

Protein Alignment Summary Table (Pol)

Name(s)	Accession No.	Region	First Author	Reference
ROD	M15390	complete genome	Guyader M	<i>Nature</i> <b>326</b> (6114), 662–669 (1987)
NIHZ	J03654	complete genome	Zagury JF	<i>PNAS</i> <b>85</b> 5941–5945 (1988)
ISY	J04498	complete genome	Franchini G	<i>PNAS</i> <b>86</b> 2433–2437 (1989)
ST	M31113	complete genome	Kumar P	<i>J Virol</i> <b>64</b> 890–901 (1990)
BEN	M30502	complete genome	Kirchhoff F	<i>Virology</i> <b>177</b> 305–311 (1990)
CAM2	D00835	complete genome	Tristem MF	<i>J Gen Virol</i> <b>72</b> 721 (1991)
D194	J04542	complete genome	Kuehnel H	<i>PNAS</i> <b>86</b> 2383–2387 (1989)
GHI	M30895	complete genome	Hasegawa A	<i>ARHR</i> <b>5</b> 593–604 (1989)
UC2	U38293	complete genome	Barnett SW	Unpublished (1997)
MDS	Z48731	complete genome	Becker M	Unpublished (1995)
U22047 KR	U22047	gag pol vif tat rev nef env	Kraus GK	<i>PNAS</i> <b>90</b> (9), 4226–4230 (1993)
FA2	L33089	complete genome	Gao F	<i>J Virol</i> <b>68</b> 7433–7477 (1994)
5132P1	L16970	partial pol integrase	Gao F	<i>ARHR</i> <b>9</b> 703–704 (1993)
S52907	S52907	pol partial	Warren TC	<i>Protein Expr Purif</i> <b>3</b> 479–487 (1992)
HIW2ALT HIV2D205	X61240	pol RT region	Dietrich U	<i>ARHR</i> <b>8</b> 1619 (1992)
UC1	L07625	complete genome	Barnett SW	<i>J Virol</i> <b>67</b> 1006–1014, (1993)
EHOA	U27200	complete genome	Rey-Cuille MA	<i>Virology</i> <b>202</b> (1), 471–476 (1994)
JA2	L33090	complete genome	Gao F	<i>J Virol</i> <b>68</b> 7433–7477 (1994)
ONFT2	L33091	partial pol integrase	Gao F	<i>J Virol</i> <b>68</b> 7433–7477 (1994)
22POLB7	M87138	partial pol	Gao F	<i>Nature</i> <b>358</b> 495–499 (1992)
FORTC2	M87111	pol, RT region, partial	Gao F	<i>Nature</i> <b>358</b> 495–499 (1992)
POLC12	M87110	pol, partial integrase	Gao F	<i>Nature</i> <b>358</b> 495–499 (1992)
MM1A11	M76764	complete genome	Luciw PA	<i>ARHR</i> <b>8</b> 395–402 (1992)
MM251	M19499	complete genome	Franchini G	<i>Nature</i> <b>328</b> 539–543 (1987)
MM32H	D01065	complete genome	Rud EW	<i>J Gen Virol</i> <b>75</b> 529 (1994)

MM239	M33262	complete genome	Regier DA	<i>ARHR</i> <b>6</b> 1221–1231 (1990)
MM142	M16403	complete genome	Chakrabarti L	<i>Nature</i> <b>328</b> 543–547 (1987)
MNE	M32741	complete genome	Benveniste RE	Unpublished (1990)
SMMH4	X14307	complete genome	Hirsch VM	<i>Nature</i> <b>339</b> 389–391 (1989)
SIVSMMH9	M80194	complete genome	Courgnaud V	<i>J Virol</i> <b>66</b> 414–419 (1992)
SIVSMMPIA	M31325	complete genome	Dewhurst S	<i>Nature</i> <b>345</b> 636–640 (1990)
STM	M83293	complete genome	Novembre FJ	<i>Virology</i> <b>186</b> 783–787 (1992)
SMMLIB1	M62651	pol	Marx PA	<i>J Virol</i> <b>65</b> 4480–4485 (1991)
SMMM7	M27256	pol	Li Y	Unpublished (1989)

II-B-9  
DEC 98





```

CONSENSUS-S-A
A.ROD
A.NIH2
A.ISY
A.ST
A.BEN
A.CAM2
A.D194
A.GH1
A.UC2
A.MDS
A.KR
A.S52907
CONSENSUS-B
B.D205
B.UC1
B.EHOA
CONSENSUS-D
D.FORTC2
CONSENSUS-SD
SD_MM1A11AA
SD_MM251
SD_MM32H
SD_MM239
SD_MM142
SD_MNE
SD_SMMH4
SD_SMMH9
SD_SMMPB14_441.A
SD_SMMPB14_15.B
SD_SMMPB16_6
CONSENSUS-STM
STM_STM
CONSENSUS-AE
AE_PAR
CONSENSUS-AB
AB_7312AR

```

ÖKLVGVINMWAQI YpGIKTKhLcrlIrGKWtI TEEVQWTELAFAETLEENP; ILSQeQHGHYYEERKELEAIVGQdOhQwCYk1HQ?ekIHKVgyakIknLthngvYrLLAqVWVKIGKE

II-B-12  
DEC 98

```

CONSENSUS-A
A.ROD
A.NIH2
A.ISY
A.ST
A.BEN
A.CAM2
A.D194
A.GH1
A.UC2
A.MDS
A.KR
A.S52907
CONSENSUS-B
B.D205
B.UC1
B.EHOA
CONSENSUS-D
D.FORTC2
CONSENSUS-SD
SD_MM1A11AA
SD_MM251
SD_MM32H
SD_MM239
SD_MM142
SD_MNE
SD_SMMH4
SD_SMMH9
SD_SMMPB14_441.A
SD_SMMPB14_15.B
SD_SMMPB16_6
CONSENSUS-STM
STM_STM
CONSENSUS-SO
SO_SMMH7
SO_SMM1IB1

```

ALVTWGRIPKfHLPVBRReIWEQWMDNYWQVYTWLpDMDFVSTPPLVLAfNlVgdIlgpGefRfYTDGSCNpKQeKAGYvTDpGcDKV?YLEDTfLNOQAELEaFaMaLlDSGPKVNIIVD

Reverse Transcriptase \ RNase H

```

586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```



# Amino Acids

CONSENSUS -A	GTWQMDCTHLEGRKLIIVAVHVASGFLTEAEVLPQesRQTFALFLKLIASHRMPITTHLHTDNGANFTSOEVKVMWAMWIGIROSFGVPEYNPQSGGVWBAAMNHHLKNOIst rReQANT?ETITVIm	942
A.ROD	-----V-----	942
A.NIHZ	-----V-----	921
A.ISY	-----V-----	920
A.ST	-----V-----	920
A.BEN	-----V-----	940
A.CAM2	-----V-----	940
A.D194	-----V-----	919
A.GH1	-----V-----	958
A.UC2	-----V-----	940
A.MDS	-----V-----	940
A.KR	-----V-----	940
A.FA2	-----V-----	940
A.5132P1	-----V-----	237
CONSENSUS-B	-----V-----	237
B.D205	-----V-----	937
B.UC1	-----V-----	939
B.EHQA	-----V-----	941
B.TA2	-----V-----	940
B.ONFT2	-----V-----	237
CONSENSUS-C	-----V-----	237
C.22POLIB7	-----V-----	237
CONSENSUS-D	-----V-----	746
D.FORFC2	-----V-----	658
D.POLC12	-----V-----	237
CONSENSUS-SD	-----V-----	237
SD_MM1A11AA	-----V-----	929
SD_MM251	-----V-----	945
SD_MM32H	-----V-----	941
SD_MM239	-----V-----	945
SD_MM142	-----V-----	945
SD_MMNB	-----V-----	941
ST1rcm	-----V-----	941
SD_SWMH4	-----V-----	90
SD_SWMH9	-----V-----	904
SD_SWMPB14.441.A	-----V-----	907
SD_SWMPB14.15.B	-----V-----	939
SD_SWMPB16.6	-----V-----	939
CONSENSUS-STM	-----V-----	907
STM_STM	-----V-----	907
CONSENSUS-SO	-----V-----	926
SO_SMMW7	-----V-----	926
CONSENSUS-AB	-----V-----	331
AB_712AP	-----V-----	330

II-B-14  
DEC 98

CONSENSUS -A AVHGMNFKRRGGIGDMTPAERLIINMIITTEHOELQFLGAKNSKLNPFVYFREGROQIMKGPGEHLLMKGGAAVIKVGtdIKvpprKAKIIRpdyGgrge?DS?shlEGARE?dGEVVA?PCQ Integrase \

A. ROD S-----S-----R--NF-K-Q-Y- L-S-----L-S-----N-

A.NIHZ S-----I-----D-----L-----II-----M-G-----M-

A.ISY V--V--V--V--N--O-----P-M-G-----M-

A.ST V--A-----R--NF--Q--Y- L-SP-----M-C---

A.BEN S-----R--NF--Q--Y- L-SP-----M-C---

A.CAM2 S-----R--NF-K-Q-Y- L-S-----L-S-----N-

A.D194 S-----R--NF--Q--Y- L-SP-----M-G-----M-

A.GH1 R--NF-K-Q-Y- L-S-----L-S-----N-

A.UC2 S-----R--NF-K-Q-Y- L-SP-----M-G-----M-

A.MDS S-----H-----R-V-S-----T-----

A.KR S-----H-----R-V-S-----T-----

CONSENSUS -B T-----F--L--FO--q--Y- E--I--E--I--h--GK-L--CSAdm--DT--Q--AR--m--QS??

B.D205 T-----F--L--FO--Q--Y- E--I--E--I--E--H--GKGL--CSADM--DT--Q--AR--M--QSD

B.UC1 A-----F--L--FO--Q--Y- E--LI--E--I--I--E--H--GK-L--CGTDM--DT--Q--AR--M--QSG-

B.EHQA T-----IV-----T--L--FO--Y- E--D--I--E--I--N--GK-L--CSADV--DTMG--AR--QSN\$

CONSENSUS -SD S-----Q-S--F--Y- E--L-----K--GK--m--s--m--DTG--AR--\$

SD\_MMI111AA S-----Q-S--F--Y- E--L-----K--GK--V--S--M--DTG--AR--\$

SD\_MM251 S-----Q-S--F--Y- E--L-----K--GK--V--S--M--DTG--AR--\$

SD\_MM32H S-----Q-S--F--Y- E--L-----K--GK--V--S--M--DTG--AR--\$

SD\_MM32H S-----Q-S--F--Y- E--L-----K--GK--V--S--M--DTG--AR--\$

SD\_MM142 S-----Q-S--F--Y- E--L-----K--GK--M--S--M--DTG--AR--\$

SD\_MME S-----L--DI--A--DIATMK--NOI--IQ--Y- E--VL-----K--GK--V--S--M--DTG--AR--\$

STVrcm S-----Q-S--F--Y- E--L-----K--GK--I--I--E-----K--GK--L--G--DTG--AR--\$

SD\_SMMH4 S-----Q-S--F--Y- E--L-----K--GK--L--G--DTG--AR--\$

SD\_SMMH9 S-----Q-S--F--Y- E--L-----K--GK--L--G--DTG--AR--\$

SD\_SMMPB14\_441.A S-----Q-S--F--Y- E--L-----K--GK--L--G--DTG--AR--\$

SD\_SMMPB14\_15.B S-----Q-S--F--Y- E--L-----K--GK--L--G--DTG--AR--\$

SD\_SMMPB16\_6 S-----Q-S--F--Y- E--L-----K--GK--L--G--DTG--AR--\$

CONSENSUS -STM S-----Q-S--F--Y- E--L-----K--GK--V--G--DT--A--

STM\_STM S-----Q-S--F--Y- E--L-----K--GK--V--G--DT--A--

CONSENSUS -SO S-----Q-S--F--Y- E--L-----K--GK--V--G--DT--A--

SO\_SMM7 S-----Q-S--F--Y- E--L-----K--GK--V--G--DT--A--

SO\_SMM11B1 S-----Q-S--F--Y- E--L-----K--GK--V--G--DT--A--

CONSENSUS -A VPEIQNKRRPRGALCSPPOGGMGVDDLQOGNIPITTRKSSRNTGILIEPNTRRMALLSCSKINLVYRRKVLDRCCVPRLCRHNT

A. BEN S-----A-----

CONSENSUS -B S-----A-----

B\_UC1 S-----A-----

1058  
1036  
1035  
1035  
1055  
1054  
1057  
1060  
1058  
1042  
1060  
1056  
1060  
1060  
1056  
1056  
158  
1019  
1022  
1054  
1054  
1041  
1041  
340  
401  
395

II-B-15  
DEC 98