限制 .....	303
Evidence for Global Warming in the Past Decade .....	314
过去十年中全球变暖的证据 .....	315
GAL4 Activates Transcription in <i>Drosophila</i> .....	320
在果蝇中 GAL4 激活转录 .....	321
Mysteries of HIV: Challenges for Therapy and Prevention .....	336
神秘的 HIV: 它的治疗和预防面临的挑战 .....	337
Human Basophil Degranulation Triggered by Very Dilute Antiserum against IgE .....	356
高度稀释的抗 IgE 抗血清引发人嗜碱性粒细胞脱颗粒 .....	357
“High-Dilution” Experiments a Delusion .....	374
高倍稀释实验的错觉 .....	375
New Semiconductor Device Physics in Polymer Diodes and Transistors .....	398
高分子二极管和晶体管中半导体器件物理学的新学问 .....	399
Resolution of Quantitative Traits into Mendelian Factors by Using a Complete Linkage Map of Restriction Fragment Length Polymorphisms .....	420
通过限制性酶切片长度多态性的完整连锁图对孟德尔因子数量性状的解析 .....	421
Observation of Cold Nuclear Fusion in Condensed Matter .....	440
凝聚态物质中观察到的冷核聚变 .....	441
Problems with the $\gamma$ -Ray Spectrum in the Fleischmann <i>et al.</i> Experiments .....	460
弗莱施曼等人实验中 $\gamma$ 射线谱的问题 .....	461
DNA Fingerprinting on Trial .....	472
DNA 指纹图谱技术在庭审中的应用 .....	473
Stratospheric Clouds and Ozone Depletion in the Arctic during January 1989 .....	500
1989 年 1 月北极上空的平流层云与臭氧亏损 .....	501
A Membrane-Targeting Signal in the Amino Terminus of the Neuronal Protein GAP-43 .....	522
神经元蛋白 GAP-43 氨基末端中的膜导向信号 .....	523

Rescue of <i>bicoid</i> Mutant <i>Drosophila</i> Embryos by Bicoid Fusion Proteins Containing Heterologous Activating Sequences.....	538
通过含异源激活序列的Bicoid融合蛋白拯救 <i>bicoid</i> 突变果蝇胚胎.....	539
High-Latitude Ozone Loss Outside the Antarctic Ozone Hole.....	564
南极臭氧洞外高纬度地区臭氧的亏损.....	565
Upper Bounds on “Cold Fusion” in Electrolytic Cells.....	590
电解池中“冷核聚变”的上限.....	591
Natural Variability of the Climate System and Detection of the Greenhouse Effect.....	634
气候系统的自然变率与温室效应的检测.....	635
Limits on the Emission of Neutrons, $\gamma$ -Rays, Electrons and Protons from Pons/Fleischmann Electrolytic Cells.....	648
庞斯/弗莱施曼电解池的中子、 $\gamma$ 射线、电子和质子发射的上限.....	649
Positioning Single Atoms with a Scanning Tunnelling Microscope.....	668
用扫描隧道显微镜定位单个原子.....	669
Ozone Loss in the Arctic Polar Vortex Inferred from High-Altitude Aircraft Measurements.....	678
北极极涡臭氧的亏损——来自高海拔航空测量的结果.....	679
Solid C <sub>60</sub> : a New Form of Carbon.....	706
固态C <sub>60</sub> : 碳的一种新形式.....	707
Light-Emitting Diodes Based on Conjugated Polymers.....	724
基于共轭聚合物的发光二极管.....	725
Evidence for a Common Central-Engine Mechanism in all Extragalactic Radio Sources.....	736
所有河外星系射电源有共同中央引擎机制的证据.....	737
Estimates of the Effect of Southern Ocean Iron Fertilization on Atmospheric CO <sub>2</sub> Concentrations.....	748
南大洋铁施肥对大气CO <sub>2</sub> 浓度的影响评估.....	749
Superconductivity at 18 K in Potassium-Doped C <sub>60</sub> .....	764
掺钾的C <sub>60</sub> 在18 K时的超导性.....	765
Synthesis from DNA of a Molecule with the Connectivity of a Cube.....	772
用DNA合成具有立方连接形式的分子.....	773



Recent Variations in Arctic and Antarctic Sea-Ice Covers.....	784
南北极海冰覆盖面积的近期变化.....	785
Predominance of Long-Wavelength Heterogeneity in the Mantle.....	800
地幔中长波长非均匀性的主导作用.....	801
Helical Microtubules of Graphitic Carbon.....	828
石墨碳的螺旋状微管.....	829
A Planetary System around the Millisecond Pulsar PSR1257+12.....	840
毫秒脉冲星 PSR1257+12 的行星系统.....	841
Photochromism Induced in an Electrolytically Pretreated MoO <sub>3</sub> Thin Film by Visible Light.....	854
电解预处理 MoO <sub>3</sub> 薄膜的可见光变色.....	855
The <i>C. elegans</i> Genome Sequencing Project: a Beginning.....	864
秀丽隐杆线虫基因组测序计划: 一个开端.....	865
Tectonic Forcing of Late Cenozoic Climate.....	886
构造运动对晚新生代气候的影响.....	887
Curling and Closure of Graphitic Networks under Electron-Beam Irradiation.....	916
石墨网格在电子束照射下的卷曲与闭合.....	917
Ordered Mesoporous Molecular Sieves Synthesized by a Liquid-Crystal Template Mechanism.....	930
利用液晶模板机制合成的有序介孔分子筛.....	931
Evidence for Massive Discharges of Icebergs into the North Atlantic Ocean during the Last Glacial Period.....	944
末次冰期大量冰山流入北大西洋的证据.....	945
Solar Cycle Length, Greenhouse Forcing and Global Climate.....	964
太阳活动周期长度、温室强迫与全球气候.....	965
Implications for Global Warming of Intercycle Solar Irradiance Variations.....	978
太阳辐照度周期性变化对全球变暖的意义.....	979
Appendix: Index by Subject	
附录: 学科分类目录.....	995

**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.

---

**H**UMAN 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)<sup>1-14</sup>. These viruses share a number of biological and structural features: tropism for OKT4<sup>+</sup> lymphocytes<sup>10,15-21</sup>, the ability to produce giant multinucleated cells in culture<sup>10,17,22-24</sup>, immunological cross-reactivity of some virally encoded proteins<sup>12,13</sup>, distant nucleic acid sequence similarities<sup>8,9,25-28</sup> and the preference for magnesium of the viral reverse transcriptase<sup>29-31</sup>. 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<sup>8,9,25,32-34</sup>. 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<sup>25,32-37</sup>, the 3' portion of which is a long open reading frame (*lor*)<sup>32,35,37</sup>. 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<sup>38,39</sup>. This phenomenon, *trans*-acting transcriptional regulation, is not observed for other retroviruses<sup>40</sup> and is probably mediated by the *lor* gene product of these retroviruses<sup>38</sup>.

# 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<sup>+</sup>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<sup>+</sup> cell population<sup>10-14,41-43</sup>. This effect is mirrored by a pronounced cytopathic rather than transforming effect of HTLV-III infection upon the OKT4<sup>+</sup> cells in lymphocyte cultures *in vitro*<sup>10,11,20</sup>. In contrast, infection with HTLV-I results in a low incidence of T-cell leukaemia lymphoma (an OKT4<sup>+</sup>-cell malignancy)<sup>1-6</sup>. There is evidence also for some degree of immunodeficiency in HTLV-I patients<sup>6,44</sup>. Infection of primary lymphocytes in culture by HTLV-I and -II results in *in vitro* transformation of predominantly OKT4<sup>+</sup> cells<sup>45,46</sup>. A cytopathic effect of HTLV-I infection on lymphocytes is apparent, but the effect is not as pronounced as that in HTLV-III<sup>16,17,45-48</sup>. 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<sup>10,49-51</sup>. Very few, if any, cell-free infectious virions can be recovered from ATLL patients or from cultures infected with HTLV-I or -II<sup>52</sup>.

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<sup>10</sup>. The complete primary nucleotide sequence of three unintegrated viral clones of 8.9, 5.3 and 3.6 kilobases (kb) in length<sup>27</sup> was determined with the partial sequence from an integrated proviral clone<sup>53</sup> (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<sup>54</sup>; 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<sub>1</sub> 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<sup>Lys</sup> is located 3' to the 5' LTR. DNA sequence homologies to the LTR of HTLV-I<sup>32</sup>, -II<sup>55</sup> and BLV<sup>56</sup> 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<sup>+</sup> 细胞数减少<sup>[10-14,41-43]</sup>。这一效应是通过 HTLV-III 感染体外淋巴细胞培养物 OKT4<sup>+</sup> 造成的细胞病变反映出来的，而非转染效应<sup>[10,11,20]</sup>。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<sup>+</sup> 细胞恶性肿瘤）<sup>[1-6]</sup>，而且也有 HTLV-I 病人存在不同程度的免疫缺陷的例子<sup>[6,44]</sup>。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<sup>+</sup> 细胞<sup>[45,46]</sup>。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<sup>[16,17,45-48]</sup>。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<sup>[10,49-51]</sup>。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<sup>[52]</sup>。

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<sup>[10]</sup>。利用一个完整前病毒克隆的部分序列<sup>[53]</sup>(图 1)，我们测定了长度<sup>[27]</sup>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<sup>[54]</sup>；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<sub>1</sub> 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<sup>Lys</sup> 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<sup>[32]</sup>、-II<sup>[55]</sup> 和 BLV<sup>[56]</sup> 的 LTR 同源的序列。

Complete Nucleotide Sequence of the AIDS Virus, HTLV-III

CLONE	NUCLEOTIDE POSITION	AMINO ACID RESIDUE
BH10	18	277
BH8	320	381
BH10	348	2846
BH8	370	406
BH10	398	2921
BH8	420	431
BH10	448	2996
BH8	470	436
BH10	498	3071
BH8	520	481
BH10	548	3146
BH8	570	506
BH10	598	3221
BH8	620	531
BH10	648	3311
BH8	670	581
BH10	698	3446
BH8	720	606
BH10	748	3521
BH8	770	631
BH10	798	3596
BH8	820	636
BH10	848	3671
BH8	870	681
BH10	898	3746
BH8	920	706
BH10	948	3821
BH8	970	731
BH10	998	3896
BH8	1020	756
BH10	1048	3971
BH8	1070	781
BH10	1098	4046
BH8	1120	806
BH10	1148	4121
BH8	1170	831
BH10	1198	4196
BH8	1220	856
BH10	1248	4271
BH8	1270	881
BH10	1298	4346
BH8	1320	906
BH10	1348	4421
BH8	1370	931
BH10	1398	4496
BH8	1420	956
BH10	1448	4571
BH8	1470	981
BH10	1498	4646
BH8	1520	1006
BH10	1548	4721
BH8	1570	1031
BH10	1598	4796
BH8	1620	1056
BH10	1648	4871
BH8	1670	1081
BH10	1698	4946
BH8	1720	1106
BH10	1748	5021
BH8	1770	1131
BH10	1798	5096
BH8	1820	1156
BH10	1848	5171
BH8	1870	1181
BH10	1898	5246
BH8	1920	1206
BH10	1948	5321
BH8	1970	1231
BH10	1998	5396
BH8	2020	1256
BH10	2048	5471
BH8	2070	1281
BH10	2098	5546
BH8	2120	1306
BH10	2148	5621
BH8	2170	1331
BH10	2198	5696
BH8	2220	1356
BH10	2248	5771
BH8	2270	1381
BH10	2298	5846
BH8	2320	1406
BH10	2348	5921
BH8	2370	1431
BH10	2398	5996
BH8	2420	1456
BH10	2448	6071
BH8	2470	1481
BH10	2498	6146
BH8	2520	1506

Table with columns for CLONE, Nucleotide position, Amino acid residue, and sequence. The table lists the complete nucleotide sequence of HIV-III, with corresponding amino acid translations and various annotations such as Hind III, Pvu II, and Eco RI sites.



BH10 CATATGTTGGCCACCACTGGCTGTCCACCAACCCACCAACCAAGAAATGATTTGGTAAATGGACA 4971
9
BH8 HlaAnvV1TrpA1ThrAlaCysValProThrAspPheCysProGlnGluValValLysValLysVal 5
BH10 GAAATTTACATGTGGAAATGACAFTGTAGGACAGATGCGATGATATACGTTATTTGGGATCAAGC 6164
12
BH8 GluAnPheAnPheAnPheValGlnMetHisGlyAlaIleIleSerLeuTrpAspGlnSer 12
BH10 CTAAGCCATGTTAAATTAACCCCACTGTTGTTAATGAGTCCACTGATTTGAGAAATGATGATTAATCC 6221
14
BH8 LeuVpProValIleValLysValLeuThrProGlnGluValIleValIleValIleValIleValIleVal 14
BH10 AAATGATGATGCGAGGAAATGATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6296
17:
BH8 AAsSerSerSerGlyValArgPheHisIleLeuGluValIleValIleValIleValIleValIleValIle 17:
BH10 AGAAGTAAAGTGCAGAAAGATGACATTTTTTATAACTGATGATTAAGCCATAGATGATGACTACTACC 6371
19:
BH8 ArgGlyValIleGlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 19:
BH10 TATACGTTGCAAGTGTGACACCTGCTATACAGAGCCGCTGCAAGAGTCTCTTGGACGATCCCATCA 6444
22:
BH8 TyrThrLeuThrSerCysAnThrSerValIleThrGlnAlaCysProIleValIleValIleValIleVal 22:
BH10 CATTATGTCGCCCGGCTTTGGCATTTCAAAATGATTAAGACGATCTCATAGACAGGACAGGATGACA 6521
24:
BH8 HisValCysIlePheIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 24:
BH10 AATGTCAACACACATGATGACATGGAATTAAGCCAGTATGATGACACTGCTGTTAAATCCAGCTG 6594
27:
BH8 AsnValIleThrValIleGlnCysThrHisGlyIleArgProIleValIleValIleValIleValIleVal 27:
BH10 CAGAGAGAGAGTAAATTAATGTCACCAATGACAGAGGATGCAAAACATTAATGATGACAGGACCA 6671
29:
BH8 AlaIleGluIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 29:
BH10 TCTGTAGAAATTAATGTCCAAACCCACACACATACCAAGAAAGATTCCTGATCCAGAGAGGAGGAG 6746
32:
BH8 SerValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 32:
BH10 GCATTTGTACATAGGAAATAAGGAAATGACAGACCACTTTTACATTAATGAGGACAAAGATTTGAA 6821
34:
BH8 AlaPheValIleThrIleGlyIleValIleValIleValIleValIleValIleValIleValIleValIle 34:
BH10 ACTTTAAACAGATGATGACAAATTAAGGACCAATTTGAAATTAATGAAACATTAATCTTTAAGCAG 6894
37:
BH8 ThrLeuValIleIlePheSerValLeuGlnIleValIleValIleValIleValIleValIleValIleVal 37:
BH10 GAGG 6971
39:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 39:
BH10 TTTAATGATCTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7046
42:
BH8 PheAnSerThrTrpPheSerThrTrpSerThrValIleValIleValIleValIleValIleValIleVal 42:
BH10 TCCCATCGAATTAACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7121
44:
BH8 LeuPheCysArgIleIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 44:
BH10 GGCACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7196
47:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 47:
BH10 GATG 7271
49:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 49:
BH10 ATTACCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7346
52:
BH8 IleIleIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 52:
BH10 GGCATTTGTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7421
54:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 54:
BH10 CCGACAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7496
57:
BH8 AlaIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 57:
BH10 TTGCATCAGCTGTCGGGATCAGACAGCTCAGGACAGGATTCGGTGGTGGAAATACCTAAAGGATCAA 7571
59:
BH8 LeuGlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 59:
BH10 CAGCTCTGGGATTTGGGATGCTGGAAACTTTTGCACACTGCTGCTGCTGATGATGATGATGATGAT 7646
42:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 42:
BH10 AATAAATCTGACAGATTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7721
44:
BH8 AsnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 44:
BH10 TTAATGATCTGTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7796
47:
BH8 LeuIleIleSerLeuIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 47:
BH10 TGGCAATTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7871
49:
BH8 TrpIleSerLeuTrpPheAnIleIleValIleValIleValIleValIleValIleValIleValIleVal 7871
BH10 GCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7946
72:
BH8 GlnIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 72:
BH10 TCGTTTCAAGCCCACTCCCAATCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8021
74:
BH8 SerPheGlnThrHisAlaProIlePheArgGlyIlePheArgGlyIlePheArgGlyIlePheArgGly 74:
BH10 GACAGACAGATCCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8096
77:
BH8 AspArgMetSerIleIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 77:
BH10 TTGACATCCACCGCTTGGAGACTTACTGTTGATTAAGCAGGATTTGGAACCTTGGAGCCAGAGGAT 8171
79:
BH8 PheSerThrHisIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 79:
BH10 GAAGCCCTCAATATGTTGGATCTCTCAAGATTTGGATGATGATGATGATGATGATGATGATGATGAT 8246
82:
BH8 GluAlaIleValIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 82:
BH10 TACACCACTTACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8321
84:
BH8 AsnAlaThrAlaIleValIleValIleValIleValIleValIleValIleValIleValIleValIleVal 84:
BH10 CCCCACCTACCTAAGAAATTAAGCAGGATTTGATTAAGATGATGATGATGATGATGATGATGATG 8396
86:
BH8 ArgHisIlePheArgGlyIleValIleValIleValIleValIleValIleValIleValIleValIleVal 86:
BH10 TAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8471
88:
BH10 ATCTGGAGCCATGAGAACTGAGGACCTCACATAGCCACACAGCAGCTAACATGATGATGATGATG 8546
91:
BH8 BH8 AGAAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8621
93:
BH10 PheIle 93:
BH10 ACGTGTGATCTTCAAGCTTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8696
96:
BH10 TATCTTGTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 8771
98:
BH10 CAGATATCCAGTACCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 8846
101:
BH10 AGGAGAGACACCACTGTTTACACCTGTGACCTGATGATGATGATGATGATGATGATGATGATGATG 8921
103:
BH8 BH8 GAGGTTGACAGCCGCTAGCATTTTACACCTGACAGCCGATGATGATGATGATGATGATGATGATGAT 8996
106:
BH10 TCGACTGATCAAGGAGCTTTCCGCTGGGAGTTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 9071
108:
BH8 BH8 AGCCCTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9146
111:
BH10 GGGAGCTC
113:
BH8 BH8 TCTGCTAGTGGGAAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9215
114:
BH8 BH8 CCCTTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 9284
117:
BH8 BH8

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 XbaI-digested DNA from HTLV-III-infected H9 cell cloned in lambdaJ1 (ref. 53). Clones BH10, BH8 and BH5 were derived from a library of SstI-digested DNA from the Hart supernatant fraction of HTLV-III-infected H9 cells cloned in lambdaGates lambdaB (ref. 27). Both libraries were screened with a cDNA probe synthesized from virion RNA using oligo(dT) as a primer<sup>26</sup>. Clones BH8, BH5 and a portion of HXB2 were sequenced as described previously<sup>21</sup>. Clone BH10 was sequenced by the method of Sanger<sup>32</sup> 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.