

HIV-1 sequences in the epidemic suggest an alternative pathway for the generation of the Long Terminal Repeats

Pierre Cappy^{1,2,3*}, Alice Moisan^{2,3}, Fabienne De Oliveira², Jean-Christophe Plantier^{2,3} and Matteo Negroni¹

¹Université de Strasbourg, CNRS, Architecture et Réactivité de l'ARN, UPR 9002, F-67000 Strasbourg, France; ²Laboratoire de Virologie, Associé au Centre National de Référence du VIH, Hôpital Charles Nicolle, CHU de Rouen, Rouen, France; ³GRAM EA2656, Faculté de Médecine-Pharmacie, Normandie Université, Rouen, France

Supplementary information

Table S1. Recombinant strains retrieved and used in the study (HIV-1/MO, HIV-1/M CRFS and URFs).

Identification	Type of recombinant	Reference	GenBank accession number
97CA-MP645	HIV-1/MO	Peeter et al., 1999, J virol	AJ239083
REC024	HIV-1/MO	Ngoupo et al., 2016, AIDS	KM438032
RBF208	HIV-1/MO	Vessièrè et al., 2010, AIDS	GQ351296
YBF274	HIV-1/MO	De Oliveira et al., 2017, Retrovirology	KX579838
REC107	HIV-1/MO	Ngoupo et al., 2016, 21th Int. AIDS Conf., 2016, Durban	N/A
BCF212	HIV-1/MO	N/A	KY359380
RBF222	HIV-1/MO	N/A	KY359382
RBF237	HIV-1/MO	N/A	KY359381
RBF240	HIV-1/MO	N/A	KY359384
subtype G	CRF	Abecasis et al., 2008, J virol	N/A
CRF03_AB	CRF	Liitsola et al.. 1998, AIDS	N/A
CRF08_BC	CRF	Huang et al., 2013, J med primatol	KC914396
CRF12_BF	CRF	Carr et al.. 2001,AIDS	AF385934-6
CRF26_AU	CRF	Vidal et al.. 2009, AIDS Res Hum Retrov.	N/A
CRF31_BC	CRF	Santos et al.. 2006, AIDS	N/A
CRF32_06/A1	CRF	Adojaan et al.. 2005, J AIDS	N/A
CRF42_BF	CRF	Struck et al., 2015, AIDS Res Hum Retrov.	EU170139
CRF45_cpx	CRF	Niama et al.. 2009, AIDS Res Hum Retrov.	N/A
CRF60_BC	CRF	Simonetti et al.. 2013, Infect Genet Evol	N/A
CRF71_BF	CRF	Pessoa et al.. 2014, PLoS One	N/A
CRF87_cpx	CRF	Hu et al.. 2016, AIDS Res Hum Retrov.	N/A
CRF88_BC	CRF	Hu et al.. 2016, AIDS Res Hum Retrov.	N/A
URF_SHI	URF	Shi et al., 2004, AIDS Res Hum Retrov.	AY445524
URF_CRN	URF	Frangè et al., 2008, Retrovirology	EU448296
URF_AUK	URF	Frangè et al., 2008, Retrovirology	EU448295
URF_WEI	URF	Wei et al., 2013, AIDS Res Hum Retrov.	KC833436
URF_GUI	URF	Gui et al., 2013, AIDS Res Hum Retrov.	KP170487

N/A: not applicable

Table S2. Concordance of the Tat/TAR couple according to the concordance of *gag* and *nef*.

	Tat=TAR	Tat≠TAR*	Total
discordant gag/nef	4	4	8
matching gag/nef	12	7	19
Total	16	11	27

*or presence of a recombination breakpoint in tat or R

H0: the two variables (Tat/TAR state and gag/nef state) are independent
p-value = 0.6753 (Fisher's Exact Test for Count Data)

Table S3. Panel of reference sequences used for the recombination analyses.

HIV-1/M subtype	Genebank accession number	HIV-1/M subtype	Genebank accession number
A1	AF004885	D	K03454
A1	AM000053	D	M22639
A1	KT152840	D	AB485648
A1	EU861977	D	JX236668
A1	JX236677	D	AB485650
A1	AF413987	F1	FJ771008
A1	JX236669	F1	AB485658
B	AF042100	G	AF084936
B	DQ007903	G	AB287004
B	K03455	CRF01	U54771
B	AB731663	CRF01	DQ859178
B	AY423381	CRF01	JX112869
B	AF086817	CRF01	AB253423
B	EF363123	CRF01	AB565504
C	AF290028	CRF02	L39106
C	AB485643	CRF02	AB231896
C	EF469243	CRF02	AB049811
C	KC156210	CRF02	AB485633
C	KC156212	CRF02	AB485636
C	KC156212	HIV-1/O	Genebank accession number
C	KC156220	O	L20587
C	AY228557	O	L20571
C	AB254153	O	AJ302647
		P	GU111555

Supplementary Figure S1

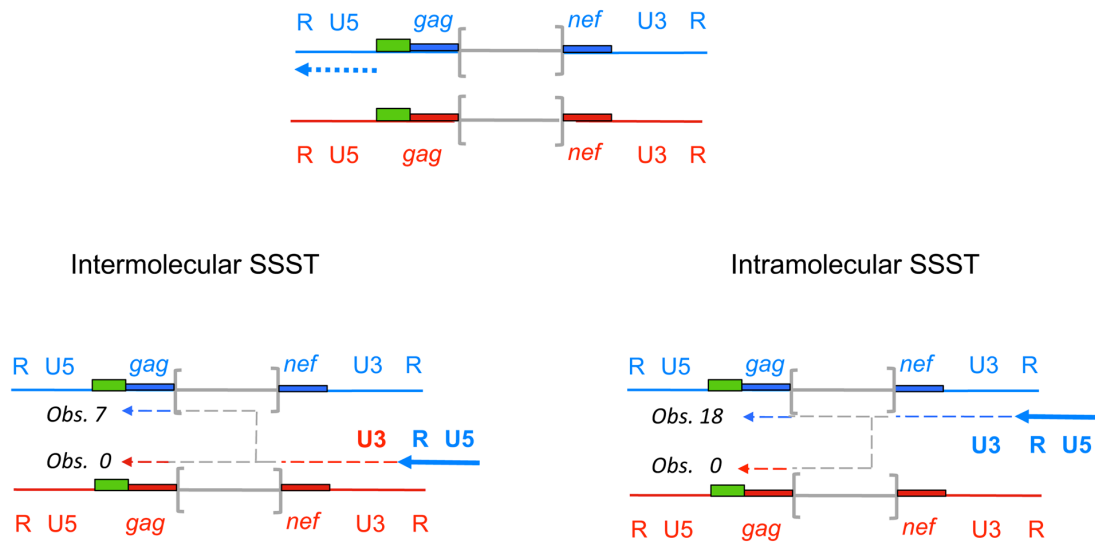


Figure S1. Pathways through which (-) DNA SSST should generate LTR with a U3/R breakpoint coupled to *discordant gag-nef* and non-recombinant LTRs coupled to *matching gag-nef*. The green rectangle represents the PBS, the boxes indicate the portions referred in the main text as *gag* and *nef* sequences for classifying proviruses into *matching gag-nef* or *discordant gag-nef*. To indicate whether they are phylogenetically matching or discordant, they are coloured in blue for the RNAs where reverse transcription is primed and in red for the other. The internal portion of the genome is drawn in grey, included in brackets. Upper drawing: synthesis of (-) strong stop DNA is indicated by a thick blue dotted arrow. Transfer of (-) strong stop DNA can then be either intermolecular or intramolecular⁵⁻⁶. In the case of intermolecular SSST, a breakpoint is generated in all cases at the U3/R junction. Depending on the number of template switching occurring during reverse transcription of internal regions, this can generate either *matching* or *discordant gag/nef* pairs. The number of cases observed in our analyses is indicated (Obs.).

Supplementary Figure S2

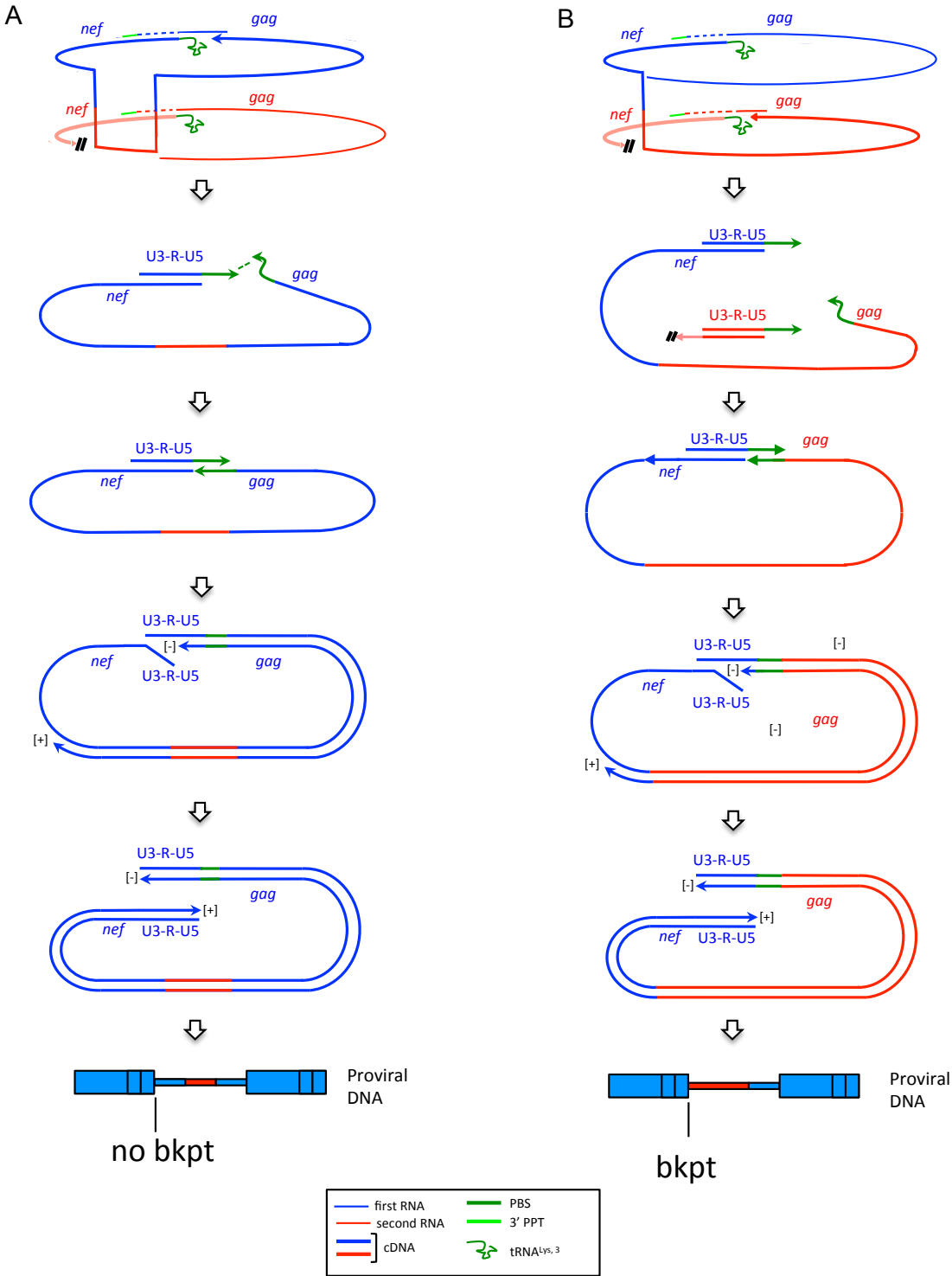


Figure S2. The description of the figure is the same as for Figure 4 of the main text. It shows that depending on the pattern of internal template switching, intramolecular (+) DNA SSST will not generate (column A) or it will generate (column B) a breakpoint at the U5/gag junction.

Supplementary sequences

Sequences of the HIV-1/MO *nef*-LTR-MA concatenated sequences.

RBF208

ATGGGAAACGCATTGAGGAGAAGCCCATTTCCAGGATGGGCAGCAATAAGAGARAGAATAAGAA
GARCTCCCCYCCCTGAGCCTGAACCATGTGCACCTGGAGTAGGACCAGTCTCCAGGGAATTAGC
AGCTAGAGGAGGGATACCCAGTTCCTTTACTCCTCAAAAYAATGCAGCCCTTGCATTCTAGAAA
GTCACCAAGAAGAAGAAGTAGGTTTCCCAGTAAGACCTCAAGTGCCTCTAAGRCCAATGACATAT
AAACTAGCATTGACCTCAGTTTTTAAAGAAAAGGGGGGACTGGATGGGCTAATTTACTCCCCT
GAKAAAGCAGAGATCCTAGATCTTTGGATACATCACACTCAGGGAATCTTCCCTGATTGGCAGTG
TTACACACCAGGACCAGGAAGTACACTAACATTTGGGTGGTTATTTAAGCTAGTACCAG
TGTCAAAAGAGGAGGCAGAAGCACTAGGAAACACAAGTGAAGGGGCTAACCTACTGCATCCAGC
TTGCAGCCATGGCTTTGGAGACGCACACGGAGAGATACTGAAATGGCAGTTTGATAGATCATTAG
CCAACACCCATGTTGCTATGATAACCAACCCTGAGCTCTTCAACAAGGACTAAAAGTCTGACAT
GAAGATAGCTGACGCTGCGGGACTTTCCAACAGAGACTGCTGACACTGCGGGGACTTTCCAGTG
TGGGAGGGACAGGGGCGGTTCCGGGAGTGGCTAACCCCTCAGATGCTGCATATAAGCAGCCGCT
TTCTGCTTGACCGGGTCTCGGTTAGAGAACCAGGTCTGAGCCCGGGAGCTCCCTGGCCTCTAG
CTGAACCCGCTGCTTAACGCTCAATAAAGCTTGCCTTGAGTGAGAAGCAGTGTGTGCTCATCTGT
TCAACCCTGGTATCTAGAGATCCCTCAGATCACGTAGACTGAGGYAGAAAATCTCTAGCAGTGGC
GCCCCAACAGGGACCGGAAAGAGAAAAGTGAACCAGGGAAGAAAACCTCCGACGCAACGGGCT
CGGCTTAGCGGAGTGCACCCGCTAAGAGGGCGAGAGGACTCACAGAAGGGGTGAGTAAWTTTGC
TGGCGGTGGCCAGACCTAGGGGAAGGGCGAAGTCCCTAGGGGAGGAAGATGGGTGCGAGTGC
GTCTGTGTTGAAAGGGAGCAAATTGGATGCATGGGAACAAATTAGGTTAAAGCCAGGATCTAAAA
AGAAATATCAATTAACATTTAGTATGGGCAAGCAGGGAGCTGGACAGATACGCATGTAATCCT
GAGCTATTAGAACTATAGAGGGTAACGAGAACTATTACAGCAGTTAGAGCCAGCTCTCAAGAC
AGGGTCAGACAGCCTGCAGTCACTCTGGAACACAATAGCAGTGCTCTGGTGTGTTACAACAGA
TATARAGTTGAAGACACACAGCAGGCGATACAAAAGTTAAAGGAAGTAATGGGGAGCAGGAAGT
CTGCAGAGGTCGCTAAGGAAGACACAGGCGAAAGGCAGAGGGGTCAAATTAC

RBF222

ATGGGTRGCAAGTGGTCAAAGGTAGCTTTGTGGGATGGYCTCAGGTTATGGAAAGAATAAGRC
AAACTCCTCCGAGGCAAACCTCCAGCAGCAACAGGAGTGGGAACAGTATCTCAAGATTTAGGT
AGACATGGAGCAATCACAAGCAGTAATACAGCARCTACTAATGCTGCCTGTGCCTGGCTAGAAGC
ACAAGAAGACGAGGAGGTAGGCTTTCCAGTCAGGCCGCGAGGTACCTTTGAGACCAATGACTTTT
AAGGCAGCRTTGTATCTCAGCCACTTTTTAAAGAAAAGGGGGGACTGGATGGGTTAATTTACTC
CAAGAAAAGACAAGAGATCCTTGATCTGTGGGTCTATCACACACAAGGATTCTTCCCAGATTGGC
AGAACTACACACCAGGACCAGGGACTAGATTCCCCTGACCTTTGGGTGGTGTTCAGCTAGAA
CCRATGGATCCAGCAGAGGTAGAGGAAGCCAATAAAGGAGAGAACAATAGTTTATTACACCCCAT
CTGTCAGCATGGAATGGAGGACGAGGACAGAGAAGTGTGCTGGTCTGGAGATTTGACAGCCACCTG
GCATGGAAACACAKCGCTCGAGAGCTGCATCCGGAGTTCTACAAAGACTGCTGACACAGAAGTT
GCTGACAGGGACTTTCCGCTGGGACTTTCCGGGGGAGGTGTGGCTTGGGAGGAGTTGGGGAGT
GGCTAACCCCTCAGAWGCTGCATATAAGCAGCCGCTTCTCGCCTGTACTGGGTCTCTCTTYTAG
ACCAGATTTGAGCCTGGGAGCTCTCTGGCTAGCAGGGAACCCACTGCTTAAGCCTCAATAAATTC
AAGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCTTTCTAGTCT
GTGTAAAAAATCTCTAGCAGTGGCGCCCCGAACAGGGACCTTTAAATWAATAGGGACTCGAAAGC
GAAAGTTCCARAGAARATCTCTCGACGCAGGGACYCGGCTGGCTGAGGTGCACGCAGCAAGAG

GCRAGAGCGGCRACCTGGTGTAGTACGCCAAAAAATTTGACTAGCGGAGGCTARAAGGARARARAT
GGGTGCGAGAGCGTCAGTGTTAAGTGGGGGAAAATTAGATACATGGGAAAGAATTCGCTTAAGG
CCAGGGGGAAAAGAAAAAATATAAACTAAAACATATAGTATGGGCAAGCAGGGAGTTGGAAAGATT
TGCACTTAACCCTGGCCTTTTAGAATCAGCAGAAGGATGTCAACAGTTAATGGAACAGTTACAGT
CAACTCTCAAGACAGGATCAGAAGAACTTAAATCATTATTTAATACCTTAGCAGTCCTTTGGTGCG
TRCATCAAAGGATGGACATAAGAGACACCAAGGAAGCCTTAGATAAACTAGAGGAAGCGCAAAT
AAGAGCAAAGCAAAGGAACAGCAGGCAACAGCTAGCGCAAGCTGCAGCCAAAATTAC

REC107

ATGGGGAATGCATGGAGTAAAAGTAAATTTGCAGGATGGTCAGCAGTAAGAGAGAGAATGAGGA
AAACCCCCACTAATAATCCTGAACCATGTGCACCTGGGGTAGGAGAAATCTCTAGACAGTTAGCA
GCTAGAGGAGAAATACCAGGTTCCCTATACTCTTCATAACAATGCAGCCTTGGCCTTCCTAGAAGA
ACATGAAGAGGGTGAAGAAGTAGGGTCCAGTAACACCTCMAGTGCCTCTTAGGCCAATGACC
TATAAAGGAGCATTGACCTCGGCTTCTTTTTAAAAGAAAAGGGAGGACTGGATGGGTTGATTTA
CTCCATAAAAAGAGCAGAGATCCTAGATCTTTGGGTGTATCRCACTCAGGGCTACTTCCCTGATT
GGCAGAACTACACACCAGGACCAGGAACCAGATACCCATTGACATTTGGATGGCCATTCAAGCTA
GTGCCAGTATCAGCAGAAGAGGCAGAAAGACTGGGCAATATGTGTGAAAGGACTACGCTCCTAC
ATCCAATCAGTAATCATGGTAGTGAAGATGAACATGGGGAAATACTGAAATGGCAGTTTGATAGTT
CAYTGGGCCGCACCCATGTTGCTTTAATAACTCACCTCATCTTCAACAAGGACTAAAATGTT
TATCTGAAGAGTGCTGACACTGCGGGACTTTCCAGCAGAGACTGCTGACGCGGGCGGGGACTTTC
CAGTGTGGGAGGGATAAGGGCGGTTCCGGGAGTGGCTAACCTCAGATGCTGCATATAAGCAG
CTGCTTCTCGCTTGTACTGGGTCTTTTTGCTAGACCAGATTGGAGCCTGGGAGCTCTGGCTA
GTAGAGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCCTGAGTGCTTCAAGTAGTGTGTGCC
ATCTGTTGTGTGACTCTGGTATCTAGAGATCCCTCAGACCACTCGAGACTGTGTAAAAATCTCTAG
CAGTGGCGCCCGAACAGGGACCCTAATAGGGACTCGAAAGCGAAAGTTCCAGAGAAGTTCTCTC
GACGCACGGACTCGGCTTGTGAGGTGCACACAGCAAGAGGCGAGAGCGGCGACTGGTGAGTA
CGCCAATTTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGAGAGCGTCAGTGTTAAGT
GGGGGAAAATTAGATGCGTGGGAGAAAATTCGGTTACGGCCAGGGGGAAAAGAAAAAATATAGAC
TAAAACATTTAGTATGGGCAAGCAGGGAGCTGGAAAGGTTCCGCACTCAACCCTGGTCTTTTAGAA
ACATCAGAAGGGTGCCAACAATAATGGAGCAGTTACAATCAACTCTCAGAACAGGATCAGAAGA
ACTTAAGTCATTGTTAATAACAATAGCCACCTTTGGTGTGTACATCAAAGGATAGATATTAAGA
CACCAAGGAAGCCTTAGATAAAAATAGAGGAAGTGA AAAAATAAGAACAAGCAAAGACACAGCAAG
CAGCCGCCGCAGGAAGCGGCAGCCAAAATTTT

YBF274

ATGGGTAACAAATGGTCAAAAAGTAGTATAGTGGGATGGCCTGCCATAAGGGAAAGAATAAGAAG
AACTAATCCAGCAGAAAATCCAGCAGCAGATGGGGTAGGAGCAGTATCACAAGACCTGGCAAAA
CATGGGGCAATCACAAGTAGCAATACAAAACATACTAATGCTGCCTGTGCCTGGCTAGAAGCACA
AGAAGAAGAGGGGGTGGGTTTTCCAGTCAAACCTCAGGTACCTCTAAGACCAATGACCTACAAG
GGAGCTTTAGATCTTAGCCACTTTTTAAAAGAAAAGGGGGGACTGGAAGGGCTAGTTCACTCCCA
AAGAAGACAAGAGATCCTTGATCTTTGGGTCTACCACACACAAGGCTACTTCCCTGATTGGCAGA
ACTACACACCAGGGCCAGGAACCAGATACCCACTATGCTTTGGGTGGTGCTTCGAGCTAGTACC
AGTTGATCCAAAGGTGGTAGAAGAAGAGACTGAAGGAGAGAGCAATTGTTTGCTACACCCATAA
GCCAGCATGGAATGGATGACTCGGAGAGAGAAGTATTAGTGTGGAGATTTAACAGCAGACTAGC
ATATGAACACAAGGCCAGAGAGAAGCATCCGGAGTTCTTTAAAGACTGCTGACACCGAGTTCTTA
AAGACTGCTGATACTGAGCTTTCTACAAGGGACTTTCCGYTGGGGACTTTCCAGAGGAGGCGTG
ACTGGGGAGGGACTGGGGAGTGGCCAACCTCAGATACTGCATATAAGCAGCCGCTTTCTGCTT
GTACCGGGTCTCGGTTAGAGAACCAGGTCTGAGCCCGGGAGCTCCCTGGCTTCTCAATAAAGCT

TGCCTTGAGTGAGAAGCAGTGTGTGCTCATCTGTTCAACCCTGGTGTCTAGAGATCCCTCAGATC
ACGTAGACTGAAGCAGAAAAATCTCTAGCAGTGGCGCCYGAACAGGGACCAGAAAAGAGAAAGTG
AAACCAGGGAAGAAAACCTCCGACGCAACGGGCTCGGCTTAGCGGAGTGCACCCGCTAAGAGG
CGAGAGGAACTCACAGAAGGGTGAGTAACTTGCTGGCGGTGCCAGACCTAGAGGAAGGGCG
AAGTCTCTAGGGGAGGAAGATGGGTGCGAGAGCGTCTGTGTTGACAGGGAGTAAATTGRATGCA
TGGGAACAAATTAGGTTAAAGCCAGGATCTAAAAAGAAATATAGACTAAAACATTTAGTATGGGCA
AGCAGGGAGCTGGAAAGATWTGCTTGTAACTCCTGAGCTACTAGAACTGCAGAAGGCAATGAGA
AACTGWTACAACAGTTAGAACCAGCTCTCAAGACAGGGTCAGAAAAGCCTRCAATCACTCTGGAA
CGCAATAGTAGTCTGGTGTGTTTATAACAGATATAAAGTTGAAGATACGCAGCAGGCAATAC
AAAAGTTAAAGGAAGTAATGGGGAACAGGAAGCCYGCGRGTGYCGCTAAGGAAGACACTAGCAC
AARGCAGATGGGTCAAACACTAC

BCF212

ATGGGCAGCAAGTGGTCAAAAAGCAGCATAGTAGGATGGCCTGAGGTTAGGGAAAGAATGAGAA
GGACTCCTCCACATAAATACTCCTCCAACCTCCAGCAGCAGTAGGAGTAGGAGCAGTATCTCAA
GATTTAGCTAGACATGGAGCAATCACAAGCAGTAATACAGCTCACACTAATGCTGACTGCGCCTG
GCTGGAAGCACAAGAAGAAGAGGAGGTAGGCTTCCAGTCAGGCCACAGGTACCTTTAAGACCC
ATGACTTATAAAGCAGCTGTAGATCTCAGTCACTTTTTAAAAGAAAAGGGGGGACTGGATGGGTT
AATTCACCTCCAGGAGAAGACAAGACATCCTTGATCTGTGGGTCTACCACACACAAGGATACTTCC
CTGATTGGCAGAATTACACGCCAGGGCCAGGGACCAGATTCCCACTAACATTTGGGTGGTGCTT
CAAGCTAGTACCAGTAGATCAAGAGGAAGTAGAGGAGGCCACTAGGGGAGAGAAACAACAGCCTG
TTACATCCTATATGCCAACATGGAATGGATGATGATGAGAGAGAAGTGTTAAAGTGGCAGTTTGA
CAGCCGCCTGGCCCTAAGACACTTCGCTAGAGAAAAGCATCCGGAGTTCTACAAAAGACTGCTGA
CACAGAAGTTGCTGACTGGGACTTTCCGCTGGGGACTTTCCAGGGGAGGTGTGGTTTGGGCGG
AGTTAAGGAGTGGCTAACCTCAGATGCTGCWTATAAGCAGCCGCTTTTCGCTTGTACTGGGTCT
CTCTTGTTAGACCAGATCGAGCCTGGGAGCTCTCTGGAAGTAGCGAGAGAACCCTGCTTAAG
CCTCWATAAGCTTGCCATTGAGTGAGAAGCAGTGTGTGCTCATCTGTTCAACCCTGGTGTCTAGA
GATCCCTCAGATCACTTAGACTGAAGCAGAAAATCTCTAGCAGTGGCGCCGAACAGGGACTTG
AAAGAGAAAAGTGAACCAGGGAAGAAATCCTCCGACGCAACGGGCTCGGCTTAGCGGAGTGCA
CCCGCTAAGAGGCGAGAGGAACTCACAGGGGTGAGTAAAATTGCTGGCGGTGCCAGACCTAG
GGGAAGGGCGAAGTCTCTAGGGGAGGAAGATGGGTGCGAGAGCGTCTGTGTTAAGGGGAAGTA
AATTGGATGCATGGGARCAAAATTAAGTTGAAGCCAGGATCTAAAAGACATATAGACTAAAACATT
TAGTTTGGGCAAGCAGGGAGCTGGACAGATTTCGCATGTAATCCTGAGCTACTAGAAAATGCAGA
GGGTAATGAGCAACTGTTACAACAGTTAGAACCAGCTCTCAAGACAGGGTCAGAAAAGCCTGCAG
TCACTCTGGAACACAATAGCAGTGTCTGGTGTGTTCAACAAAGAATTAGAGTTGGAGATACGCA
GCAGGCAGTACAAAAGTTAAGGGAGGAAATGGAGAAAAGGAAGCCTAAAATACTGCTAAGGAA
GACACAACTCAAGGCAGACAGGTCAAATTAC

RBF237

ATGGGTRGCAAGTGGTCAAAAAGCAGTATAGTTGGATGGCCTGCTGTAAGGGAAAGACTAAGAC
AAACCCCTGTAGCAGCAGAKGGGGTGGGAGCAGTGTCTCAAGATTTAGATAAACATGGGGCAAT
TACAAGCAACAATACTAGGGCTACCAATGCTGACTTGGCCTGGCTGGAAGCGCARGAGGATGAG
GACGTAGGTTTTYCCAGTCAGACCTCAGGTACCTTTAAGGCCAATGACTTATAAAGCAGCTGTGCA
TCTCAGCCACTTTTTAAAAGAAAAGGGGGGACTGGATGGGTTAACTTACTCCAGGAGAAGAAAAG
AAATCCTTGATTTGTGGGTCTACCACACACAAGGCTTCTTCCCTGACTGGCAGTGTACACACCA
GGGCCAGGAACYAGGTTCCCACTGACCTTTGGATGGTGCTTCAAGCTAGTACCAGTTGATCCAG
AGGAAGTAGAAAGGGCYAATGAAGGAGAGAACTGTTTATTACATCCTATGAGCCTACATGGA
ATGGAGGACGACGACARGGAAGTGTAAAGATGGAAGTTTGACAGCACCCCTAGCACTGAGACT

TAGCCAAAGAGAAACATCCGGAGTACTACAAAGACTGAGAAACTGCTGACACAGAAGATCCTAAA
AGGGACTTTCCACTGGGGACTTTCCAGAAGGGCGTACCCAAGGGCGGGACCGGGGAGTGGCTA
MCCCTCAGATGCTGCATATAAGCAGCTGCTTTTCGCCTGTACTGGGTCTCTCTTGCTARACCAGA
TTTGAGCCTGGGAGCTCTCTGACTAGCAGGGAACCCACCTTGCCTTGAGTGCTTCAAGTAGTGT
GTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTCAGACCACTCTAGATAGTGTA AAAA
TCTCTAGCAGTGGCGCCCGAACAGGGACTTGAAAGTGAAAGTTAATAGGGACTCGAAAAGCGAAA
GTTCCAGAGAAGATCTCCCGACGCAGGACTCGGCTTGCTGAGGTGCACRCAGCAAGAGGCGAG
AGCGGCGACTGGTGAGTACGCCAATTTTTGACTAGCGGAGGCTAGAAGGGGAGAGATGGGTGC
GAGAGCGTCTGTRTTGT CAGGAAGCAAATTGGATGCATGGGAACAAATTAGGTTAAAGCCAGGAT
CTAAAAAGAAATATAGGCTAAAACATTTAGTATGGGCAAGCAGGGAGCTGGACAGATTGCTTGT
AATCCTGAGCTACTAGAACTGCAGAGGGTAATGAGAACTGTTACAACAGTTAGAGCCAGCTCT
TAAGACAGGGTCAGAAAGCCTGCAGTCACTCTGGAACACAATAGCAGTGCTCTGGTGTGTTCCACA
ATAGATTTAARGTTGAAGAYACACAGCAGGCAATACAAAAGCTAAAACAAGAAAAGGGGAAGTCT
GTGGGTACCGCTAAGGAAGACACAAGCGCAAGGCAGACAGGTCAAATTAC

RBF240

ATGGGGAACGCATTTAAGAAAAGYAGCTTTGTRGGATGGCCAGCAGTAAGAGAAAAGAATAAGAR
GAGCTCACAAAGGAGCCTGAACCTGAACCATGTGCACCKGGAGTGGGAGAACTCTCCAGGGCRG
TAGCAAATAGAGGAGGGATACCAAGTTCCTTTACTCCTCAAAAACAATGCAGCCCTTGCAATCCTA
GAAAGTCACCAAGAAGATGAAGAGGGAGTAGTTTTCCAGTCAAACCTCAAGTACCTGTAAGGC
CAATGACCTATAAAGGAGCATTGACCTCAGCTTCTTTTTAAAAGAAAAGGGAGGACTGGATGGG
TTAATTTACTCAACTAAGAGAGCAGAGATCCTGGATCTTTGGGTGCATCACACTCAGGGATTCT
CCCWGATTGGCAGA ACTACACACCAGGACCAGGA ACTAGATACCCACTGACATTTGGATGGTTG
TTTAAACTAGTACCAGTGT CAGAAGCTGAGGCAGAAGACATGGGCACTGAAGGTGACAACGCTA
GTCTCCTACATCCAGCTTGTGATCATGGCTTTGAAGATCCACACGGAGAAATACTGAAATGGCAG
TTTGATAGAACATTAGGCWTCATCATGTTGCTCAGAGAACCACCCAGAGCTCTTCTCTAAGGA
CTAAACTGCTAACTGAAGATTGCTGACACCGCGGGACTTTCCAGYAGAGACTGCTGACMCTGC
AGGGACTTTCCAGCGTGGGAGGGATAAAGGGCGGTCCGGGGAGTGGCTAACCTCAGAAGCTG
CATATAAGCAGCCGCTTTCTGCTTGTACTGGGTCTCAGTTAGAGAACCAGGTTAGAGCCCGGA
GCTCCCTRGCCTCTGGCTGAACCCACTGCTTAACGCTCAATAAACTTGCCTTRMGTGAGAAGCA
GTGTGTGCTCATCTGTTCAACCCTGGTAGCTAGAGATCCCTCAAATCACGTAAACTGAAGCAGAA
AAATCTCTAGCAGTGGCGCCC GAACAGGGACGCTGAAAGAGAAAGTGAACCAGGGAAGAAAAC
CTCCGACGCAACGGGCTCGGCTTAGCGGAGTGCACCTGCTAAGAGGCGAGAGGAGCTCACAGA
GGGTGAGTAAATTTGCTGGCAGGTGGCCAGACCTAGGGGAAGGGCGAAGTCCCTAGGGGAGGA
AGATGGGTGCGAGAGCGTCAGTGTTAAGGGGAAGCAAATTGRATGCATGGGAACAAATTAGGTT
RAAGTCAGGATCTAAAAAGAMATATCAGCTAAAACATTTAGTATGGGCAAGCAGGRAGCTGRATA
RATTCGCATGTAATCCTGAGCTTCTAGAACTGTAGAGGGAACTGARAAATTATTACAACAGCTAG
ARCCAGCTCTCAAGACAGGGTCAGAAAGCCTGCAGTCACTCTGGAACACTATAGCAGTGCTCTG
GTGTGTT CATAACA ACTTTAGAGTAGGAGATACGCAGCAGGCAATACAAAAGTTACAAGAAGAAA
TGAAGAAAAGGAAACCTGTAGATACTGCTAAGGAAA ACTCAAGGCAGACAGGTCA