

## HIV2/SIV POL

### Protein Alignment Summary Table (Pol)

Name	Accession No.	Region	First Author	Reference
ROD	M15390	complete genome	Guyader,M.	Nature 326 (6114), 662-669 (1987)
NHIZ	J03654	complete genome	Zagury,J.F.	PNAS 85, 5941-5945 (1988)
ISY	J04498	complete genome	Franchini,G.	PNAS 86, 2433-2437 (1989)
ST	M31113	complete genome	Kumar,P.	J. Virol. 64, 890-901 (1990)
BEN	M30502	complete genome	Kirchhoff,F.	Virology 177, 305-311 (1990)
CAM2	D00835	complete genome	Tristem,M.F.	J. Gen. Virol. 72, 721 (1991)
D194	J04542	complete genome	Kuehnel,H.	PNAS 86, 2383-2387 (1989)
GHI	M30895	complete genome	Hasegawa,A.	ARHR 5, 593-604 (1989)
UC2	U38293	complete genome	Barnett,S.W.	Unpublished (1997)
MDS	Z48731	gag pol vif tat rev nef env	Becker,M.	Unpublished
U22047 KR	U22047	complete genome	Kraus,G.K.	PNAS 90 (9), 4226-4230 (1993)
FA2	L33089	partial pol integrase	Gao,F.	J. Virol. 68, 7433-7477 (1994)
5132P1	L16970	pol partial	Gao,F.	ARHR 9, 703-704 (1993)
S52907	S52907	pol RT region	Warren,T.C.	Protein Expr. Purif. 3, 479-487 (1992)
HIV2ALT HIV2D205	X61240	complete genome	Dietrich Et Al,U.	ARHR 8, 1619 (1992)
UC1	L07625	complete genome	Barnett,S.W.	J. Virol. 67, 1006-1014, (1993)
EHOA	U27200	complete genome	Rey-Cuille,M.A.	Virology 202 (1), 471-476 (1994)
JA2	L33090	partial pol integrase	Gao,F.	J. Virol. 68, 7433-7477 (1994)
ONFT2	L33091	partial pol integrase	Gao,F.	J. Virol. 68, 7433-7477 (1994)
22POLB7	M87138	partial pol	Gao,F.	Nature 358, 495-499 (1992)
FORTC2	M87111	pol, RT region, partial	Gao,F.	Nature 358, 495-499 (1992)
POLC12	M87110	pol, partial integrase	Gao,F.	Nature 358, 495-499 (1992)
MM1A11	M76764	complete genome	Luciw,P.A.	ARHR 8: 395-402 (1992)
MM251	M19499	complete genome	Franchini,G.	Nature 328, 539-543 (1987)
MM32H	D01065	complete genome	Rud,E.W.	J. Gen. Virol. 75, 529 (1994)
MM239	M33262	complete genome	Regier,D.A.	ARHR 6, 1221-1231 (1990)
MM142	M16403	complete genome	Chakrabarti,L.	Nature 328, 543-547 (1987)
MNE	M32741	complete genome	Benveniste,R.E.	Unpublished (1990)
SMMH4	X14307	complete genome	Hirsch,V.M.	Nature 339, 389-391 (1989)
SIVSMMH9	M80194	complete genome	Cournaud,V.	J. Virol. 66, 414-419 (1992)

II-B-8

DEC 97

## HIV2/SIV POL

SIVSMMPIA	M31325	complete genome	Dewhurst, S.	Nature 345, 636-640 (1990)
STM	M83293	complete genome	Novembre, F.J.	Virology 186, 783-787 (1992)
SMMLIB1	M62651	pol	Marx, P.A.	J. Virol. 65, 4480-4485 (1991)
SMMM7	M27256	pol	Li, Y. et al.	Unpublished (1989)
PA2	L33093	partial pol RT	Gao, F.	J. Virol. 68, 7433-7477 (1994)

HIV2/SIV POL

CONSENSUS-A VLEIWKGGTLLGETVSPSTOKTGLIEMWQ?RTHGKLIpRkTNG?FRRdgp?gkeapQlPrgrssasGADtNStpr?ssagsgelIAdakeagAT. ?ettiQrDp?gl tApragrdt?QR?dRGLAaPQFSLM
-> Pol reading frame (-1 from Gag) p15
p15 peptide \ Protease
121
90
90
110
110
90
128
110
110
110
110
110

CONSENSUS-B
D205
UC1
EHODACG
--M--k--V-M--g--VRLI--S--hd--as--s--IC--Dep-r-HDTSGdTIcAPCRSSSSGDAekLlHad-E--E-Ep-E-L-G-G--F--
--M--KN--V-M--G--VRTI--S--HD--AS--S--IC--DPP-R-HDTSGdDTIcAPCRSSSSGDAEKlHAD-E--TE-EP-E-L-G-G--F--
--M--KT--V-M--W--VRTI--S--HD--AS--S--IC--DPP-R-HDTSGdDTIcAPCRSSSSGDAEKlHAD-E--E-Ep-E-L-G-G--F--
--A-M--TA--V-M--G--VRLI--S--GTP--DSAICA-DEP-IRHDTSGdDStIcTPCRSSSRGDAEKlHADtREB-EGbQ-E-L-G-G--F--

CONSENSUS-SD
MML111AA
MM251
MM32H
MM239
MML42
MNE
SMH4
SMH9
SMMPBJ14.441
SMMPBJ14.15
SMMPBJ6.6
--e??-CKAMQ-PK--m--kngPCY-QM--q--g--?W-m--F-H-?AS--a--CS-Rg?-C-Sak-LH-?GQ?????.-E-KQ-BAL-G-G--F--
--BR--CKAMQ-PK--M--KNGPCY-QM--Q--G--PWSM--F-H-S-AS--A-CS-RGP-C-SAK-LH-VGQA-RK.....TE-KQ-BAL-G-G--F--
--B--CKAMQ-PK--M--KNGPCY-QM--Q--G--PWSM--F-H-S-AS--A-CS-RGP-C-SAK-LH-VGQA.....E-KQ-BAL-G-G--F--
--RRR--CKAMQ-PK--M--KNGPCY-QM--Q--G--PWSM--F-H-S-AS--A-CS-RGP-C-SAK-LH-VGQA-RK.....E-KQ-BAL-G-G--F--
--BR--CKAMQ-PK--M--KNGPCY-QM--Q--G--PWSM--F-H-S-AS--A-CS-RGP-C-SAK-LH-VGQA-RK.....E-KQ-BAL-G-G--F--
--E-R--CKAMQ-PK--M--KNGPCY-QM--KQ--G--PW-L--F-H-S-AS--A-CS-R-T-C-SAK-LH-LGQA.....E-KQ-EAL-G-G--F--
--BR--CKAMQNPk--M--KNGPCY-QM--Q--R--LW-M--F-H-NAS--A-CS-RGP-C-SAKKlH-VGQT.....E-KQKAl-G-G--F--
--G--AW-M--F-H-DAS--CS-RG-C-S-E-LHEDGQ.....EGbQ-E-L-G-G--F--
--M--G--AW-M--F-H-DAS--CSXRG-C-S-E-LHEDGQ.....EGbQ-E-L-G-GNG-F--
--R--CK-MQ-P\$-----V-KS-PCY-\$M--SG--AW-M--F-H-DAS--CS-RG-C-S-E-LHEDGQ.....EGbQ-E-L-G-GNG-F--
--R--CK-MQ-P\$-----V-KS-PCY-\$M--SG--AW-M--F-H-DAS--CS-RG-C-S-E-LHEDGQ.....EGbQ-E-L-G-GNG-F--
M-----SG--AW-M--F-H-DAS--CS-RG-C-S-E-LHEDGQ.....EGbQ-E-L-G-GNG-F--

CONSENSUS-STM
STM
RPE-R-----TTGpSD-QM-----G-----LW-M-----H-NTS--A-CPSR-P-C-SAEKlH--GQE.....-E-EQE--L-G-G--F--
RPE-R-----TTGpSD-QM-----G-----LW-M-----H-NTS--A-CPSR-P-C-SAEKlH--GQE.....-E-EQE--L-G-G--F--







**HIV2/SIV POL**

CONSENSUS- A LAEAELEEN?iIIsQeQEGHYYQEEKELBATVgKdQdnQMtYKIHQ?ekILkVgKyakIknhthngvrLLaGvVQKIGKALVIGR4PKFHLVPEReLWEQWMDnYwQVTW1PdMDFvSTPPLVRLaFN  
 ROD R-----R-----E-----E-----V-----I-----I-----I-----I-----  
 NIH2 R-----R-----K-----K-----V-----G-----G-----IC-D-KYPYOR-----D-----  
 ISY K-----K-----K-----K-----V-----G-----V-----G-----V-----  
 ST K-----K-----C-----Y-----I-----S-GH-----E-----E-----D-----I-----  
 BEN R-----R-----K-----Y-----I-----S-GH-----E-----E-----D-----I-----  
 CAM2 R-----R-----K-----Y-----I-----S-----A-----G-RV-----H-----K-----V-----  
 DL94 K-----K-----S-----E-----I-----S-----A-----G-RV-----H-----K-----V-----  
 GH1 R-----R-----D-----Y-----I-----N-----E-----E-----D-----I-----  
 UC2 R-----R-----D-----Y-----I-----S-----E-----E-----D-----I-----  
 MDS R-----RV-----D-----Y-----I-----S-----E-----E-----D-----I-----  
 KR R-----K-----R-----C-----G-----G-----T-----V-----G-----T-----V-----  
 S52907 R-----R-----K-----R-----C-----G-----G-----T-----V-----G-----T-----V-----  
 CONSENSUS- B -----Q-K-E-E-----S-K-gVP-----NLA-----Gn-----V-----V-----H-----E-v-----D-TD-----E-----I-----Y-  
 D205 -----Q-K-E-E-----S-K-RVP-----NLA-----GN-V-----V-----V-----H-----E-V-----D-TD-----E-----I-----Y-  
 UC1 -----Q-K-E-E-----S-K-GVP-----NLA-----GNR-----V-----V-----E-V-----D-TD-----E-----I-----Y-  
 HDACG -----FQ-K-E-E-----S-K-GVP-----NLA-----GD-----V-----V-----E-W-----D-TD-----E-----I-----Y-  
 CONSENSUS- D M-----Y-K-K-----C-----S-P-----V-----V-----S-----ED-----H-----I-----I-----V-S-I-DV-----SD-----E-----V-  
 FORTC2 M-----Y-K-K-----C-----S-P-----V-----V-----S-----ED-----H-----I-----I-----V-S-I-DV-----SD-----E-----V-  
 CONSENSUS- SD M-----Y-K-K-----C-----G-P-----I-S-----S-----ED-----F-----H-I-----I-----QV-----2dV-----TD-----E-----I-----V-  
 MM1A11AA M-----Y-K-K-----C-----G-P-----I-N-----S-----ED-----F-----H-I-----I-----QV-----KDV-----TD-----E-----N-I-----V-  
 MM251 M-----Y-K-K-----C-----G-P-----I-S-----S-----ED-----F-----H-I-----I-----QV-----DV-----TD-----E-----I-----V-  
 MM32H M-----Y-K-K-----C-----G-P-----I-S-----S-----ED-----F-----H-I-----I-----QV-----KDV-----TD-----E-----I-----V-  
 MM239 M-----Y-K-K-----C-----G-P-----I-S-----S-----ED-----F-----H-I-----I-----QV-----KDV-----TD-----E-----I-----V-  
 MM142 M-----Y-K-K-----C-----S-P-----I-S-----S-----ED-----F-----H-I-----I-----QV-----KDV-----TD-----E-----I-----V-  
 MNE M-----Y-K-K-----C-----G-P-----I-N-----S-----ED-----F-----H-I-----I-----QV-----KDV-----TD-----E-----I-----V-  
 SMMH4 M-----Y-K-K-----C-----G-PI-----I-S-----S-----ED-V-----F-V-----EV-----I-----TD-----E-----I-----V-  
 SMMH9 M-----Y-K-P-C-----G-P-P-I-S-----S-----ED-----F-----H-I-----I-----QV-R-----x-I-----TD-----E-----I-----V-  
 SMMPB14.441 M-----Y-K-K-----C-----G-P-----I-S-----S-----ED-----F-----H-----I-----I-----QV-R-----I-----TD-----E-----I-----V-  
 CONSENSUS- STM M-----YA-K-K-----Q-R-D-P-----V-----V-----S-----ED-----F-----H-I-----I-----I-----Q-----V-----V-----AD-----E-----V-  
 STM M-----YA-K-K-----Q-R-D-P-----V-----V-----S-----ED-----F-----H-I-----I-----I-----Q-----V-----V-----AD-----E-----V-  
 CONSENSUS- SO AR? -??-??-??-?-----V-  
 SMMLIB1 X-----E-#-I-----V-  
 SMMM7 ARP-IRP-HCSRI-----V-  
 CONSENSUS- E -S-----A-K-----G-R-R-EN-----L-N-----S-----GDR-----FPQ-----S-----  
 PAR -S-----A-K-----G-R-R-EN-----L-N-----S-----GDR-----FPQ-----S-----





# HIV2/SIV POL

	RNase H	\	Integrase	
CONSENSUS-A	VLFLEKLEPAAQEHKcYHsn?kEL?HKFgI?PdIVARQIVNtCaqcQQKGBAIIHGQVhnaIcHWQMDcTHLeGkIITVAVHVASGFIEAeVLPQesgRQTALFLKLASRWPTLTHLHTDNGANP?TSQEVKMK			893
ROD	V-S-S-N-S		S	871
NIH2	I-S-N-K		HV	870
ISY	R-G-A-L-N			870
ST	V-S-L-K		T	890
BEN	II-T-L		S	890
CAM2	T-V-C-D		Y	869
D194	I-T		I-V	908
GHI	R-M-T		I-V	890
UC2	I-I		I-V	890
MDS	V-S-L		D	890
KR	V-S-L-K		D	890
FA	R-I-T		I-V	187
5132P1	I-T-A		TK	187
S52907	N		TK	560
	--\$\$			
CONSENSUS-B	G-V-V		K-S-DK	887
D205	G-V-V		K-S-DK	889
UC1	G-V-V		L-K-S-DK	891
EHQACG	I-N-V		V-S-DK	890
JA	G-V-V		HS-K-S-DK	187
ON	G-V-V		K-S-DK	187
CONSENSUS-C	RF-A-T		N-K-HV	187
POLB7	RF-A-T		N-K-HV	187
CONSENSUS-D	D-D		I-VF-L-K-ID-DK-HR	696
FORTC2	D-D		I-VF-L-K-ID-DK-HR	658
POLC12	D-D		I-VF-L-K-ID-DK-HR	187
CONSENSUS-SD	?-V-V		L-R-K-D-DK-H	880
MM1A11AA	D-V		VF-L-RI-D-DK-H	895
MM251	D-V		VF-L-RI-D-DK-H	891
MM32H	D-V		VF-L-RI-D-DK-H	895
MM239	D-V		VF-L-RI-D-DK-H	895
MM142	S-I		VF-L-R-K-D-DK-H	891
MNE	D-V		VF-L-R-K-D-DK-H	891
SMMH4	K-V		VF-L-R-K-D-DK-H	854
SMMH9	V-K		VF-XL-R-K-D-DK-HL-x	857
SMMPBUI4.441	V-VF		L-R-K-D-DK-H	889
CONSENSUS-STM	V-VF		R-K-D-DK-H	876
STM	V-VF		R-K-D-DK-H	876
CONSENSUS-SO	V-VF		??-R-K-D-DK-?	234
SMM1IB1	I-VF		Y-R-K-D-DK-H	280
SMMM7	V-VF		R-R-K-D-DK-R-T-RI-I	282

CONSensus-A VAWMI:GIEQSFQVRYNPOGQVVEAMNHHKKNQI srr IReQANT?FTIYImAVHcMNFKRKRGIGDWTPaER1 INM1TEBQEIQPLqakNSK1knpFrVYfREGRdQUMKGPgELLmKGDGAV1VKVgtdIK  
 ROD ----T-----I-----I-----S-----D-----L-----E-----L----- 1022  
 NIH2 --V--T-----D-----V-----I-----I-----V-----V-----N-----Q----- 1001  
 ISY --V--V-----E-----M-----V-----V-----V-----A----- 1000  
 ST ----T-----V-----I-----R-----NP-----Q-----Y----- 1020  
 BEN --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1020  
 CAM2 --V--T-----I-----V-----S-----R-----NP-----Q-----Y----- 999  
 D1.94 --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1038  
 GH1 --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1020  
 UC2 --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1020  
 MDS --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1020  
 KR --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 1020  
 FA --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 237  
 5132P1 --V--T-----I-----V-----S-----R-----NP-----K-----Q-----Y----- 237

CONSensus-B --V--T-----D-----D-----VSI-----V-----V-----F-----L-----FQ-----q-----Y----- 1017  
 D205 --V--T-----D-----L-----D-----VSI-----V-----T-----F-----L-----FQ-----Q-----Y----- 1019  
 UC1 --V--T-----D-----D-----VSI-----V-----A-----F-----L-----FQ-----Q-----Y----- 1021  
 EHOACG --V--T-----E-----D-----D-----VSI-----V-----T-----IV-----T-----L-----FQ-----Y----- 1020  
 JA --V--RT-----D-----D-----VSI-----V-----X-----D-----D-----VSL-----V-----X----- 237  
 ON --V--T-----D-----D-----VSL-----V-----X----- 237

CONSensus-C --V--V--HT-----D-----M-----M-----X----- 237  
 POLB7 --V--V--HT-----D-----M-----M-----X----- 237

CONSensus-D --V--T-----DK-----D-----SV-----X----- 746  
 POLC12 --V--T-----DK-----D-----SV-----X----- 237

CONSensus-SD --A--HT-----D-----S-----S-----T-----?-----Q-----S-----F-----Y----- 1008  
 MM1A11AA --A--HT-----D-----SV-----S-----T-----?-----Q-----S-----F-----Y----- 1025  
 MM251 --A--HT-----D-----SV-----S-----T-----?-----Q-----S-----F-----Y----- 1021  
 MM32H --A--HT-----D-----SV-----S-----T-----?-----Q-----S-----F-----Y----- 1025  
 MM239 --A--HT-----D-----SV-----S-----T-----?-----Q-----S-----F-----Y----- 1025  
 MM142 --A--HT-----D-----SV-----S-----T-----?-----Q-----S-----F-----Y----- 1021  
 MNE --A--HT-----D-----SM-----L-----L-----Q-----S-----F-----Y----- 1021  
 SMMH4 --A--T-----T-----SI-----I-----I-----Q-----S-----F-----Y----- 984  
 SMMH9 --AX--T-----x-----SI-----I-----I-----Q-----S-----F-----Y----- 987  
 SMMPB14.441 --A--T-----T-----SI-----I-----I-----Q-----S-----F-----Y----- 1019

CONSensus-STM --A--T-----T-----D-----D-----V-----V-----K-----L-----V-----Q-----S-----F-----Y----- 1006  
 STM --A--T-----T-----D-----D-----V-----V-----K-----L-----V-----Q-----S-----F-----Y----- 1006

CONSensus-SO --?--A?--?T--?-----?-----?-----?-----?-----Q-----S-----F-----?-----E----- 340  
 SMM1IB1 --AN--T-----T-----DK-----D-----SI-----T-----V-----Q-----S-----F-----Y----- 395  
 SMMM7 --\$--AD--HT--K-----R-----E-----SV-----Q-----S-----F-----Y----- 401

HIV2/SIV POL

Integrase \

CONSSENSUS-A vVPRRKAk1I rDYgR qe?DdS?ghL RgAR E?dGEVA?PQVPELIQMKRRRGALCSPPQGGMGWMDLQOGNI PTRKKS SRNTGILLEPWRKRMALLSCSKTINLVYRKVLDRCYPRLCRHPVT 1141  
 ROD II-----M-G-----M- 1036  
 NIHZ K---M-G-----M- 1035  
 ISY -I-----P--M-G-----M- 1035  
 ST II-----K---M-G-N-----M-C----- 1055  
 BEN L--SP-----L-S-----N----- 1142  
 CAM2 II-----L-S-----E----- 1034  
 D194 -I-----L-----E----- 1073  
 GH1 -I-----L-----E----- 1055  
 UC2 I-----L--SP-----T----- 1055  
 MDS I-----DM--GP-----T----- 1055  
 KR I-----R-V--S-----T----- 1055

CONSSENSUS-B -I-----h---GK-L-CsAdm-DT-Q.AR-m-QS??--A 1058  
 D205 -I-----H---GKGL-CSADM-DT-Q.AR-M-QSD 1057  
 UC1 -I-----H---GK-L-CGTTDM-DT-Q.AR-M-QSG-----A 1064  
 EHOCAG -I-----N---GK-L-CSADV-DTWQ.AR--QSN\$ 1058

CONSSENSUS-SD -----K---GK-----s--m-DTG-.AR--\$ 1042  
 MML11AA -----K---GK-M-S--M-DTG-.AR--\$ 1060  
 MM251 -----K---GK-V-S--M-DTG-.AR-- 1056  
 MM32H -----K---GK-V-S--M-DTG-.AR-- 1060  
 MM239 -----K---GK-V-S--M-DTG-.AR-- 1056  
 MML42 -----K---GK-M-S--M-DTG-.AR-- 1056  
 MNE -----K---GK-V-S--M-DTG-.AR-- 1056  
 SMMH4 -----K---GK-L-G--DTG-.AR-- 1019  
 SMMH9 -----K---GK-L-G--DTG-.AR-- 1022  
 SMMPB14.441 -----K---GK-L-G--DTG-.AR--\$ 1054

CONSSENSUS-STM -----K---GK-V--G--DT--.A----- 1041  
 STM -----K---GK-V--G--DT--.A----- 1041