

Supertype restriction	HIV dataset	HIV ^{EpiJen} dataset
A1	ISERILSTY	
A1	RRGWEVLKY	
A1		ISERILGTY
A1		RRGWEALKY
A1	WIYHTQGYF	WIYHTQGYF
A1	YAPPISGQI	
A1	YFPDWQNYT	
A2	AALDLSHFL	
A2	AAVDLSHFL	AAVDLSHFL
A2	AEWDRLHPV	
A2		AEWDRVHPV
A2	AIIRILQQL	AIIRILQQL
A2		AITSSNTAA
A2	ALAALITPK	ALAALITPK
A2	ALHHVAREL	
A2	ALKHRAYEL	
A2	ALQDSGLEV	ALQDSGLEV
A2	ALTSSNTAA	
A2	ALVEICTEL	
A2	ALVEICTEM	ALVEICTEM
A2	ATLEEMMTA	ATLEEMMTA
A2	ATNADCAWL	
A2		ATNAACAWL
A2	AVLSVVNRV	
A2		AVLSIVNRV
A2	DTVLEDINL	
A2	EIYKRWIIL	EIYKRWIIL
A2	EMMTACQGV	
A2	ETAYFILKL	
A2	FIMIVGGLV	FIMIVGGLV
A2	FPRPWLHGL	
A2	GLADQLIHL	
A2	GMGSPQILV	
A2	HIGPGRAFY	
A2		FPRIWLHGL
A2		GLGQHIYET
A2		GVGSPQILV
A2	IISLWDQSL	IISLWDQSL
A2	ILKDPVHGV	
A2	ILRIPVHGV	
A2	ILRSCNTSV	
A2	ILVESPAVL	
A2	KAACWWAGI	
A2		ILVESPTVL
A2	KIEELRQHL	KIEELRQHL
A2	KIKPPLPSV	KIKPPLPSV
A2	KLKAGYVT	KLKAGYVT
A2	KLTPLCVSL	KLTPLCVSL
A2	KLTSNTSV	KLTSNTSV
A2	KLVGKLNWA	KLVGKLNWA
A2	KVIEVAQGA	
A2	LEWRFDSSL	LEWRFDSSL
A2	LLDTIAIAV	
A2	LLNATAIAV	LLNATAIAV

A2	LLQLTVWGI	LLQLTVWGI
A2	LLQYSQEL	LLQYSQEL
A2	LLWKGEAV	LLWKGEAV
A2	LQLPIERL	
A2	LTFGWCFKL	
A2		LQLPPLERL
A2	LTFGWCYKL	LTFGWCYKL
A2	LVGPTVNI	LVGPTVNI
A2	LWVTVYGV	LWVTVYGV
A2	MTNNPPIPV	MTNNPPIPV
A2	MTSNPPIPV	
A2	MTYKAALDL	
A2	NIWATHACV	
A2		MTYKAAVDL
A2	NTVATLYCV	NTVATLYCV
A2		NVWATHACV
A2	PLVKLWYQL	PLVKLWYQL
A2	QMHEDIISL	QMHEDIISL
A2		RAACWWAGI
A2	RAMASDFNL	RAMASDFNL
A2	RILQQLFI	RILQQLFI
A2	RIRQGLERA	
A2		RIRQGLERI
A2	RLRDLLLIV	RLRDLLLIV
A2	RLVNGSLAL	RLVNGSLAL
A2	RLVSGFLAL	
A2	RMYSPTSIL	RMYSPTSIL
A2	RTLNAWVKV	RTLNAWVKV
A2	RVIEVLQRA	
A2	SLFNTVATL	
A2	SLGQHIYET	
A2		RVIEVVQGA
A2	SLVKHHMYV	SLVKHHMYV
A2	SLYNTIATL	
A2	TLNAWVKVI	
A2	TLNFPISPI	
A2	TLQEQIGWM	TLQEQIGWM
A2	TLTSCNTSV	
A2	TMGAASITL	
A2	VLAEAMSQV	
A2		TMGAASMTL
A2		VTLWQRPLV
A2	VTVYGVVPV	VTVYGVVPV
A2	VVAIIAIV	
A2	WLWYIKIFI	
A2	YIKIFIMIV	
A2		VVAIIAIV
A2		WLWYIKLFI
A2		YIKLFIMIV
A2	YLAWVPAHK	YLAWVPAHK
A2	YPLTFGWCF	
A2		YPLTFGWCY
A2	YQYMDDLIV	YQYMDDLIV
A2	YTAFTIPSV	
A2	YVDRFYKTL	YVDRFYKTL
A3	AFHHVAREK	

A3		AFHHVAREL
A3	ALVEICTEM	ALVEICTEM
A3	AQNPEIVIIY	
A3	AVFIHNFKR	AVFIHNFKR
A3	DLSHFLKEK	DLSHFLKEK
A3	EYRKILRQR	
A3	FVNTPLVK	
A3	FYVDGAANR	
A3	GIPHPAGLK	GIPHPAGLK
A3	HIGPGRIFY	
A3	HMYISKKAK	
A3		HMYVSGKAR
A3	HSFNCGGEF	HSFNCGGEF
A3	ILGLNKIVR	ILGLNKIVR
A3	ISLWDQSLK	
A3	IVNRVRQGY	
A3	KILYQSNPY	
A3	KIQNFRVYY	
A3		KLLYQSNPP
A3	KLTEDRWNK	KLTEDRWNK
A3	KQNPDIIVY	
A3	KVYLAWVPA	KVYLAWVPA
A3	KYCWNLLQY	
A3	MAVFIHNFK	MAVFIHNFK
A3	MTKILEPFR	MTKILEPFR
A3	MVHQAI SPR	MVHQAI SPR
A3	NTPVFAIKK	NTPVFAIKK
A3	PLTFGW CYK	PLTFGW CYK
A3	QIIEQLIKK	QIIEQLIKK
A3	QIYAGIKVK	
A3	QIYPGIKVR	QIYPGIKVR
A3	RILSTYLGR	
A3	RVLKQVTEK	
A3	SIFQSSMTK	
A3		RILGTYLGR
A3	SLWDQSLKP	SLWDQSLKP
A3	SVITQACPK	
A3	TACNNCYCK	
A3		TACTNCYCK
A3	TIKIGGQLK	TIKIGGQLK
A3	TLFCASDAK	TLFCASDAK
A3	TLYCVHQKI	
A3	TRYPLTFGW	
A3	VSFEPPIH	VSFEPPIH
A3	YLAWVPAHK	YLAWVPAHK

Supplementary table: Comparing the HIV and HIV^{EpiJen} datasets.

	Epitope restricted to the A1 supertype
	Epitope restricted to the A2 supertype
	Epitope restricted to the A3 supertype
	Epitope is only found in the HIV dataset
	Epitope is only found in the HIV ^{EpiJen} dataset
	Epitope is found in both datasets