

Table S1

1/2

siRNA name	Sense strand	Antisense strand	siRNA efficacy prediction Ui-Tei, Reynolds, Amarzguiou			Relative luc activity (%)
siFL-5*	GAGGCCAAAAACAUAAGAAA	UCUUUAUGUUUUUGGCGUCUU	yes	yes	yes	17
siFL-6*	ACGCCAAAAACAUAAGAAA	UUCUUUAUGUUUUUGGCGUCUU			yes	7
siFL-12*	AAAACAUAAGAAAGGCCCGG	GGCCUUUCUUUAUGUUUUUG				93
siFL-14*	AACAUAAGAAAGGCCCGCG	CCGGCCUUUCUUUAUGUUUU				81
siFL-18*	UAAAGAAAGGCCCGGCCCAU	GGCGCCGGGCCUUUCUUUAUG				93
siFL-24*	AAGGCCCGGCCCAUUCUAUC	UAGAAUGGCGCCGGCCUUUC				83
siFL-29	CCGGCCCAUUCUAUCCGCUG	GCGGAUAGAAUGGCGCCGGC				47
siFL-30*	CGGCGCAUUCUAUCCGCUGG	AGCGGAUAGAAUGGCGCCGGG			yes	54
siFL-31	GGCGCAUUCUAUCCGCUGGA	CAGCGGAUAGAAUGGCGCCGG				22
siFL-33	CGCAUUCUAUCCGCUGGAAG	UCCAGCGGAUAGAAUGGCGCC			yes	14
siFL-34	GCCAUUCUAUCCGCUGGAAGA	UCCAGCGGAUAGAAUGGCGC			yes	31
siFL-35	CCAUUCUAUCCGCUGGAAGAU	CUCCAGCGGAUAGAAUGGCG				18
siFL-36*	CAUUCUAUCCGCUGGAAGAUG	UCUCCAGCGGAUAGAAUGGC	yes			11
siFL-37	AUUCUAUCCGCUGGAAGAUGG	AUCUCCAGCGGAUAGAAUGG				44
siFL-40	CUAUCCGCUGGAAGAUGGAAC	UCCAUCUCCAGCGGAUAGAA	yes	yes		9
siFL-42*	AUCCGCUGGAAGAUGGAACCG	GUUCCAUCUCCAGCGGAUAG				87
siFL-47	CUGGAAGAUGGAACCGCUGGA	CAGCGGUUCCAUCUCCAGCG				17
siFL-48*	UGGAAGAUGGAACCGCUGGAG	CCAGCGGUUCCAUCUCCAGC				81
siFL-49	GGAAGAUGGAACCGCUGGAGA	UCCAGCGGUUCCAUCUCCAG			yes	8
siFL-53	GAUGGAACCGCUGGAGAGCAA	GCUCUCCAGCGGUUCCAUCUU				17
siFL-54*	AUGGAACCGCUGGAGAGCAAC	UGCUCUCCAGCGGUUCCAUCU				64
siFL-55	UGGAACCGCUGGAGAGCAACU	UUGCUCUCCAGCGGUUCCAUC				9
siFL-179	GUUCGGUUGGCAGAAGCUAUG	UAGCUUCUGCCAACCGAACGG	yes			5
siFL-184*	GUUGGCAGAAGCUAUGAAACG	UUUCAUAGCUUCUGCCAACCG	yes	yes	yes	6
siFL-257	CAAUUCUUUAUGCCGUGUUUG	ACACCGCAUAAAGAAUUGAA				53
siFL-265*	UAUGCCGGUUGGCGCGUUUG	CGGCCCCAACACCGCAUAAA				75
siFL-272*	GUGUUGGCGCGUUUAUUUAUC	UAAUAUACGCGCCAACACCG	yes	yes	yes	13
siFL-295*	AGUUGCAGUUGCGCCCGCGAA	CGCGGGCGCAACUGCAACUCC				68
siFL-296	GUUGCAGUUGCGCCCGCGAAC	UCGCGGGCGCAACUGCAACUC				23
siFL-309*	CCGCGAACGACAUUUUAUUAUG	UUUAUUAUGUCGUUCGCGGGC	yes	yes	yes	6
siFL-329	GAACGUGAAUUGCUAACAGU	UGUUGAGCAAUUCAGUUCAU	yes	yes		7
siFL-370	CGUGGUGUUCGUUUCAAAAA	UUUGGAAACGAACACCAACGGU	yes		yes	3
siFL-387	AAAAGGGUUGCAAAAAAUUU	AUUUUUUGCAACCCUUUUUG		yes		5
siFL-411*	ACGUGCAAAAAAAGCUCCAA	GGGAGCUUUUUUUGCAGUUC				72
siFL-412	CGUGCAAAAAAAGCUCCAAU	UGGGAGCUUUUUUUGCAGUUC			yes	9
siFL-421	AAAGCUCCAAUCAUCCAAAA	UUUGAUGAUUGGGAGCUUUU		yes		21
siFL-423	AGCUCCAAUCAUCCAAAAA	UUUUGAUGAUUGGGAGCUUU		yes	yes	17
siFL-428*	CCAAUCAUCCAAAAAUUAUU	UAAUUUUUUGAUGAUUGGGA	yes	yes	yes	4
siFL-431	AUCAUCCAAAAAUUAUUUAUC	UAAUAUUUUUUGAUGAUUG		yes		63
siFL-507	CUCAUCUACCUCCCGUUUUUA	AAACCGGGAGGUAUGAGAU			yes	13
siFL-515*	CCUCCCGUUUUUAUGAAUAC	AUUCAUUAAAACCGGGAGGUA	yes	yes	yes	25
siFL-575	GCACUGAUCAUGAACUCUCU	AGGAGUUCUAUGAUCAGUGCAA	yes			9
siFL-583	CAUGAACUCUCUGGAUCUAC	AGAUCAGAGGAGUUCUAUGAU	yes		yes	26
siFL-630	CUCAUAGAACUGCCUGCGUGA	ACGCAGGAGUUCUAUGAGGC			yes	39
siFL-655	CUCGAUGCAGAGAUCCUAU	AGGAUCUCUGGCAUGCGAGAA	yes		yes	5
siFL-658*	GCAUGCCAGAGAUCCUAUUUU	AAUAGGAUCUCUGGCAUGCGA	yes	yes	yes	5
siFL-694	UCCGGAUACUGCGAUUUUAAG	UAAAAUCGAGUAUCCGGAAU		yes	yes	2
siFL-695*	CCGGAUACUGCGAUUUUAAGU	UUAAAAUCGAGUAUCCGGAA	yes	yes	yes	5
siFL-721	UCCAUCUACACGGUUUUGG	AAAACCGUGAUGGAAUGGAAC		yes		12
siFL-734*	GGUUUUGGAAUGUUUAUCACA	UAGUAAACAUCCAAAACCGU	yes		yes	6
siFL-749	ACUACACUCGGAUUAUUUGAUA	UCAAUAUCCGAGUGUAGUAA		yes		3
siFL-750	CUACACUCGGAUUAUUUGAUAU	AUCAAUAUCCGAGUGUAGUA	yes	yes		8
siFL-758	GGAUUUUGAUAUGUGGAUUU	AUCCACAUAUCAAUAUCCGA	yes		yes	2
siFL-765	UGAUAUGUGGAUUUCGAGUCG	ACUCGAAAUCCACAUAUCAA				19
siFL-766	GAUAUGUGGAUUUCGAGUCGU	GACUCGAAAUCCACAUAUCAA				50
siFL-767	AUAUGUGGAUUUCGAGUCGUC	CGACUCGAAAUCCACAUAUCAA				38
siFL-768	UAUGUGGAUUUCGAGUCGUCU	ACGACUCGAAAUCCACAUAUCAA				25
siFL-769	AUGUGGAUUUCGAGUCGUCUU	GACGACUCGAAAUCCACAUAU				14
siFL-770	UGUGGAUUUCGAGUCGUCUUA	AGACGACUCGAAAUCCACAUA				13
siFL-771	GUGGAUUUCGAGUCGUCUUA	AAGACGACUCGAAAUCCACAUA	yes		yes	2

Table S1

2/2

siRNA name	Sense strand	Antisense strand	siRNA efficacy prediction			Relative luc activity (%)
			Ui-Tei,	Reynolds,	Amarzguioui	
siFL-772	UGGAUUUCGAGUCGUCUUAU	UAAGACGACUCGAAAUCCACA		yes		3
siFL-773	GGAUUUCGAGUCGUCUUAUUG	UUAAGACGACUCGAAAUCCAC	yes	yes	yes	3
siFL-774*	GAUUUCGAGUCGUCUUAUUGU	AUUAAGACGACUCGAAAUCCA	yes	yes	yes	1
siFL-775	AUUUCGAGUCGUCUUAUUGUA	CAUUAAGACGACUCGAAAUCC				37
siFL-776	UUUCGAGUCGUCUUAUUGUAU	ACAUUAAGACGACUCGAAAUC				31
siFL-777	UUCGAGUCGUCUUAUUGUAUA	UACAUUAAGACGACUCGAAAU		yes		7
siFL-778	UCGAGUCGUCUUAUUGUAUAG	AUACAUUAAGACGACUCGAAA				15
siFL-779	CGAGUCGUCUUAUUGUAUAGA	UAUACAUUAAGACGACUCGAA	yes	yes	yes	1
siFL-780	GAGUCGUCUUAUUGUAUAGAU	CUAUACAUAUAGACGACUCGA				3
siFL-781	AGUCGUCUUAUUGUAUAGAUU	UCUAUACAUAUAGACGACUCG				35
siFL-782	GUCGUCUUAUUGUAUAGAUUU	AUCUAUACAUAUAGACGACUC	yes		yes	54
siFL-783	UCGUCUUAUUGUAUAGAUUUG	AAUCUAUACAUAUAGACGACU				6
siFL-857	CUGGUGCCAACCCUAUUCUCC	AGAAUAGGGUUGGCACCAGCA	yes		yes	24
siFL-891*	GCACUCUGAUUGACAAAUAACG	UAUUUGUCAUUCAGAGUGCUU	yes	yes	yes	5
siFL-903	ACAAAUACGAUUUAUCUAAUU	UUAGAUAUAUUCGUAUUUGUCA		yes		14
siFL-904*	CAAAUACGAUUUAUCUAAUUU	AUUAGAUAAAUCGUAUUUGUC	yes	yes	yes	13
siFL-998	CCAGGUUACAGGCAAGGAUUAU	AUCCUUGCCUGAUACCCUGGCA	yes		yes	5
siFL-1044*	UUCUGAUUACACCCGAGGGGG	CCCUCGGGUGUAUUCAGAAUA				75
siFL-1048	GAUUACACCCGAGGGGAUGA	AUCCCCUCGGGUGUAUUCAG				6
siFL-1054	ACCCGAGGGGAUGAUAUAAACC	UUUAUCAUCCCCUCGGGUGU		yes	yes	10
siFL-1063	GGAUUAUAAACCGGGCGCGGU	CGCGCCCGUUUAUCAUCCCC				6
siFL-1064*	GAUGAUAAACCGGGCGCGGUC	CCGCGCCCGUUUAUCAUCCCC				46
siFL-1065	AUGAUAAACCGGGCGCGGUCG	ACCGCGCCCGUUUAUCAUCC				15
siFL-1074	CGGGCGCGGUCGUAAGUUG	ACUUUACCGACCGCGCCCGGU	yes		yes	5
siFL-1075*	GGGCGCGGUCGUAAGUUGU	AACUUUACCGACCGCGCCCGG	yes		yes	8
siFL-1076	GCGCGGUCGUAAGUUGUU	CAACUUUACCGACCGCGCCCG				3
siFL-1081	GGUCGUAAGUUGUCCAUUUUU	UGGAACAACUUUACCGACCCG	yes			2
siFL-1085*	GGUAAAGUUGUCCAUUUUUUU	AAAAUGGAACAACUUUACCGA	yes		yes	5
siFL-1110	CGAAGGUUGUGGAUCUGGAUA	UCCAGAUCACAACCUUCGCU	yes	yes		1
siFL-1111	GAAGGUUGUGGAUCUGGAUAC	AUCCAGAUCACAACCUUCGCU	yes			4
siFL-1120	GGAUUCGGAUACCGGAAAAC	UUUCCCGGAUCCAGAUCACC		yes	yes	6
siFL-1137	AAACGUCGGCGUUAUCAA	UGAUUAACGCCAGCGUUUUUC		yes		13
siFL-1186*	GAUUUUGUCCGGUUAUGUAAA	UACAUAAACCGGACAUAAUCAU	yes	yes		2
siFL-1190	AUGUCCGGUUAUGUAAAACAU	UGUUUACAUAACCGGACAUAA		yes		103
siFL-1228	CUUGAUUGACAAGGAUGGAUG	UCCAUCUUGUCAAUCAAGGC				26
siFL-1234	UGACAAGGAUGGAUGGCUACA	UAGCCAUCCAUCCUUGUCAAU		yes		3
siFL-1308*	GCCUGAAGUCUCUGAUUAAGU	UUAAUCAGAGACUUCAGGCGG	yes	yes	yes	4
siFL-1346	GCUCCCGCUGAAUUGGAAUCC	AUUCCAAUUCAGCGGAGCCA	yes		yes	3
siFL-1352	GCUGAAUUGGAAUCCAUUUUG	AGAUGGAUUCCAAUUCAGCGG	yes		yes	11
siFL-1429	CGGUGAACUUCGCCCGCCCGU	GGCGGCGGGAAGUUCACCGGC				21
siFL-1430*	GGUGAACUUCGCCCGCCCGUU	CGGCGGCGGGAAGUUCACCGG				82
siFL-1431	GUGAACUUCGCCCGCCCGUUUG	ACGGCGGCGGGAAGUUCACCG				16
siFL-1431	GUGAACUUCGCCCGCCCGUUUG	ACGGCGGCGGGAAGUUCACCG				18
siFL-1439*	CCCGCCCGGUGUUGUUUUUG	AAACAACAACCGCGCGGGAA			yes	62
siFL-1440*	CCGCCCGGUGUUGUUUUUGG	AAAACAACAACCGCGCGGGGA	yes		yes	33
siFL-1441*	CGCCCGGUGUUGUUUUUGGA	CAAAACAACAACCGCGCGGGG				41
siFL-1482	AAAAAGAGAUCGUGGAUUACG	UAAUCCACGAUCUCUUUUUCC		yes		3
siFL-1509	GUCAAGUAACAACCGGAAAA	UUCGCGGUUGUUAUCUUGACUG			yes	42
siFL-1512	AAGUAAACAACCGGAAAAAGU	UUUUUCGCGGUUGUUAUCUUGA		yes	yes	12
siFL-1522	CGCGAAAAAGUUGCGCGGAGG	UCCGCGCAACUUUUUCGCGGU			yes	56
siFL-1539	GAGGAGUUGUUGUUGGACG	UCCACAAACAACAUCUCCUCCG	yes			2
siFL-1565	CCGAAAGGUCUUACCGAAAA	UUCCGGUAAGACCUUUCGGUA			yes	12
siFL-1586	CUCGACGCAAGAAAAUCAGA	UGAUUUUUUCUUGCGUCGAGUU	yes	yes	yes	12
siFL-1601	AUCAGAGAGAUCCUCAUAAAG	UUAUGAGGAUCUCUCUGAUUU		yes	yes	6

Table S2

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AF193275	97BL006	A1	BY	1997
AF457052	KER2008	A1	KE	2000
AF457053	KER2009	A1	KE	2000
AF457055	KER2012	A1	KE	2000
AF457068	KNH1207	A1	KE	2000
AF457069	KNH1209	A1	KE	2000
AF457070	KNH1211	A1	KE	2000
AF457077	KSM4024	A1	KE	2000
AF457079	KSM4030	A1	KE	2000
AF457080	MSA4069	A1	KE	2000
AF457081	MSA4070	A1	KE	2000
AF457083	MSA4072	A1	KE	2000
AF457084	MSA4076	A1	KE	2000
AF457086	MSA4079	A1	KE	2000
AF457089	NKU3005	A1	KE	2000
AF539405	ML170_1986	A1	KE	1986
AF004885	Q23_17	A1	KE	1994
AF457063	KNH1088	A1	KE	1999
AF457065	KNH1135	A1	KE	1999
AF457075	KSM4021	A1	KE	1999
AF069670	SE7253	A1	SE	1994
AF069671	SE7535	A1	SE	1994
AF069669	SE8538	A1	SE	1995
AF069673	SE8891	A1	SE	1995
AF107771	UGSE8131	A1	SE	1995
AF361872	97TZ02	A1	TZ	1997
AF361873	97TZ03	A1	TZ	1997
AF413987	98UA0116	A1	UA	2000
AB098333	UG029	A1	UG	-
AB098331	UG031	A1	UG	-
M62320	U455	A1	UG	1985
U51190	92UG037	A1	UG	1992
AF484507	98UG57134	A1	UG	1998
AF484508	98UG57135	A1	UG	1998
AF484509	98UG57136	A1	UG	1998
AF484512	98UG57142	A1	UG	1998
AF484478	99UGA07072	A1	UG	1999
AF484493	99UGG03379	A1	UG	1999
AF286240	97CDKFE4	A2	CD	1997
AF286241	97CDKS10	A2	CD	1997
AF286238	97CDKTB48	A2	CD	1997
AF286237	94CY017_41	A2	CY	1994
U23487	contaminant_MANC	B	-	-
AY037269	ARMS008	B	AR	2000
AY037268	ARCH054	B	AR	1998
AY037282	ARMA132	B	AR	1999
AF538302	1181	B	AU	-
AF538303	1182	B	AU	-
AF538304	C24	B	AU	-
AF538305	C42	B	AU	-

Table S2

<b>Accession</b>	<b>Name</b>	<b>Subtype</b>	<b>Country code (ISO3166)</b>	<b>Sampling year</b>
AF538306	C76	B	AU	-
AF538307	C92	B	AU	-
AF146728	VH	B	AU	-
AF042100	MBC200	B	AU	1986
AF042101	MBC925	B	AU	1987
AF042103	MBCC54	B	AU	1995
AF042104	MBCC98	B	AU	1996
AF042105	MBCD36	B	AU	1996
AY037270	BOL0122	B	BO	1999
AY173956	BZ167	B	BR	1989
AY180905	CNHN24	B	CN	-
U71182	RL42	B	CN	-
U43096	D31	B	DE	1986
U43141	HAN	B	DE	1986
AY173959	EC003	B	EC	1989
AY173960	EC102	B	EC	1989
AF256210	S61K15	B	ES	1989
K03455	HXB2	B	FR	1983
M26727	OYI	B	GA	1988
D10112	CAM1	B	GB	1983
AJ271445	GB8	B	GB	1986
AF224507	WK	B	KR	1997
AB097870	mSTD101	B	MM	1999
U34604	3202A21	B	NL	1986
AY173951	BK132	B	TH	1990
AF086817	TWCYS	B	TW	1994
AY268493	Ac_06	B	US	2001
M17451	RF	B	US	1983
K02007	SF2	B	US	1983
M17449	MNCG	B	US	1984
M38431	NY5CG	B	US	1984
AY352275	SF33	B	US	1984
AF004394	AD87	B	US	1986
U63632	JRFL	B	US	1986
M93258	YU2	B	US	1986
L02317	BC	B	US	1987
AF286365	WR27	B	US	1988
U39362	P896	B	US	1989
AY173952	US1	B	US	1990
AY173953	US2	B	US	1990
AY173954	US3	B	US	1990
AY173955	US4	B	US	1990
U69591	WCIPR	B	US	1990
U21135	WEAU160	B	US	1990
AF069140	DH12_3	B	US	1991
AY308761	1299_d22	B	US	1996
AY308760	1333_d2	B	US	1996
AB078005	ARES2	B	US	1997
AY247251	Ac_06	B	US	1997
U52953	92BR025	C	BR	1992

Table S2

3/10

<b>Accession</b>	<b>Name</b>	<b>Subtype</b>	<b>Country code (ISO3166)</b>	<b>Sampling year</b>
AF286228	98BR004	C	BR	1998
AF443088	00BW07621	C	BW	2000
AF443089	00BW076820	C	BW	2000
AF443090	00BW087421	C	BW	2000
AF443091	00BW147127	C	BW	2000
AF443092	00BW16162	C	BW	2000
AF443093	00BW1686	C	BW	2000
AF443094	00BW17593	C	BW	2000
AF443095	00BW17732	C	BW	2000
AF443096	00BW17835	C	BW	2000
AF443097	00BW17956	C	BW	2000
AF443098	00BW18113	C	BW	2000
AF443099	00BW18595	C	BW	2000
AF443100	00BW18802	C	BW	2000
AF443101	00BW192113	C	BW	2000
AF443102	00BW20361	C	BW	2000
AF443103	00BW20636	C	BW	2000
AF443104	00BW20872	C	BW	2000
AF443105	00BW2127214	C	BW	2000
AF443106	00BW21283	C	BW	2000
AF443107	00BW22767	C	BW	2000
AF443108	00BW38193	C	BW	2000
AF443109	00BW38428	C	BW	2000
AF443110	00BW38713	C	BW	2000
AF443111	00BW3876_9	C	BW	2000
AF443112	00BW3886_8	C	BW	2000
AF443113	00BW3891_6	C	BW	2000
AF443114	00BW3970_2	C	BW	2000
AF443115	00BW5031_1	C	BW	2000
AF110959	96BW01B03	C	BW	1996
AF110962	96BW0402	C	BW	1996
AF110967	96BW0502	C	BW	1996
AF290028	96BW06	C	BW	1996
AF110969	96BW1104	C	BW	1996
AF110972	96BW1210	C	BW	1996
AF110974	96BW15C02	C	BW	1996
AF110976	96BW16B01	C	BW	1996
AF110980	96BW17	C	BW	1996
AF443075	96BWM032	C	BW	1996
AF443074	96BWMO1_5	C	BW	1996
AF443076	98BWMC122	C	BW	1998
AF443077	98BWMC134	C	BW	1998
AF443078	98BWMC14A3	C	BW	1998
AF443079	98BWMO1410	C	BW	1998
AF443080	98BWMO18D5	C	BW	1998
AF443081	98BWMO36A5	C	BW	1998
AF443082	98BWMO37D5	C	BW	1998
AF443083	99BW393212	C	BW	1999
AF443084	99BW46424	C	BW	1999
AF443085	99BW4745	C	BW	1999

Table S2

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AF443086	99BW47547	C	BW	1999
AF443087	99BWMC168	C	BW	1999
U46016	ETH2220	C	ET	1986
AF286233	98IS002	C	IL	1998
AB023804	93IN101	C	IN	1993
AF067157	93IN904	C	IN	1993
AF067158	93IN905	C	IN	1993
AF067154	93IN999	C	IN	1993
AF067159	94IN11246	C	IN	1994
AF286223	94IN476	C	IN	1994
AF067155	95IN21068	C	IN	1995
AF286231	98IN012	C	IN	1998
AF286232	98IN022	C	IN	1998
AY049708	01IN565_10	C	IN	1999
AF457054	KER2010	C	KE	2000
AB097871	mIDU101_3	C	MM	1999
AF361874	97TZ04	C	TZ	1997
AF361875	97TZ05	C	TZ	1997
AF286234	98TZ013	C	TZ	1998
AF286235	98TZ017	C	TZ	1998
AY463228	J38MA	C	ZA	2000
AY463223	J54Ma	C	ZA	2001
AY118165	97ZA003	C	ZA	1997
AY118166	97ZA009	C	ZA	1997
AF286227	97ZA012	C	ZA	1997
AY158533	98ZA445	C	ZA	1998
AY158534	98ZA502	C	ZA	1998
AY158535	98ZA528	C	ZA	1998
AY162223	TV001	C	ZA	1998
AY162224	TV002	C	ZA	1998
AY162225	TV012	C	ZA	1998
AF411967	99ZACM9	C	ZA	1999
AY043176	CTSC2	C	ZA	1999
AY043173	DU151	C	ZA	1999
AY043175	DU422	C	ZA	1999
AY043174	DU179	C	ZA	1999
AY463229	1119MB	C	ZA	1999
AY463217	1134MB	C	ZA	2000
AY463218	1176MB	C	ZA	2000
AY463219	1192M3M	C	ZA	2000
AY463220	1195MB	C	ZA	2000
AY463221	1210MB	C	ZA	2000
AY463222	1228MB	C	ZA	2000
AY463224	1162MB	C	ZA	2000
AY463225	1170MB	C	ZA	2000
AY463236	1214MB	C	ZA	2000
AY463226	1217MB	C	ZA	2000
AY463227	1225MB	C	ZA	2000
AY463230	1165MB	C	ZA	2000
AY463231	1168MB	C	ZA	2000

Table S2

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AY463232	1171MB	C	ZA	2000
AY463233	1178MB	C	ZA	2000
AY463234	1197MB	C	ZA	2000
AY463235	2004MB	C	ZA	2001
AY463237	2134MB	C	ZA	2001
AF411966	ZASW7	C	ZA	1999
AF286224	96ZM651	C	ZM	1996
AF286225	96ZM751	C	ZM	1996
K03454	ELI	D	CD	1983
M27323	NDK	D	CD	1983
U88822	84ZR085	D	CD	1984
M22639	Z2Z6	D	CD	1985
AY371155	01CM_0009BBY	D	CM	2001
AY371156	01CM_0175BA	D	CM	2001
AY371157	01CM_4412HAL	D	CM	2001
AF133821	MB2059	D	KE	-
AF457090	NKU3006	D	KE	2001
AJ488926	MN011	D	TD	1999
AJ488927	MN012	D	TD	1999
AJ320484	92UG001	D	UG	1992
U88824	94UG114	D	UG	1994
AF484502	98UG57128	D	UG	1998
AF484504	98UG57130	D	UG	1998
AF484505	98UG57131	D	UG	1998
AF484506	98UG57132	D	UG	1998
AF484511	98UG57140	D	UG	1998
AF484514	98UG57143	D	UG	1998
AF484513	98UG57146	D	UG	1998
AF484518	99UGA03349	D	UG	1999
AF484477	99UGA07412	D	UG	1999
AY304496	99UGA08483	D	UG	1999
AF484480	99UGB21875	D	UG	1999
AF484481	99UGB25647	D	UG	1999
AF484483	99UGB32394	D	UG	1999
AF484485	99UGD23550	D	UG	1999
AF484486	99UGD26830	D	UG	1999
AF484487	99UGE08364	D	UG	1999
AF484515	99UGE13613	D	UG	1999
AF484489	99UGE23438	D	UG	1999
AF484519	99UGF03726	D	UG	1999
AF484490	99UGF05734	D	UG	1999
AF484494	99UGG10555	D	UG	1999
AF484495	99UGG35093	D	UG	1999
AF484497	99UGJ27597	D	UG	1999
AF484516	99UGJ32228	D	UG	1998
AF484498	99UGK09259	D	UG	1999
AF484499	99UGK09958	D	UG	1999
AF077336	VI850	F1	BE	1993
AY173957	BZ126	F1	BR	1989
AY173958	BZ163	F1	BR	1989

Table S2

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AF005494	93BR020_1	F1	BR	1993
AF075703	FIN9363	F1	FI	1993
AJ249238	MP411	F1	FR	1996
AY371158	02CM_0016BBY	F2	CM	2002
AJ249236	MP255	F2	CM	1995
AJ249237	MP257	F2	CM	1995
AF377956	CM53657	F2	CM	1997
AF084936	DRCBL	G	BE	1996
AY371121	01CM_4049HAN	G	CM	2001
AF423760	X558	G	ES	2000
AF450098	X138	G	ES	1999
AF061641	HH8793_12_1	G	FI	1993
U88826	92NG083	G	NG	1992
AF061642	SE6165	G	SE	1993
AF190127	VI991	H	BE	1993
AF190128	VI997	H	BE	1993
AF005496	056	H	CF	1990
AF082394	SE7887	J	SE	1993
AF082395	SE7022	J	SE	1994
AJ249235	EQTB11C	K	CD	1997
AJ249239	MP535	K	CM	1996
AF197340	90CF11697	01_AE	CF	1990
U51188	90CF402	01_AE	CF	1990
AF197341	90CF4071	01_AE	CF	1990
AY008718	97CNGX_11F	01_AE	CN	1997
AY008714	97CNGX2F	01_AE	CN	1997
AB052995	93JP_NH1	01_AE	JP	1993
AB070352	NH25	01_AE	JP	1993
AF259955	CM235	01_AE	TH	1990
U54771	CM240	01_AE	TH	1990
AF197338	93TH057	01_AE	TH	1993
AF197339	93TH065	01_AE	TH	1993
U51189	93TH253	01_AE	TH	1993
AF164485	93TH9021	01_AE	TH	1993
AB032740	95TNIH022	01_AE	TH	1995
AB032741	95TNIH047	01_AE	TH	1995
AY125894	97TH6_107	01_AE	TH	1997
AY371122	01CM_0002BBY	02_AG	CM	2001
AY371123	01CM_0005BBY	02_AG	CM	2001
AY371124	01CM_0008BBY	02_AG	CM	2001
AY371131	01CM_0074NY	02_AG	CM	2001
AY371137	01CM_0131NY	02_AG	CM	2001
AY371132	01CM_0158ND	02_AG	CM	2001
AY371133	01CM_0191ND	02_AG	CM	2001
AY371134	01CM_0925MO	02_AG	CM	2001
AY371136	01CM_1237NG	02_AG	CM	2001
AY371138	01CM_1475MV	02_AG	CM	2001
AY371142	01CM_4410HAL	02_AG	CM	2001
AY371125	02CM_0013BBY	02_AG	CM	2002
AY371126	02CM_0014BBY	02_AG	CM	2002



Table S2

<b>Accession</b>	<b>Name</b>	<b>Subtype</b>	<b>Country code (ISO3166)</b>	<b>Sampling year</b>
AY371127	02CM_0015BBY	02_AG	CM	2002
AY371139	02CM_1669LE	02_AG	CM	2002
AY371140	02CM_1677LE	02_AG	CM	2002
AY371146	02CM_1901LE	02_AG	CM	2002
AY371128	02CM_1970LE	02_AG	CM	2002
AY371129	02CM_2162SA	02_AG	CM	2002
AY371130	02CM_2348SA	02_AG	CM	2002
AY371141	02CM_4082STN	02_AG	CM	2002
AJ286133	97CM_MP807	02_AG	CM	1997
AF377954	CM52885	02_AG	CM	1997
AF377955	CM53658	02_AG	CM	1997
AY271690	pBD6_15	02_AG	CM	1999
AY151001	ECU41	02_AG	EC	-
AY151002	ECU42	02_AG	EC	-
AF063223	DJ263	02_AG	FR	1991
AF063224	DJ264	02_AG	FR	1991
AB049811	97GHAG1	02_AG	GH	1997
L39106	IBNG	02_AG	NG	-
AF107770	SE7812	02_AG	SE	1994
AJ251056	MP1211	02_AG	SN	1998
AJ251057	MP1213	02_AG	SN	1998
AF414006	98BY10443	03_AB	BY	2000
AF193276	KAL153	03_AB	RU	1997
AF193277	RU98001	03_AB	RU	1998
AF049337	CY032	04_cpx	CY	1994
AF119820	97PVCH	04_cpx	GR	1991
AF119819	97PVMY	04_cpx	GR	1997
AF193253	VI1310	05_DF	BE	-
AF076998	VI961	05_DF	BE	1993
AY227107	X492	05_DF	ES	1999
AF064699	BFP90	06_cpx	AU	1996
AJ288982	95ML127	06_cpx	ML	1995
AJ245481	95ML84	06_cpx	ML	1995
AJ288981	97SE1078	06_cpx	SN	1997
AF503396	CNGL179	07_BC	CN	-
AF286226	97CN001	07_BC	CN	1997
AX149771	CN54	07_BC	CN	1997
AF286230	98CN009	07_BC	CN	1998
AY008715	97CNGX_6F	08_BC	CN	1997
AY008716	97CNGX_7F	08_BC	CN	1997
AY008717	97CNGX_9F	08_BC	CN	1997
AF286229	98CN006	08_BC	CN	1998
AY093605	96GH2911	09_cpx	GH	1996
AY093603	95SN1795	09_cpx	SN	1995
AY093604	95SN7808	09_cpx	SN	1995
AY093607	99DE4057	09_cpx	US	1999
AF289548	96TZ_BF061	10_CD	TZ	1996
AF289549	96TZ_BF071	10_CD	TZ	1996
AF289550	96TZ_BF110	10_CD	TZ	1996
AY371149	01CM_0186ND	11_cpx	CM	2001

Table S2

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AY371150	01CM_4041HAN	11_cpx	CM	2001
AY371151	02CM_2190SA	11_cpx	CM	2002
AY371153	02CM_4118STN	11_cpx	CM	2002
AF492624	1816	11_cpx	CM	1995
AF492623	4496	11_cpx	CM	1996
AJ291718	MP818	11_cpx	CM	1997
AJ291719	MP1298	11_cpx	FR	1999
AJ291720	MP1307	11_cpx	FR	1999
AF179368	GR17	11_cpx	GR	-
AF408629	A32879	12_BF	AR	1997
AF408630	A32989	12_BF	AR	1997
AF385936	ARMA159	12_BF	AR	1999
AF385934	URTR23	12_BF	UY	1999
AF385935	URTR35	12_BF	UY	1999
AY371154	02CM_3226MN	13_cpx	CM	2002
AF460972	1849	13_cpx	CM	1996
AF460974	4164	13_cpx	CM	1996
AF423758	X475	14_BG	ES	2000
AF423759	X477	14_BG	ES	2000
AF450096	X605	14_BG	ES	2000
AF450097	X623	14_BG	ES	2000
AF423756	X397	14_BG	ES	1999
AF423757	X421	14_BG	ES	1999
AF529572	02TH_OUR1331	15_01B	TH	2002
AF529573	02TH_OUR1332	15_01B	TH	2002
AF530576	99TH_R2399	15_01B	TH	1999
AF516184	99TH_MU2079	15_01B	TH	1999
AF286239	97KR004	16_A2D	KR	1997
AF067156	95IN21301	A1C	IN	1995
AF457061	KISII5011	A1C	KE	2000
AF457087	MSA4080	A1C	KE	2000
AF539406	ML170_1995	A1C	KE	1995
AF457064	KNH1097	A1C	KE	1999
U88823	92RW009_06	A1C	RW	1992
AF071474	SE9488	A1C	SE	1996
AF361871	97TZ01	A1C	TZ	1997
AF361876	97TZ06	A1C	TZ	1997
AF361878	97TZ08	A1C	TZ	1997
AF361879	97TZ09	A1C	TZ	1997
AF484491	99UGF25926	A1C	UG	1999
AF484501	99UGK30889	A1C	UG	1999
AF457058	KER2021	A1D	KE	2000
AF457078	KSM4028	A1D	KE	2000
AF457082	MSA4071	A1D	KE	2000
AF457059	KISII5003	A1D	KE	1999
AF457073	KSM4015	A1D	KE	1999
AF075701	SE6954	A1D	SE	1993
AF071473	SE7108	A1D	SE	1994
AF442569	TZBFL0011	A1D	TZ	1996
AF442566	TZBFL0086	A1D	TZ	1997

Table S2

9/10

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AF442570	TZBFL0088	A1D	TZ	1997
AF484503	98UG57129	A1D	UG	1998
AF484521	98UG57137	A1D	UG	1998
AF484510	98UG57139	A1D	UG	1998
AF484522	98UG57147	A1D	UG	1998
AF484482	99UGB26587	A1D	UG	1999
AF484479	99UGC06443	A1D	UG	1999
AF484517	99UGC38442	A1D	UG	1999
AF484488	99UGE22831	A1D	UG	1999
AF484492	99UGF27390	A1D	UG	1999
AF484496	99UGJ21953	A1D	UG	1999
AF484520	99UGK38855	A1D	UG	1999
AY037267	ARCH003	BF	AR	2000
AF408631	A050	BF	AR	1997
AF408632	A32878	BF	AR	1997
AY037266	ARCH014	BF	AR	1998
AF408626	A025	BF	AR	1999
AF332867	A027	BF	AR	1999
AF408627	A047	BF	AR	1999
AF408628	A063	BF	AR	1999
AY037278	ARMA006	BF	AR	1999
AY037283	ARMA029	BF	AR	1999
AY037275	ARMA036	BF	AR	1999
AY037277	ARMA037	BF	AR	1999
AY037281	ARMA038	BF	AR	1999
AY037280	ARMA097	BF	AR	1999
AY037271	BOL0137	BF	BO	1999
AY037272	URTR17	BF	UY	1999
AB097865	mIDU502	01B	MM	2000
AB097867	mCSW104	01B	MM	1999
AF490973	1269	01B	TH	-
AY167123	CM237	01B	TH	-
AF362994	NP1623	01B	TH	-
AY082968	TH1326	01B	TH	-
AF468970	TH283	01B	TH	-
AF490974	TH9_95	01B	TH	-
AF457088	NKU3004	A1A2D	KE	2000
AF457056	KER2017	A1CD	KE	2000
AY371163	01CM_1152NG	A1U	CM	2001
AF411964	CM4	A1CDGKU	ZA	1999
AF075702	SE8603	A1CD	SE	1995
AJ237565	97NOGIL3	A1DHK	NO	1997
X04415	MAL	A1DK	CD	1985
AY371164	01CM_1404MV	A1F1	CM	2001
AF377959	CM53379	A1F1GHJU	CM	1997
AJ276596	VI1197	A1G	BE	1994
AF377957	CM53392	A1G	CM	1997
AF457062	KNH1043	A1G	KE	1999
U88825	92NG003	A1G	NG	1992
AF076474	VI354	A1GHU	GA	-

Table S2

10/10

Accession	Name	Subtype	Country code (ISO3166)	Sampling year
AF192135	BW2117	A1GJ	BW	1998
U76035	Z321	A1GU	CD	76
AF069672	SE6594	A1U	SE	1993
AF411965	DU178	A2C	ZA	1998
U86780	ZAM184	A2C	ZM	1990
AF457051	KER2003	A2D	KE	1999
AF457072	KSM4001	A2D	KE	1999
AF316544	97CDKP58	A2G	CD	1997
AY371168	01CM_4008HAN	AD	CM	2001
AY371160	02CM_3163MN	AF2	CM	2002
AY371166	01CM_0989MO	AGU	CM	2001
AY371169	02CM_1918LE	AGU	CM	2002
AY371161	01CM_1278NG	AHJU	CM	2001
AY371162	01CM_1296NG	AHJU	CM	2001
AP005206	HH069	BC	CN	2000
AP005207	HH086	BC	CN	2000
AB097873	mIDU103	BC	MM	1999
AB097869	mIDU106	BC	MM	1999
AF005495	93BR029_4	BF1	BR	1993
AF423755	X254	BG	ES	1999
AY074891	00BWMO351	CD	BW	2000
AF457085	MSA4077	CD	KE	2000
AF361877	97TZ07	CD	TZ	1997
AF484500	99UGK10192	CD	UG	1999
AF076475	VI1126	F2KU	BE	1994
AY371159	01CM_0001BBY	01A	CM	2001
AY371165	02CM_1867LE	01A	CM	2002
AY037284	CM53122	01A1	CM	1997
AB097872	mCSW105	01A1	MM	1999
AY371170	01CM_0908MO	01ADF2	CM	2001
AY371167	01CM_0130NY	01AF2U	CM	2001
AB097866	mCSW503	01BC	MM	2000
AB097868	mIDU107	01BC	MM	1999
AY262830	NP1809	01C	TH	-
AJ404325	97DCKTB49	01GHJKU	CD	1997
AY371145	01CM_0190MA	0102A	CM	2001
AF184155	G829	02A	GH	-
AB052867	AG2	02A1	GH	1997
AY371143	02CM_2339SA	02A1U	CM	2002
AJ276595	VI1035	02C	BE	1993
AY371147	02CM_3228MN	02G	CM	2002
AJ293865	B76	06A1	BJ	-
AF286236	83CD003	U	CD	1983
AF457101	90CD121E12	U	CD	1990
AY046058	GR303	U	GR	1999

Table S3

1/3

HXB2 coordinate	Target gene	Target site	Conservation (%)	Reference
64	U3, nef	AAGGCTACTTCCCTGATTGGC	34.5	Dave (2004)
67	U3, nef	GCTACTTCCCTGATTGGCAGAACTA	18.6	Yamamoto (2006)
72	U3, nef	TTCCCTGATTGGCAGAACTACACAC	17.9	Yamamoto (2006)
130	U3, nef	GATGGTGCTTCAAGCTAGTAC	28.3	Capodici (2002)
184	U3, nef	AAGGAGAGAACACCAGCTTGT	0.4	Chang (2005)
291	U3, nef	AAGGCCCGAGAGCTGCATCCGG	0.9	Suzuki (2005)
321	U3, nef	AAGACTGCTGACATCGAGCTT	1.1	Suzuki (2005)
348	U3	AAGGGACTTTCGCTGGGGAC	23.5	Suzuki (2005)
395	U3	AACTGGGGAGTGGCGAGCCCT	0.2	Suzuki (2005)
456	R, TAR	GTCTCTCTGGTTAGACCAGAT	16.2	Yoshinari (2004)
466	R, TAR	TTAGACCAGATCTGAGCCTGG	16.1	Yoshinari (2004)
468	R, TAR	AGACCAGATCTGAGCCTGG	16.1	Jacque (2002)
483	R, TAR	CTGGGAGCTCTCTGGCTAACT	9.5	Yoshinari (2004)
493	R, TAR	TCTGGCTAACTAGGGAACCCA	11.3	Yoshinari (2004)
515	R, poly A	TGCTTAAGCCTCAATAAAGCT	85.3	Yoshinari (2004)
551	R, U5	AAGTAGTGTGTGCCCGTCTGT	65.1	Chang (2005)
626	U5, PBS	TCTCTAGCAGTGGCGCCCG	92.0	Nishitsuji (2006)
634	PBS	GGTGGCGCCCGAACAGGGACCTT	17.0	Han (2004)
635	PBS	GTGGCGCCCGAACAGGGAC	90.1	Das (2004)
636	PBS	TGGCGCCCGAACAGGGACT	54.2	Das (2004)
769	U5, Ψ	GCGGAGGCTAGAAGGAGAG	88.1	Das (2004)
774	U5, Ψ, gag (AUG)	GGCTAGAAGGAGAGAGATG	91.0	Das (2004)
780	U5, gag (AUG)	AAGGAGAGAGATGGGTGCGAGAGCGTC	93.2	ter Brake (2006)
788	U5, gag (AUG)	AGATGGGTGCGAGAGCGTC	92.9	Westerhout (2006a)
896	gag (p17)	GGGCAAGCAGGGAGCTAGAAC	21.5	Capodici (2002)
896	gag (p17)	GGGCAAGCAGGGAGCTAGAACTT	21.5	Han (2004)
1475	gag (p24)	GAGAACCAAGGGGAAGTGACA	44.3	Capodici (2002)
1600	gag (p24)	AATAAAATAGTAAGAATGTA	77.3	ter Brake (2006)
1605	gag (p24)	AATAGTAAGAAATGTATAGCCC	75.9	Chang (2005)
1793	gag (p24)	AAGCATTGGGACCAGGAGCGA	0.0	Hu (2002)
1817	gag (p24)	TAGAAGAAATGATGACAGCATG	82.3	ter Brake (2006)
1820	gag (p24)	AAGAAATGATGACAGCATGTC	62.1	Chang (2005)
2060	gag (p7)	AAGATTGTACTGAGAGACAGG	9.3	Novina (2002)
2060	gag (p7)	AAGATTGTACTGAGAGACAGGCT	9.5	Suzuki (2005)
2062	gag (p7)	GATTGTACTGAGAGACAGG	9.3	Song (2003)
2062	gag (p7)	GATTGTACTGAGAGACAGG	9.3	Lee (2005)
2066	gag (p7)	GTACTGAGAGACAGGCTAA	9.1	Pusch (2003)
2077	gag (p7,p1), pol, reboosomal slip site	CAGGCTAATTTTTTAGGGAA	70.6	ter Brake (2006)
2167	gag (p6), pol	GAGGAGAGCTTCAGGTTTGGG	4.2	Lin (2004)
2315	pol (prot)	AGCTCTATTAGATACAGGA	44.5	Paul (2003)
2316	pol (prot)	GCTCTATTAGATACAGGAGCA	47.1	Scherer (2004)
2328	pol (prot)	ACAGGAGCAGATGATACAGT	59.5	ter Brake (2006)
2375	pol (prot)	ATGGAAACCAAAAATGATAGG	81.9	ter Brake (2006)
2559	pol (RT)	CCTATTGAGACTGTACCAG	8.0	Paul (2003)
2586	pol (RT)	AAGCCAGGAATGGATGGCCCA	63.5	Chang (2005)
2748	pol (RT)	GACAGTACTAAATGGAGAATT	0.0	Sabariegos (2006)
2960	pol (RT)	TGAGACACCAGGGATTAGA	22.8	Surabhi (2002)
2961	pol (RT)	GAGACACCAGGGATTAGAT	22.8	Surabhi (2002)
3092	pol (RT)	TCAATACATGGATGATTTGTA	26.5	Huelsmann (2006)
4168	pol (p15)	AAGGAATTGGAGGAAATGAAC	47.2	Chang (2005)

Table S3

HXB2 coordinate	Target gene	Target site	Conservation (%)	Reference
4173	pol (p15)	ATTGGAGGAAATGAACAAGT	47.6	ter Brake (2006)
4285	pol (int)	GGAGAGCAATGGCTAGTGA	58.7	Nishitsuji (2006)
4537	pol (int)	AATTAGCAGGAAGATGGCCAG	67.0	Chang (2005)
4539	pol (int)	TTAGCAGGAAGATGGCCAGT	67.4	ter Brake (2006)
4650	pol (int)	ATTCCTTACAATCCCCAAG	50.7	ter Brake (2006)
4776	pol (int), cPPT	CACAATTTTAAAGAAAAGGGGGATTGG GGGG	90.7	ter Brake (2006)
4809	pol (int)	TACAGTGCAGGGGAAAGAATA	84.4	ter Brake (2006)
4811	pol (int)	CAGTGCAGGGGAAAGAATA	84.4	Westerhout (2006a)
4884	pol (int), CTS	AAAATTCAAAATTTTCGGGT	77.3	ter Brake (2006)
4888	pol (int), CTS	TTCAAAAATTTTCGGGTTTATT	80.6	ter Brake (2006)
4893	pol (int), CTS	AATTTTCGGGTTTATTACAG	78.5	ter Brake (2006)
4953	pol (int)	CTCTGAAAGGTGAAGGGGCAGTAGTAAT	84.0	ter Brake (2006)
4960	pol (int)	AAGGTGAAGGGGCAGTAGTAA	69.8	Chang (2005)
4966	pol (int)	AAGGGGCAGTAGTAATAGAAG	0.0	Hu (2002)
5040	pol (int), vif	TATGGAAAACAGATGGCAGGTG	90.7	ter Brake (2006)
5138	vif	GGAAAGCTAAGGACTGGTT	0.2	Jacque (2002)
5138	vif	GGAAAGCTAAGGACTGGTT	0.2	Joshi (2005)
5195	vif	G TTCAGAAGTACACATCCC	60.8	Lee (2005)
5323	vif	AGCACACAAGTAGACCCTGA	24.4	Jacque (2002)
5323	vif	AGCACACAAGTAGACCCTG	24.4	Jacque (2002)
5323	vif	AGCACACAAGTAGACCCTG	24.4	Nishitsuji (2004)
5481	vif	CTTGGCACTAGCAGCATT	5.6	Jacque (2002)
5852	tat	CTAGAGCCCTGGAAGCATC	7.8	Surabhi (2002)
5864	tat	AAGCATCCAGGAAGTCAGCCT	10.3	Dave (2004)
5887	tat	AACTGCTTGTACCAATTGCTA	1.8	Coburn (2002)
5889	tat	CTGCTTGTACCAATTGCTATT	5.0	Boden (2003a,b)
5889	tat	CTGCTTGTACCAATTGCTA	3.2	Boden (2004a)
5889	tat	CTGCTTGTACCAATTGCTATT	5.0	Huelsmann (2006)
5915	tat	AAGTGTGCTTTTCATTGCCAAG	5.4	Lee (2003)
5954	tat, rev	GCCTTAGGCATCTCCTATG	1.2	Das (2004)
5966	tat, rev	TCCTATGGCAGGAAGAAGCGGAG	81.5	ter Brake (2006)
5967	tat, rev	CCTATGGCAGGAAGAAGCG	56.8	Das (2004)
5969	tat, rev	TATGGCAGGAAGAAGCGGA	79.9	Surabhi (2002)
5980	tat, rev	GAAGCGGAGACAGCGACGA	45.5	Surabhi (2002)
5983	tat, rev	GCGGAGACAGCGACGAAGAGC	7.4	Lee (2002)
5983	tat, rev	GCGGAGACAGCGACGAAGAGC	7.4	Li (2003)
5983	tat, rev	GCGGAGACAGCGACGAAGAGC	7.4	Scherer (2004)
5983	tat, rev	GCGGAGACAGCGACGAAGAGC	7.4	Joshi (2005)
5983	tat, rev	GCGGAGACAGCGACGAAGAGC	7.4	Li (2005)
5994	tat, rev	GACGAAGAGCTCATCAGAACA	0.2	Boden (2004b)
6205	vpu	AAGAGCAGAAGACAGTGGCAA	53.6	Chang (2005)
6213	vpu, env	AAGACAGTGGCAATGAGAGTG	52.2	Chang (2005)
7152	env	AAGAGGGGACCAGGGAGAGCAC	0.0	Park (2003)
7368	env	AAGTTCTACTGTAATTCAACAC	0.0	Park (2003)
7384	env	CAACACAACCTGTTTAATAGTA	4.0	Scherer (2004)
7460	env	CACAATCACACTCCCATGCAG	1.0	Scherer (2004)
7464	env	AAGACACTCCCATGCAGAATAC	0.0	Park (2003)
7497	env	AAGATGTGGCAGGAAGTAGGAC	0.2	Park (2003)
7635	env	GGAGGAGGCGATATGAGGGAC	0.4	Scherer (2004)
7798	env	CAGCAGGAAGCACTATGGGCGC	75.8	ter Brake (2006)

Table S3

3/3

HXB2 coordinate	Target gene	Target site	Conservation (%)	Reference
7858	env	TGTCTGATATAGTGCAGCAGC	0.0	Scherer (2004)
7916	env	TCTGTTGCAACTCACAGTCTG	24.1	Scherer (2004)
7924	env	AACTCACAGTCTGGGGCATCA	11.9	Dave (2004)
8425	tat, rev, env	AAGAAGAAGGTGGAGAGAGAG	7.1	Dave (2004)
8472	rev, env	AACGGATCCTTGGCACTTATC	0.4	Coburn (2002)
8483	rev, env	GGCACTTATCTGGGACGAT	0.4	Lee (2005)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Lee (2002)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Banerjea (2003)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Li (2003)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Scherer (2004)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Unwalla (2004)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Joshi (2005)
8509	rev, env	GCCTGTGCCTCTTCAGCTACC	40.6	Unwalla (2006)
8558	rev, env	AACGAGGATTGTGGAACCTTCT	8.7	Dave (2004)
8759	env	AAGAATAAGACAGGGCTTGGA	9.5	Dave (2004)
8762	env	AATAAGACAGGGCTTGGAAG	10.9	Dave (2004)
8817	nef	AAGTAGTGTGATTGGATGGCC	0.4	Chang (2005)
8844	nef	AAGGGAAAGAAATGAGACGAGC	6.3	Dave (2004)
8914	nef	CAUGGAGCAAUCACAAGUATT	0.0	Sabariegos (2006)
8960	nef	GTGCCTGGCTAGAAGCACA	11.6	Jacque (2002)
8960	nef	GTGCCTGGCTAGAAGCACA	11.6	Das (2004)
8960	nef	GTGCCTGGCTAGAAGCACA	11.6	Westerhout (2005)
8960	nef	GTGCCTGGCTAGAAGCACA	11.6	Westerhout (2006a,b)
8999	nef	TTCCAGTCACACCTCAGGTAC	1.6	Scherer (2004)
9063	nef, U3	CTTTTTTAAAAGAAAAGGGGGGACTGGA	90.7	ter Brake (2006)
9069	nef, U3	AAAAGAAAAGGGGGGACTGGA	90.7	Dave (2004)
9078	nef, U3	GGGGGGACTGGAAGGGCTA	14.1	Das (2004)
9081	nef, U3	GGGACTGGAAGGGCTAATT	10.6	Westerhout (2006a)
9083	nef, U3	GACTGGAAGGGCTAATTCA	2.6	Nishitsuji (2006)

\* For the target sequence longer than 21 nt, conservation was calculated for each possible 21 nt subsequence, and the highest value is shown. Target sequence shorter than 21 nt was extended to 21 nt for calculating their conservations.

Table S4

1/5

HXB2 coordinate	Target gene	Target site (21 bp)	Conservation (%)	siRNA efficacy prediction Ui-Tei, Reynolds, Amarzguioui		
412	U3, TATA	CCCTCAGATGCTGCATATAAG	74.6			
413	U3, TATA	CCTCAGATGCTGCATATAAGC	75.1			
414	U3, TATA	CTCAGATGCTGCATATAAGCA	74.6	yes		
415	U3, TATA	TCAGATGCTGCATATAAGCAG	74.6			
416	U3, TATA	CAGATGCTGCATATAAGCAGC	75.8			
505	R, TAR	GGGAACCCACTGCTTAAGCCT	77.8	yes		
506	R, TAR	GGAACCCACTGCTTAAGCCTC	81.3			
507	R, TAR, poly A	GAACCCACTGCTTAAGCCTCA	84.9			
508	R, TAR, poly A	AACCCACTGCTTAAGCCTCAA	85.3			yes
509	R, TAR, poly A	ACCCACTGCTTAAGCCTCAAT	84.9	yes	yes	yes
510	R, TAR, poly A	CCCACTGCTTAAGCCTCAATA	84.9	yes	yes	yes
511	R, TAR, poly A	CCACTGCTTAAGCCTCAATAA	85.2		yes	
512	R, poly A	CACTGCTTAAGCCTCAATAAA	85.4	yes	yes	yes
513	R, poly A	ACTGCTTAAGCCTCAATAAAG	84.7			
514	R, poly A	CTGCTTAAGCCTCAATAAAGC	85.3			
515	R, poly A	TGCTTAAGCCTCAATAAAGCT	85.3	yes		
516	R, poly A	GCTTAAGCCTCAATAAAGCTT	85.7			
517	R, poly A	CTTAAGCCTCAATAAAGCTTG	85.7			
518	R, poly A	TTAAGCCTCAATAAAGCTTGC	86.7			
519	R, poly A	TAAGCCTCAATAAAGCTTGCC	86.6			
520	R, poly A	AAGCCTCAATAAAGCTTGCCCT	93.9			yes
521	R, poly A	AGCCTCAATAAAGCTTGCCCTT	94.2	yes		yes
522	R, poly A	GCCTCAATAAAGCTTGCCCTTG	95.0			
523	R, poly A	CCTCAATAAAGCTTGCCCTTGA	94.2			
524	R, poly A	CTCAATAAAGCTTGCCCTTGAG	94.6			
525	R, poly A	TCAATAAAGCTTGCCCTTGAGT	87.5			
526	R, poly A	CAATAAAGCTTGCCCTTGAGTG	87.5			
527	R, poly A	AATAAAGCTTGCCCTTGAGTGC	85.5			
528	R, poly A	ATAAAGCTTGCCCTTGAGTGCT	79.5			
552	U5	AGTAGTGTGTGCCCGTCTGTT	70.3			
576	U5, PAS	TGACTCTGGTAACTAGAGATC	71.5			
577	U5, PAS	GACTCTGGTAACTAGAGATCC	81.3			
578	U5, PAS	ACTCTGGTAACTAGAGATCCC	81.9			
579	U5, PAS	CTCTGGTAACTAGAGATCCCT	87.1			yes
580	U5, PAS	TCTGGTAACTAGAGATCCCTC	87.1			
581	U5, PAS	CTGGTAACTAGAGATCCCTCA	87.7			yes
582	U5, PAS	TGGTAACTAGAGATCCCTCAG	88.2			
583	U5, PAS	GGTAACTAGAGATCCCTCAGA	88.2		yes	
584	U5, PAS	GTAACCTAGAGATCCCTCAGAC	87.7			
585	U5	TAACTAGAGATCCCTCAGACC	83.7			
622	U5, PBS	AAAATCTCTAGCAGTGGCGCC	90.2			
623	U5, PBS	AAATCTCTAGCAGTGGCGCCC	93.1			
624	U5, PBS	AATCTCTAGCAGTGGCGCCCG	92.0			
625	U5, PBS	ATCTCTAGCAGTGGCGCCCGA	92.1			
626	U5, PBS	TCTCTAGCAGTGGCGCCCGAA	93.3			
627	U5, PBS	CTCTAGCAGTGGCGCCCGAAC	92.8			
628	U5, PBS	TCTAGCAGTGGCGCCCGAACA	92.4			
629	U5, PBS	CTAGCAGTGGCGCCCGAACAG	92.4			
630	U5, PBS	TAGCAGTGGCGCCCGAACAGG	93.8			
631	U5, PBS	AGCAGTGGCGCCCGAACAGGG	93.3			



Table S4

2/5

HXB2 coordinate	Target gene	Target site (21 bp)	Conservation (%)	siRNA efficacy prediction Ui-Tei, Reynolds, Amarzguioui	
632	U5, PBS	GCAGTGGCGCCCGAACAGGGGA	90.0		
633	U5, PBS	CAGTGGCGCCCGAACAGGGAC	90.1		
682	U5	TCTCTCGACGCAGGACTCGGC	85.2		
683	U5	CTCTCGACGCAGGACTCGGCT	88.1		
684	U5	TCTCGACGCAGGACTCGGCTT	88.5		
685	U5	CTCGACGCAGGACTCGGCTTG	88.5		
686	U5	TCGACGCAGGACTCGGCTTGC	88.2		
687	U5	CGACGCAGGACTCGGCTTGCT	88.5		
688	U5	GACGCAGGACTCGGCTTGCTG	90.0		
689	U5	ACGCAGGACTCGGCTTGCTGA	88.2	yes	yes
761	U5, Ψ	TTTACTAGCGGAGGCTAGAA	85.2		yes
762	U5, Ψ	TTGACTAGCGGAGGCTAGAAG	86.9		
763	U5, Ψ	TGACTAGCGGAGGCTAGAAGG	87.6		
764	U5, Ψ	GACTAGCGGAGGCTAGAAGGA	88.0	yes	
765	U5, Ψ	ACTAGCGGAGGCTAGAAGGAG	88.0		
766	U5, Ψ	CTAGCGGAGGCTAGAAGGAGA	88.4		yes
767	U5, Ψ	TAGCGGAGGCTAGAAGGAGAG	88.1		
768	U5, Ψ	AGCGGAGGCTAGAAGGAGAGA	87.8		yes
769	U5, Ψ	GCGGAGGCTAGAAGGAGAGAG	88.3		
770	U5, Ψ, gag (AUG)	CGGAGGCTAGAAGGAGAGAGA	88.6	yes	yes
771	U5, Ψ, gag (AUG)	GGAGGCTAGAAGGAGAGAGAT	88.3		yes
772	U5, Ψ, gag (AUG)	GAGGCTAGAAGGAGAGAGATG	91.0		
773	U5, Ψ, gag (AUG)	AGGCTAGAAGGAGAGAGATGG	91.4		
774	U5, Ψ, gag (AUG)	GGCTAGAAGGAGAGAGATGGG	91.4		
775	U5, Ψ, gag (AUG)	GCTAGAAGGAGAGAGATGGGT	91.1		
776	U5, Ψ, gag (AUG)	CTAGAAGGAGAGAGATGGGTG	90.4		
777	U5, Ψ, gag (AUG)	TAGAAGGAGAGAGATGGGTGC	90.4		
778	U5, Ψ, gag (AUG)	AGAAGGAGAGAGATGGGTGCG	90.1		
779	U5, Ψ, gag (AUG)	GAAGGAGAGAGATGGGTGCGA	90.4		
780	U5, gag (AUG)	AAGGAGAGAGATGGGTGCGAG	92.1		
781	U5, gag (AUG)	AGGAGAGAGATGGGTGCGAGA	92.1		yes
782	U5, gag (AUG)	GGAGAGAGATGGGTGCGAGAG	92.1		
783	U5, gag (AUG)	GAGAGAGATGGGTGCGAGAGC	93.2		
784	U5, gag (AUG)	AGAGAGATGGGTGCGAGAGCG	92.5		
785	U5, gag (AUG)	GAGAGATGGGTGCGAGAGCGT	93.2		
786	U5, gag (AUG)	AGAGATGGGTGCGAGAGCGTC	92.9		
787	U5, gag (AUG)	GAGATGGGTGCGAGAGCGTCA	89.8		
1597	gag (p24)	TTAAATAAAATAGTAAGAATG	72.4		
1598	gag (p24)	TAAATAAAATAGTAAGAATGT	77.3		
1599	gag (p24)	AAATAAAATAGTAAGAATGTA	76.9		yes
1600	gag (p24)	AATAAAATAGTAAGAATGTAT	77.7		
1601	gag (p24)	ATAAAATAGTAAGAATGTATA	77.7		yes
1602	gag (p24)	TAAAATAGTAAGAATGTATAG	77.5		
1603	gag (p24)	AAAATAGTAAGAATGTATAGC	75.9		
1604	gag (p24)	AAATAGTAAGAATGTATAGCC	75.9		
1605	gag (p24)	AATAGTAAGAATGTATAGCCC	75.9		
1606	gag (p24)	ATAGTAAGAATGTATAGCCCT	75.9		
1817	gag (p24)	TAGAAGAAATGATGACAGCAT	82.3	yes	yes
1818	gag (p24)	AGAAGAAATGATGACAGCATG	82.3		
2075	gag (p7, p1), pol, ribosomal slip site	GACAGGCTAATTTTTTAGGGA	70.6	yes	

Table S4

3/5

HXB2 coordinate	Target gene	Target site (21 bp)	Conservation (%)	siRNA efficacy prediction		
				Ui-Tei,	Reynolds,	Amarzguioui
2077	gag (p7, p1), pol, ribosomal slip site	CAGGCTAATTTTTTAGGGAAA	70.0	yes		yes
2251	gag (p6), pol (prot)	TCCCTCAAATCACTCTTTGGC	79.9			
2252	gag (p6), pol (prot)	CCCTCAAATCACTCTTTGGCA	81.9	yes		yes
2328	pol (prot)	ACAGGAGCAGATGATACAGTA	81.3		yes	
2329	pol (prot)	CAGGAGCAGATGATACAGTAT	77.7	yes	yes	yes
2330	pol (prot)	AGGAGCAGATGATACAGTATT	77.9	yes		yes
2331	pol (prot)	GGAGCAGATGATACAGTATTA	79.1		yes	yes
2332	pol (prot)	GAGCAGATGATACAGTATTAG	78.5			
2333	pol (prot)	AGCAGATGATACAGTATTAGA	77.3	yes		yes
2334	pol (prot)	GCAGATGATACAGTATTAGAA	79.7		yes	
2335	pol (prot)	CAGATGATACAGTATTAGAAG	78.3			
2336	pol (prot)	AGATGATACAGTATTAGAAGA	77.7			
2375	pol (prot)	ATGGAAACCAAAAATGATAGG	81.9			
2487	pol (prot)	CCTACACCTGTCAACATAATT	72.0			
2488	pol (prot)	CTACACCTGTCAACATAATTG	72.2			
2489	pol (prot)	TACACCTGTCAACATAAATTGG	72.2			
2490	pol (prot)	ACACCTGTCAACATAAATTGGA	73.0			
2611	pol (RT)	TTAAACAATGGCCATTGACAG	70.0			
2613	pol (RT)	AAACAATGGCCATTGACAGAA	73.0			
2614	pol (RT)	AACAATGGCCATTGACAGAAG	73.0			
2615	pol (RT)	ACAATGGCCATTGACAGAAGA	73.6		yes	
2696	pol (RT)	AATTGGGCTGAAAATCCATA	74.8			
3302	pol (RT)	CTGGACTGTCAATGATATACA	70.2	yes	yes	yes
4173	pol (p15)	ATTGGAGGAAATGAACAAGTA	76.7		yes	
4174	pol (p15)	TTGGAGGAAATGAACAAGTAG	76.5			
4175	pol (p15)	TGGAGGAAATGAACAAGTAGA	76.5	yes		yes
4176	pol (p15)	GGAGGAAATGAACAAGTAGAT	71.1			yes
4177	pol (p15)	GAGGAAATGAACAAGTAGATA	71.1	yes		yes
4178	pol (p15)	AGGAAATGAACAAGTAGATAA	70.5	yes	yes	yes
4179	pol (p15)	GGAAATGAACAAGTAGATAAA	70.7		yes	
4182	pol (p15)	AATGAACAAGTAGATAAATTA	71.7		yes	
4183	pol (p15)	ATGAACAAGTAGATAAATTAG	71.9			
4184	pol (p15)	TGAACAAGTAGATAAATTAGT	71.9			
4749	pol (int)	ACAGCAGTACAAATGGCAGTA	74.9			
4750	pol (int)	CAGCAGTACAAATGGCAGTAT	74.1	yes		yes
4751	pol (int)	AGCAGTACAAATGGCAGTATT	74.1	yes	yes	yes
4752	pol (int)	GCAAGTACAAATGGCAGTATTC	71.5			
4753	pol (int)	CAGTACAAATGGCAGTATTCA	71.5	yes	yes	yes
4754	pol (int)	AGTACAAATGGCAGTATTCAT	71.5			
4776	pol (int), cPPT	CACAATTTTAAAAGAAAAGGG	89.8			
4777	pol (int), cPPT	ACAATTTTAAAAGAAAAGGGG	89.8			
4778	pol (int), cPPT	CAATTTTAAAAGAAAAGGGGG	89.8			
4779	pol (int), cPPT	AATTTTAAAAGAAAAGGGGGG	90.5			
4780	pol (int), cPPT	ATTTTAAAAGAAAAGGGGGGA	90.5			
4781	pol (int), cPPT	TTTTAAAAGAAAAGGGGGGAT	90.5			
4782	pol (int), cPPT	TTTAAAAGAAAAGGGGGGATT	90.7			
4783	pol (int), cPPT	TTAAAAGAAAAGGGGGGATTG	90.7			
4784	pol (int), cPPT	TAAAAGAAAAGGGGGGATTGG	90.7			
4785	pol (int), cPPT	AAAAGAAAAGGGGGGATTGGG	90.3			
4786	pol (int), cPPT	AAAGAAAAGGGGGGATTGGGG	89.4			

Table S4

4/5

HXB2 coordinate	Target gene	Target site (21 bp)	Conservation (%)	siRNA efficacy prediction Ui-Tei, Reynolds, Amarzguioui		
4787	pol (int), cPPT	AAGAAAAGGGGGATTGGGGG	86.6			
4788	pol (int), cPPT	AGAAAAGGGGGATTGGGGG	78.1			
4789	pol (int), cPPT	GAAAAGGGGGATTGGGGGGT	78.1			
4790	pol (int), cPPT	AAAAGGGGGATTGGGGGGTA	78.1			
4791	pol (int), cPPT	AAAGGGGGATTGGGGGGTAC	76.9			
4792	pol (int), cPPT	AAGGGGGATTGGGGGGTACA	76.7			yes
4793	pol (int), cPPT	AGGGGGATTGGGGGGTACAG	80.2			
4794	pol (int), cPPT	GGGGGGATTGGGGGGTACAGT	80.4	yes		yes
4795	pol (int), cPPT	GGGGGATTGGGGGGTACAGTG	80.4			
4796	pol (int), cPPT	GGGGATTGGGGGGTACAGTGC	80.0			
4797	pol (int), cPPT	GGGATTGGGGGGTACAGTGCA	79.3			
4798	pol (int), cPPT	GGATTGGGGGGTACAGTGCAG	78.9			
4799	pol (int), cPPT	GATTGGGGGGTACAGTGCAGG	79.1			
4800	pol (int), cPPT	ATTGGGGGGTACAGTGCAGGG	76.1			
4801	pol (int)	TTGGGGGGTACAGTGCAGGGG	75.9			
4802	pol (int)	TGGGGGGTACAGTGCAGGGGA	75.5			yes
4803	pol (int)	GGGGGGTACAGTGCAGGGGAA	73.5			yes
4804	pol (int)	GGGGGTACAGTGCAGGGGAAA	72.9			yes
4805	pol (int)	GGGGTACAGTGCAGGGGAAAG	73.1			
4806	pol (int)	GGGTACAGTGCAGGGGAAAGA	72.3	yes	yes	
4807	pol (int)	GGTACAGTGCAGGGGAAAGAA	72.9			
4808	pol (int)	GTACAGTGCAGGGGAAAGAAT	74.5		yes	
4809	pol (int)	TACAGTGCAGGGGAAAGAATA	84.4	yes	yes	yes
4810	pol (int)	ACAGTGCAGGGGAAAGAATAA	71.5		yes	yes
4811	pol (int)	CAGTGCAGGGGAAAGAATAAT	71.5	yes		yes
4812	pol (int)	AGTGCAGGGGAAAGAATAATA	73.7		yes	
4813	pol (int)	GTGCAGGGGAAAGAATAATAG	73.9			
4814	pol (int)	TGCAGGGGAAAGAATAATAGA	74.1	yes	yes	yes
4883	pol (int), CTS	AAAAATTCAAATTTTCGGGT	77.3			
4884	pol (int), CTS	AAAATTCAAATTTTCGGGTT	76.7			
4885	pol (int), CTS	AAATTCAAATTTTCGGGTTT	78.5			
4886	pol (int), CTS	AATTCAAATTTTCGGGTTTA	78.7			
4887	pol (int), CTS	ATTCAAATTTTCGGGTTTAT	79.3		yes	
4888	pol (int), CTS	TTCAAATTTTCGGGTTTATT	80.6	yes		yes
4889	pol (int), CTS	TCAAATTTTCGGGTTTATTA	76.9		yes	
4890	pol (int), CTS	CAAATTTTCGGGTTTATTAC	77.5			
4891	pol (int), CTS	AAAATTTTCGGGTTTATTACA	77.7			
4892	pol (int), CTS	AAATTTTCGGGTTTATTACAG	78.5			
4953	pol (int)	CTCTGGAAAGGTGAAGGGGCA	82.2			yes
4954	pol (int)	TCTGGAAAGGTGAAGGGGCAG	82.0			
4955	pol (int)	CTGGAAAGGTGAAGGGGCAGT	82.2			yes
4956	pol (int)	TGGAAAGGTGAAGGGGCAGTA	84.0			
4957	pol (int)	GGAAAGGTGAAGGGGCAGTAG	83.6			
4958	pol (int)	GAAAGGTGAAGGGGCAGTAGT	83.4			
4959	pol (int)	AAAGGTGAAGGGGCAGTAGTA	72.1		yes	yes
4961	pol (int)	AGGTGAAGGGGCAGTAGTAAT	70.1	yes		yes
5040	pol (int), vif	TATGGAAAACAGATGGCAGGT	90.7			
5041	pol (int), vif	ATGGAAAACAGATGGCAGGTG	90.1			
5966	pol (int), vif	TCCTATGGCAGGAAGAAGCGG	81.5			
5967	pol (int), vif	CCTATGGCAGGAAGAAGCGGA	79.9			

Table S4

5/5

<b>HXB2 coordinate</b>	<b>Target gene</b>	<b>Target site (21 bp)</b>	<b>Conservation (%)</b>	<b>siRNA efficacy prediction Ui-Tei, Reynolds, Amarzguioui</b>
5970	pol (int), vif	ATGGCAGGAAGAAGCGGAGAC	70.2	
7654	env (gp120)	ACAATTGGAGAAGTGAATTAT	70.7	
7655	env (gp120)	CAATTGGAGAAGTGAATTATA	70.5	yes
7796	env (gp41)	AGCAGCAGGAAGCACTATGGG	71.1	
7797	env (gp41)	GCAGCAGGAAGCACTATGGGC	71.7	
7798	env (gp41)	CAGCAGGAAGCACTATGGGCG	75.4	
7799	env (gp41)	AGCAGGAAGCACTATGGGCGC	75.8	
7830	env (gp41)	ACGCTGACGGTACAGGCCAGA	72.1	yes
7831	env (gp41)	CGCTGACGGTACAGGCCAGAC	72.9	
9063	nef, 3'PPT	CTTTTAAAAAGAAAAGGGGGG	86.1	
9064	nef, 3'PPT	TTTTTAAAAAGAAAAGGGGGGA	89.5	
9065	nef, 3'PPT	TTTTTAAAAAGAAAAGGGGGGAC	90.2	
9066	nef, 3'PPT, U3	TTTAAAAAGAAAAGGGGGGACT	90.1	
9067	nef, 3'PPT, U3	TTAAAAAGAAAAGGGGGGACTG	90.3	
9068	nef, 3'PPT, U3	TAAAAAGAAAAGGGGGGACTGG	90.5	
9069	nef, 3'PPT, U3	AAAAGAAAAGGGGGGACTGGA	90.7	