# Report

# APOBEC4, a New Member of the AID/APOBEC Family of Polynucleotide (Deoxy)cytidine Deaminases Predicted by Computational Analysis

Igor B. Rogozin<sup>1</sup>
Malay K. Basu<sup>1</sup>
I. King Jordan<sup>1</sup>
Youri I. Pavlov<sup>2-4</sup>
Eugene V. Koonin<sup>1,\*</sup>

<sup>1</sup>National Center for Biotechnology Information NLM; National Institutes of Health; Bethesda, Maryland USA;

<sup>2</sup>Eppley Institute for Research in Cancer; <sup>3</sup>Department of Biochemistry and Molecular Biology; <sup>4</sup>Department of Pathology and Microbiology; University of Nebraska Medical Center; Omaha, Nebraska USA

\*Correspondence to: Eugene V. Koonin; National Center for Biotechnology Information NLM; National Institutes of Health; Bethesda, Maryland 20894 USA; Tel.: 301.435.5913; Fax: 301.435.7794; Email: koonin@ncbi.nlm.nih.gov

Received 07/01/05; Accepted 07/06/05

Previously published as a *Cell Cycle* E-publication: http://www.landesbioscience.com/journals/cc/abstract.php?id=1994

### **KEY WORDS**

cytidine deaminase, DNA/RNA modification, phylogenetic analysis, editing enzyme, innate immunity, APOBEC1

### **ABSTRACT**

Using iterative database searches, we identified a new subfamily of the AID/APOBEC family of RNA/DNA editing cytidine deaminases. The new subfamily, which is represented by readily identifiable orthologs in mammals, chicken, and frog, but not fishes, was designated APOBEC4. The zinc-coordinating motifs involved in catalysis and the secondary structure of the APOBEC4 deaminase domain are evolutionarily conserved, suggesting that APOBEC4 proteins are active polynucleotide (deoxy)cytidine deaminases. In reconstructed maximum likelihood phylogenetic trees, APOBEC4 forms a distinct clade with a high statistical support. APOBEC4 and APOBEC1 are joined in a moderately supported cluster clearly separated from AID, APOBEC2 and APOBEC3 subfamilies. In mammals, APOBEC4 is expressed primarily in testis which suggests the possibility that it is an editing enzyme for mRNAs involved in spermatogenesis.

## INTRODUCTION

Cytidine deaminases (CDAs; EC 3.5.4.5) catalyze the deamination of cytidine to uridine and are important in the pyrimidine salvage pathway in prokaryotes and eukaryotes. These enzymes contain a zinc-coordinating domain with the characteristic motif (H/C)xE ... PCxxC (x stands for any residue). The Zn ion in the active site plays a central role in the proposed catalytic mechanism by activating a water molecule to form a hydroxide ion that performs a nucleophilic attack on the substrate.  $^{1-3}$ 

The cytidine deaminase superfamily also includes the AID/APOBEC family, which is a vertebrate-specific RNA/DNA editing expansion of deaminases. One of the best characterized modes of mRNA-editing is cytidine to uridine (C > U) deamination catalyzed by APOBEC1. APOBEC1 is the catalytic component of a complex that edits apolipoprotein B mRNA by catalyzing the C6666U deamination which creates a premature stop codon and causes tissue-specific production of a truncated apolipoprotein B polypeptide chain.<sup>4,5</sup> The AID/APOBEC protein family contains four subfamilies (AID, APOBEC1, APOBEC2, and APOBEC3) and includes several members with an experimentally confirmed capability of deaminating cytosine to uracil in single-stranded polynucleotides, while fulfilling diverse physiological functions.<sup>6-8</sup> AID functions in the adaptive humoral immune response, namely, somatic hypermutation of the immunoglobulin V gene and switch recombination of the immunoglobulin C gene. 9,10 Two members of the diverse APOBEC3 subfamily (human APOBEC3G and APOBEC3F) are involved in an innate pathway of restriction of retrovirus infection, presumably, by deaminating cytosines in viral first-strand cDNA replication intermediates<sup>7,11,12</sup> although more complex models have also been proposed.<sup>13</sup> The physiological functions of APOBEC2<sup>14,15</sup> and of other APOBEC3s are unknown.<sup>7,8,12</sup> However, the expansion and rapid evolution of APOBEC3 proteins in the primate lineage (there are eight human APOBEC3 genes) suggests that at least some of these proteins might have antiviral functions.<sup>8</sup>

APOBEC1 was the first member of the family to be discovered<sup>4,5</sup> and has become the paradigm for subsequent studies. However, phylogenetic analysis indicates that APOBEC1 is a recent evolutionary arrival whereas AID and APOBEC2 are the ancestral family members.<sup>8</sup> We analyzed distant similarities of AID/APOBEC protein sequences and identified a previously undetected subfamily of AID/APOBEC homologs which we provisionally named APOBEC4. Phylogenetic analysis suggests that this protein subfamily is most closely related to APOBEC1, however, it has a wider phyletic distribution similar to that of AID/APOBEC2 and, accordingly, appears to have emerged early in vertebrate evolution.

### **MATERIALS AND METHODS**

The non-redundant (nr) database of National Center for Biotechnology Information (http://ncbi.nlm.nih.gov/) and vertebrate genomes at the ENSEMBL web site (http://www.ensembl.org/) were searched using the BLASTP program. 16 Nucleotide genome sequences were searched using TBLASTN with protein sequences as queries. Iterative sequence similarity searches were performed using PSI-BLAST with a single sequence used as the query and

with default parameters. <sup>16</sup> Each search was run for a minimum of three iterations or to convergence. Multiple alignments were generated using the MUSCLE program with 50 iterations. <sup>17</sup> The resulting multiple alignment was corrected manually using the PSI-BLAST results, the known three dimensional structures of CDAs, and predicted secondary structure of APOBECs as additional guides. Protein secondary structure was predicted using JPRED. <sup>18</sup>

Phylogenetic analyses were performed using minimum evolution (least-square) and maximum likelihood methods. To generate the input file, all columns containing gaps were either deleted pairwise or entirely deleted from the corrected alignment. Minimum evolution trees were constructed using

either MEGA3<sup>19</sup> with the Poisson correction model and pairwise deletion of gaps and 1000 bootstrap replicates or using the FITCH program of the PHYLIP package<sup>20</sup> with 1000 bootstrap replicates, after complete removal of all gapped columns. Maximum likelihood trees were generated using a two-step procedure. At the first step, a minimum evolution tree was generated using FITCH, and at the second step, the topology of this tree was used as the input to PROTML<sup>21</sup> to produce a maximum likelihood tree using local rearrangements. The statistical significance of the internal nodes of the resulting maximum likelihood tree was then determined using relative estimate of logarithmic likelihood bootstrap (RELL-BP) as implemented in PROTML.<sup>21</sup> In the second method, an initial tree was constructed using the PROTML program of the PHYLIP package, with star decomposition. The tree topology was used as a guide to generate maximum likelihood trees using the PhyML program with 100 bootstrap replicates generated from the input alignment.<sup>22</sup> The consensus tree of these 100 bootstrapped trees was derived using the CONSENSE program of the PHYLIP package to obtain the bootstrapped the full maximum likelihood tree. Both methods produced phylogenetic trees with the same topology with respect to the main branching.

### **RESULTS AND DISCUSSION**

Identification of APOBEC4 by sequence similarity searches. A PSI-BLAST search with the human AID sequence as the query returns hits to AID/APOBEC proteins in the first and second iterations. Unexpectedly, from the third iteration onward, this and other searches with AID/APOBEC sequences started to recover uncharacterized vertebrate proteins (Table 1). Reciprocal searches using these sequences as queries readily recovered the AID/APOBEC sequences, suggesting that these proteins and AID/APOBEC proteins are, indeed, homologs. We called this newly discovered protein subfamily APOBEC4. BLAST searches with the human APOBEC4 sequence (GI: 44888831) as a query readily identified the presence of the APOBEC4 gene in the *Xenopus tropicalis* genome. By contrast, these searches failed to identify APOBEC4 protein in any of the available (nearly) complete fish genomes.

Sequence and structural features of APOBEC4. To further examine the sequence and structural conservation between CDA, AID, APOBEC1, APOBEC2, APOBEC3, and APOBEC4, we constructed a multiple alignment

Table 1 The APOBEC4 subfamily of predicted cytidine deaminases

Species	Name	Refseq	GI	ESTs
Human	Hypothetical protein	NM_203454	44888831	8 testis, 1 brain, 1 uterus
Macaca fasicularis	Testicular cDNA	-	17026052	-
Mouse	Hypothetical protein	XM_355245	38073497	2 testis
Rat	Hypothetical protein	XP_573474	62945336	1 testis
Cow	Hypothetical protein	XP_613244	61875234	-
Chicken	Hypothetical protein	XP_426631	50751204	-
Xenopus tropicalis	Hypothetical protein GENSCAN00000137932	-	-	-

Table 2 Number of amino acid substitutions per site in human and mouse orthologs of the AID/APOBEC family

Protein	AID	APOBEC1	APOBEC2	APOBEC3	APOBEC4
Number of substitutions	0.09	0.31	0.10	1.03	0.30
Standard error	0.03	0.05	0.03	0.12	0.05

The number of amino acid substitutions and standard errors were calculated using the Poisson correction as implemented in MEGA3.<sup>19</sup> In the case of APOBEC3, the human APOBEC3F and mouse APOBEC3 proteins were used.

of 4 CDA sequences and 45 AID/APOBEC sequences (Fig. 1). In the case of the dimeric APOBEC3 proteins,8 the more conserved C-terminal domain, which determines the specificity of retrovirus hypermutation induced by human APOBEC3F and APOBEC3G,<sup>23</sup> was used for this analysis. The alignment shows notable conservation of the Zn-coordinating motif, (H/C)xE... PCx<sub>2-6</sub>C (Fig. 1). However, the proline residue, which is present in the middle of the HxE motif in most of the APOBEC4 sequences, aligns with an alanine in the CDA sequences whereas AID/APOBEC sequences contain several other amino acids in this position (Fig. 1). Interestingly, the Xenopus tropicalis APOBEC4 contains an alanine in the middle of the HxE motif (Fig. 1) suggesting that this could be the ancestral state of the motif. A distinctive feature of APOBEC4 is the insertion of four amino acids between the conserved cysteines of the PC...C motif; the presence of this insert complicates the detection of APOBEC4 in sequence similarity searches (see above). It should be noted that even longer inserts are present in this motif in several deaminases outside the CDA/AID/APOBEC superfamily,<sup>24</sup> which is compatible with the prediction that APOBEC4 is an active deam-

Structural homology models based on *E. coli* or yeast CDA structures have been previously proposed for APOBEC1, AID and APOBEC3G.  $^{25\text{-}27}$  The deaminase domain of APOBEC4 conforms to the b1b2a1b3a2b4a3b5 arrangement (a designates an  $\alpha$ -helix and b designates a  $\beta$ -strand) typical of the AID/APOBEC family rather than the b1b2a1b3a2b4b5 arrangement (with the a3 helix missing) seen in the CDAs (Fig. 1). The additional a3 helix is a signature of the AID/APOBEC family. With the exception of this helix, the predicted secondary structural elements of APOBEC4 show a nearly perfect correspondence with the elements derived from the 3D structure of CDAs, supporting the notion that these proteins contain a domain with the same fold (Fig. 1).

Phylogenetic analysis and evolutionary implications. To gain insight into the evolution of the AID/APOBEC family and, in particular, the origin of APOBEC4, we constructed phylogenetic trees from the multiple alignment shown in Figure 1. The APOBEC1, APOBEC2 and APOBEC4 subfamilies each formed a clade with high statistical support (data not shown). Unexpectedly, the members of the APOBEC3 subfamily did not form a distinct clade but instead were interspersed with the AID proteins (data not shown) although previous phylogenetic analyses suggested monophyly of

CDA_Homsa_263657					FPVGAALLTQEGRIFK				
CDA_Sacce_6323274 CDA Thema 4981379					FRVGCSILTNNDVIFT FRVGAALLTKSGKIYT				
CDA_Thema_4981379 CDA_Bacsu_80258	3	aa	OFITTEALKARD	-KAIAKISGI	FOVGAALLTKSGKIII	GCNT	ENSSIGLTVCAE	RVAVEKAV	SEG-I
CDA_Bacsu_00230	,	aa	НИНИНИНИНИН	-MATAL ISK	EEEEEEE EEE	EE		HHHHHHH HH	
APOBEC3 Musmu 26340722	204	aa		LCYYHRMKP	LCYQLEQFNGQAPLK				
APOBEC3 Crilo 48474310	204	aa	EEEFYSQFYNQRVKH	LCYYHRMKP?	LCYQLEQFNGQAPLK	GCLL	SEKGKQHAE	ILFLDKIR	SME-I
APOBEC3F_Homsa_24416443									
APOBEC3G_Macni_48476259									
APOBEC3G_Gorgo_50254066					LCYEVERLHNDTWVLLNQ				
APOBEC3G_Pantr_48476269					LCYEVERLHNDTWVLLNQ				
APOBEC3G_Lagla_48476319					LCYKVEHLRNGTWVPLHQ				
APOBEC3G_Sagla_48476309					CLCYEAEHLHSGTWVPLHQ				
APOBEC3G_Ponpy_48476299 APOBEC3B Pantr 55661948					/LCYKVERLHNDTWVLLNQ: VLCYEVKIRRGHSNLLWDT				
AROBEC3E_Pantr_55661948 AROBEC3_Canfa_57093121					LCYQLKPHEGSVIAK				
APOBEC3C Homsa 9294747					VLCFTVEGIKRRSVVSWKT				
APOBECSA Pantr 55661364					FLPCGIGGPDLSGSQ				
APOBEC3D Homsa 22907041	13				VLCYEVKIKRGRSNLLWDT				
AID Canfa 50979250					LCYVVKRRDSATSFSLDF				
AID Homsa 22297288	7	aa	RRKFLYOFKNVR	-WAKGRRET	LCYVVKRRDSATSFSLDF	GYLR	NKNGCHVE	LLFLRYIS	DWD-
AID Musmu 6753018	7	aa	OKKFLYHFKNVR	-WAKGRHETY	LCYVVKRRDSATSCSLDF	GHLR	NKSGCHVE	LLFLRYIS	DWD-
AID Galga 50729359	7	aa	RKLFLYNFKNLR	-WAKGRRETY	LCYVVKRRDSATSCSLDF	GYLR	NKMGCHVE	VLFLRYIS	AWD-
AID Ictpu 40949661	10	aa	QRKFIYHYKNVR	-WARGRNETY	CLCFVVKKRNSPDSLSFDF	GHLR	NRSGCHVE	LLFLSYLG	V
AID Danre 46487636	11				LCFVVKRRIGPDSLSFDF				
AID_Takfu_41016736	>1	aa	KFIYHYKNVR	-WARGRHETY	LCFVVKRRVGPDTLTFDF	GHLR	NRSGCHVE	LLFLRYLG	A
AID_Tetni_47221672	>4	aa	RKKFLYHYKNVR	-WARGRHETY	LCFVVKRRVGPDTLTFDF	GHLR	NRNGCHVE	LLFLRYLG	A
APOBEC2_Danre_61651784					FLCYLVDHGGEGLMR				
APOBEC2_Ratno_27681627	48	aa	VNFFKFQFRNVE	-YSSGRNKTI	FLCYVVEAQSKGGQVQATQ	GYLE	DEHAGAHAE	EAFFNTIL	PA
APOBEC2_Canfa_57094914					FLCYVVEAQGKGGQVQASR				
APOBEC2_Galga_50760475					FLCYVVETQGKESKTSR				
APOBEC2_Tetni_47228640					LLCFRVDTPGGSTEPLK				
APOBEC2_Xentr_49523039	51	aa	ASSEMBODENIUS	-ISSGRNKT	LCYTVERPEGQVFH	GILE	DEHASAHAE	PAPETSVL	ryf-
APOBEC2_Xenla_49256526	54	aa	ASSEMFORKNVE	-YSSGRNKT	CLLYEIKWGNQNIWRH	GYLE	DEHVSAHAE	EAFFTSVL	PQF-
APOBEC1_Mondo_23396444	19	aa	PWEFVAFFNPQE	TRKET(	CLLYEINWGGRHSVWR	umo	SNQNTSQHAE	INFMERFT	AERE
APOBEC1_Musmu_13624299 APOBEC1 Mesau 12002871	19	aa	PHEFEVER DERE	IRKEI(	CLLYEIRWGGRHNIWR	HTS	ONTSDHUE	VNFLEKFT	CEDA
APOBEC1_Mesau_12002871 APOBEC1 Ponpy 48476239	31	22	CMEEDAFFDQGE	IRKEI	CLLYEIKWGMSRKIWRSSG	nig	VNISKHVE	UNEIRKFI	CEDE
APOBEC1_Ponpy_46476239 APOBEC1_Orycu_627785	10	22	DWFFFUFFDDOF	IRKEI	CLLYEIKWGASSKTWRSSG		KNIINHVE	VNFIRKFT	CECE
APOBEC1 Homsa 2696116					CLLYEIKWGMSRKIWRSSG				
APOBEC1 Ratno 6978519	19	aa	PHEFEVEFDPRE	IRKETO	CLLYEINWGGRHSIWR	HTS	ONTNKHVE	VNFIEKFT	TERY
APOBEC4 Bosta 61875234					TFYELKTSSGSLVQK				
APOBEC4 Ratno 62945336					TFYELRSSSGNLIQK				
APOBEC4 Macfa 17026052	46	aa	EFCOIFGFPYGT	TYPOTKHI	TFYELKTSSGSLVQK	GHAS	SCTGNYIHPE	SMLFEMNGYL	DSAI
APOBEC4 Homsa 44888831					TFYELKTSSGSLVQK				
APOBEC4 Musmu 38073497					TFYELRSSSKNLIQK				
APOBEC4_Galga_50751204		aa	EFT.DAFGEDCDT	TRUDOTHUI	LLFYELKSFSGTVVQK	CHAT	CCAFODMUDE	SMI.FEADGVI.	DAVI
PODEGA Verebro									
APOBEC4_xentr	46	aa	EFYEAFGFPYGP	TMPENKQI	LIFYEVKDFSGTNIQK		NCISSNIHAE	SILFEDSGYL	DALL
APOBEC4_Xentr	46	aa		TMPENKQI			NCISSNIHAE		
_			EFYEAFGFPYGP HHHEE	TMPENKQI	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2	GQVT	NCISSNIHAE	SILFEDSGYL HHHHH a1	DALL EE
			EFYEAFGFPYGP HHHEE -FRAIAIASDMODDF:	TMPENKQI I ISPOGA	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CROVMREFGTNWP-VYMT-	GQVT	NCISSNIHAE	SILFEDSGYL HHHHH al	DALI EE 16
	 G		EFYEAFGFPYGP HHHEE -FRAIAIASDMQDDF: -WKCMVICGDSEDQC	TMPENKQI I ISPCGA VSPCGV	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2 CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFPIVMLN	GQVT	NCISSNIHAE	SILFEDSGYL HHHHH al PDGTYIVMTVQ GSRSKVMTMG	DALI EE 16
	G		EFYEAFGFPYGP HHHEE -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKT-	TMPENKQI ISPCGA VSPCGV	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFPIVMLN CRQVLYEF-SDDLDVIMA-	GQVT	NCISSNIHAE	SILFEDSGYL HHHHH al PDGTYIVMTVQ GSRSKVMTMG	DALL EE 16 13 14
	G		EFYEAFGFPYGP HHHEE -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV-	TMPENKQI I ISPCGA VSPCGV -APCGA -SPCGA	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT CRQVLVEF-SDDLDVIMA- GRQVISELCTKDVIVVLT-	GQVT	NCISSNIHAE	SILFEDSGYL HHHHH al PDGTYIVMTVQ PGSRSKVMTMG RDGNFEIVKLK LQGQIKEMTVE	DALL EE 16 13 14
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC' -FVAIAIASDSPDKTFQMLAVADTPGPV- EEEEEE	TMPENKQI ISPCGA VSPCGVAPCGA SPCGA HH	LIFYEVKDFSGTNIQK EEEEEEE EEEE bl b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFPIVMLN CRQVIVEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE	GQVT	KPISSNIHAE	SILFEDSGYL HHHHH al PDGTYIVMTVQ PGSRSKVMTMG KDGNFEIVKLK QGQIKEMTVE EEEEE	16 13 14 18
	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMIAVAADTPGPV- EEEEEEQVTITCYLTW	TMPENKQI ISPCGA VSPCGVAPCGASPCGA HH -SPCPN	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  GROVMREFGTNWP-VYMT- CROFINEFVVKDFPIVMLN CROVLYEF-SDDLDVIMA- GROVISELCTKDVIVVLT- HHHHHHHH EEEEE GAWQLAAFKRDRPDLILHI	GQVT		SILFEDSGYL HHHHH al PDGTYIVMTVQ TGSRSKVMTMG RDGNFEIVKLK LQGQIKEMTVE EEEEE GGILVDVMDLP	16 13 14 18
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Crilo_48474310	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPKT: -FQMLAVAADTPGPV- EEEEEEQVTITCYLTWQVTITCYLTW	ISPCGA VSPCGVAPCGASPCGAHH -SPCPN	LIFYEVKDFSGTNIQK EEEEEEE EEEE bl D2  CRQVMREFGTNWP-VYMT- CRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI GAWRLAAFKRDRPDLILHI GAWRLAAFKRDRPDLILHI	GQVT		SILFEDSGYL HHHHH al PDGTYIVMTVQ PGSRSKVMTMG RDGNFEIVKLK APGQIKEMTVE EEEEE GGILVDVMDLP BGILVDVMDLP	DALI EE 16 13 14 18 44 59
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258 APOBEC3_Musmu_26340722 APOBEC3_Crilo_48474310 APOBEC3_THomsa_24416443	G  N		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLTW	TMPENKQI ISPCGA VSPCGVAPCGA SPCGA HH -SPCPN SPCPN	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFPIVMLN CRQVLVEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAAFKARPRDRLILHI CAGEVAAFLARHSNVNLTI	GQVT  YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT-		SILFEDSGYL HHHHH al PDGTYIVMTVQ GSRSKVMTMG EDGNFEIVKLK LOGQIKEMTVE EEEEE GGILVDVMDLP GGLVDVMDLP LGASVEIMGYK	16 13 14 18 44 59
CDA_Homsa_263657  CDA_Sacce_6323274  CDA_Thema_4981379  CDA_Bacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Criio_48474310  APOBEC3F_Homsa_24416443  APOBEC3G_Macni_48476259	G  N		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMIAVAADTEGPV- EEEEEEQVTITCYLTWTNYEVTWYTSWQQYRYTCFTSW	TMPENKQI ISPCGA VSPCGVSPCGA HH -SPCPNSPCPNSPCPS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GROVMREFGTNWP-VYMT- CROFINEFVVKDFPIVMLN CROVLYEF-SDDLDVIMA- GROVISELCTKDVIVVLT- HHHHHHHH EEEEE GAWQLAAFKRDRPDLILHI CAWRLAAFKRDRPDLILHI CAGEVAAFLARHSNVNLTI GAGEWAKFISNNEHVSLCI GAGEMAKFISNNEHVSLCI	GQVT		SILFEDSGYL HHHHH al PDGTYIVMTVQ GSRSKVMTMG GDGNFEIVKLK QGQIKEMTVE EEEEE GILVUVMDLP GGILVUVMDLP GGASVEIMGYK GGAKIAMMNYS	DALI EE 16 13 14 18 44 59 39 43
CDA_Homsa_263657  CDA_Sacce_6323274  CDA_Thema_4981379  CDA_Bacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Crilo_48474310  APOBEC3F_Homsa_24416443  APOBEC3G_Macni_4847629  APOBEC3G_Gorgo_50254066	G  N		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWTNYEVTWYTSWQDYRVTCFTSWQDYRVTCFTSW	TMPENKQI ISPCGA /SPCGASPCGA SPCPN SPCPESPCFS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- RQRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNKEWSLCI CAQEMAKFISNKEWSLCI	GQVT YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR-		SILFEDSGYL HHHHH a1 PDGTYIVMTVQ GSRSKVMTMG RDGNFEIVKLK LQGQIKEMTVE EEEEE GGILVDVMDLP GGASVEIMGYK RDGAKIAMMNYS RGAKIAMMNYS RGAKIAMMYS	DALI EE 16 13 14 18 44 59 39 43 43
CDA Homsa 263657  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 Crilo 48474313  APOBEC3 Homsa 24416443  APOBEC3G Macni 48476259  APOBEC3G Fantr 48476269	G N H		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEQVTITCYLTWQVTITCYLTWQDYRVTCFTSWQDYRVTCFTSWQDYRVTCFTSW	TMPENKQI ISPCGA YSPCGVSPCGASPCFA SPCPNSPCFSSPCFSSPCFS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLVEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI	GQVT		SILFEDSGYL HHHHH a1 PDGTYIVMTVQ GSRSKVMTMG RDGNFEIVKLK QGQIKEMTVE EEEEE GGILVDVMDLP GGASVEIMGYK GGAKISIMTYS GGAKISIMTYS	16 13 14 18 44 59 39 43 43
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Crilo_48474310 APOBEC3F_Homsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319	G N H A		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFOMLAVAADTFGPV- EEEEEEQVTITCYLTWTNYEVTWYTSWQQYRVTCFTSWQDYRVTCFTSW	TMPENKQI ISPCGA	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GROVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- GROVISELCTKDVIVVLT- HHHHHHH EEEE GAWQLAAFKRDRPDLILHI CAWGLAAFKRDRPDLILHI CAGEVAAFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAGEKVAEFLQENPHVNLHI	YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- SAARIYDYQRG-		SILFEDSGYL HHHHH a1 PGGTYIVMTVQ GGSRSKVMTMG RDGNFEIVKLK AGGIKEMTVE EEEEE GGLVDVMDLP GGASVEIMGYK GGAKIAMNYS GGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGAKISIMTYS	DALI EE 16 13 14 18 44 59 39 43 43 43
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Crilo_48474310 APOBEC3G_Macni_48476259 APOBEC3G_Gorgo_50254066 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Sagla_48476309	G N H A		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTITCYLIWTNYEVTWYTSWQDYRVTCFTSWQDYRVTCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISW	TMPENKQI I ISPCGA SPCGV APCGA HH SPCPN SPCPS SPCFS SPCFS SPCFS SPCFS	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEE EOO CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKBORPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEWSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAGEVAEFLLENPHVNLHI	YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- SAARIYDDQGR- SAARIYDYRG- FAARIYDYRG-		SILFEDSGYL HHHHH al PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK **QQIKEMTVE EEEEE GILVDVMDLP **GILVDVMDLP **GASVEIMGYK **GAKISIMTYS **GAKISIMTYS **GAKISIMTYS **GAFISIMKYS	16 13 14 18 44 59 39 43 43 43
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Flomsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Gorgo_50254064 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Lagla_48476309 APOBEC3G_Sagla_48476309 APOBEC3G_Ponpy_48476299	G N H A		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLTWQVTITCYLTWQDYRVTCFTSWQDYRVTCFTSWQDYRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSW	TMPENKQI ISPCGA SPCGA HHSPCPNSPCPS SPCFSSPCFSSPCFSSPCFSSPCFSSPCFSSPCFS	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEE EOO CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKBORPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEWSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAGEVAEFLLENPHVNLHI	TYTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- SAARIYDDQGR- SAARIYDYRG- FAARIYDYRG- FAARIYDDQGR-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK LQGQIKEMTVE EEEEE GGILVDVMDLP KGASVEIMGYK KGAKISIMTYS KGAKISIMTYS KGAPISMKYS KGAPISMKYS KAPISMKYS KAPISMYD	166 13 144 18 44 59 39 43 43 43 43
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_FLomsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Sagla_48476319 APOBEC3G_Sagla_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476399 APOBEC3G_Ponpy_48476399 APOBEC3G_Ponpy_48476399 APOBEC3G_Pantr_55661948	G N H H K Y		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAWADTPGPV- EEEEEEQVTITCYLIWQVTITCYLIWQVTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISWCTYRVTCFISWKCFQITWFVSWQKFEITCYVTW	TMPENKQI I ISPCGA SPCGV APCGA HH SPCPN SPCPS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFR TPCPD	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKEDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNKHVSLCI CAQEWAKFISNNKHVSLCI CAQEWAKFISNNKHVSLCI CAQEWAKFISNNKHVSLCI CAQEWAKFISNNKHVSLCI CACKVAEFLQENPHVNLHI CAQEWAKFISNNGHVSLCI CAQEVAEFLHENPHVNLHI CAQEWAKFISNNGHVSLCI CACKALAFVADHPHISIRL			SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQGQIKEMTVE EEEEE GILVDVMDLP CGASVEIMGYK GGAKISIMTYS KGAKISIMTYS KGARISIMTYS KGAPISMMKYS KEAKISIMTYD KGARVKIMDDE KGARVKIMDDE KGARVKIMDDE KGARVKIMDDE KGARVKIMDDE	DALI EE 16 13 14 18 18 44 59 39 43 43 43 43 118
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Flomsa_24416443 APOBEC3_Macni_48476259 APOBEC3G_Macni_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Tonga_703121 APOBEC3C_Canfa_57093121 APOBEC3C_Canfa_57093121	G N H A K Y N		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTFGPV EEEEEQVTITCYLTWQVTITCYLTWQUTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISW	TMPENKQI ISPCGA SPCGA HH -SPCPN SPCPN SPCFS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLVEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- CAMPLAAFKRDRPDLILHI CAGEVAAFFLARHSNVNLTI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFLARHSNVNLTI CAQEMAKFISNNCHVSLCI CAQEWAEFLARHSNVNLTI CAQEMAKFISNNQHVSLCI CAGEVAEFLARHSNVNLTI CAKKLIAFVNDHPHISLRI CAKKLIAFVNDHPHISLRI CAKKLIAFVNDHPHISLRI CAGEVAEFLARHSNVNLTI	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- SAARIYDYRG- FAARIYDYRG- FAARIYDYRG- FAARIYDYWRC- SAARLYYYWRC- FASRLYFHWRQ- FTARLYFFWRQ-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSRSKVMTMG RDGNFEIVKLK QGQIKEMTVE GEEEE GGILVDVMDLP RGASVEIMGYK GGAKISIMTYS GGAKISIMTYS GGAFISMMKYS KGAPISMMKYS KGAPISMMYS KGAPISMAS KGAPISMA	16 13 14 18 44 59 39 43 43 43 43 43 118 143
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3F_Homsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Sagla_48476319 APOBEC3G_Pantr_55661948 AROBEC3C_Gafa_57093121 APOBEC3C_Homsa_9294747 APOBEC3C_Pantr_55661364	G N H A K Y S		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFOMLAVAADTFGPV- EEEEEQVTITCYLTWQVTTCYLTWQQYRVTCFTSWQDYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQTYRVTCFTSWQKFEITCYVTW	TMPENKQI ISPCGA YSPCGVAPCGA HH -SPCPNSPCPSSPCFS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHH EEEE GAWQLAAFKRDRPDLILHI CAWELAAFKRDRPDLILHI CAGEWAEFLARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNWHVSLCI CAQEMAKFICHNHVNLHI CAQEVAEFLHENPHVNLHI CAQEVAEFLHENPHVNLHI CAGEVAEFLACHHNVILTI CAKKLIAFVNDHPHISLRL CAGEVAEFLARHSNVNLTI CAGGVAEFLARHSNVNLTI CAGGVAEFLGENTHVRIRI	YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYRFG- FAARIYDYRFG- FAARIYDYRFG- FAARIYDYRFG- FAARIYTYRFR- FASRLYFHWRQ- FTARLYYFWRQ- FTARLYYFYUP-		SILFEDSGYL HHHHH a1 PGGTYIVMTVQ GGSRSKVMTMG RDGNFEIVKLK QGGIKEMTVE GEEEE GGILVDVMDLP GGASVEIMGK GGAKIAMMNYS GGAKIAMMNYS GGAKIAMMYS GGAFISMMKYS GGAPISMMKYS GGAPISMMKYS GGAPISMMKYS GGAPISMMYS GGAPISMMS	16 13 14 18 44 59 39 43 43 43 43 43 43 278
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Flomsa_24416443 APOBEC3G_Gorgo_50254066 APOBEC3G_Pantr_48476259 APOBEC3G_Pantr_48476309 APOBEC3G_Pantr_55661948 APOBEC3G_Pantr_55661948 APOBEC3G_Fantr_55661948 APOBEC3G_Fantr_55661948 APOBEC3G_Homsa_9294747 APOBEC3C_Homsa_9294747			EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTITCYLTWQVTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISWCTYRVTCFISWCKFEITCYVTWKKFEITCYVTWTKYQVTWYTSW	TMPENKQI ISPCGA SPCGV APCGA HH SPCPN SPCPS SPCFS	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA CRQVISELCTKDVIVVLT HHHHHHHH EEEEE CAWQLAAFKEDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEWSLCI CAQEMAKFISNNEWSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNGHVSLCI CAQEWAEFLENDHVNLHI CAQEMAKFISNNGHVSLCI CVAKLAKFLAEHENVTLTI CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLQENTHVRLRI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLARHSNVNLTI CAGGVAFLERHPNVTLTI	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDVRFG- FAARIYDVRFG- FAARIYDDQGR- SAARLYYYBRFG- FASRLYFHWRQ- FTARLYYFQYP- FAARIYDDPL- FAARIYDYPD- FAARIYDYPL- SAARLYYYBRR-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK AQGQIKEMTVE EEEEE GGILVDVMDLP CGASVEIMGYK GGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGARISIMTYS AGARISIMTYS AGARISIMTYS GAPISMMKYS AGARISIMTYD GARVKIMDDE GGAVUSIMTYD AGARVKIMDDE AGARVKIMDYE	DALL EE  16 13 14 18 44 59 39 43 43 43 43 43 43 43 43 118 1443 278
CDA Homsa 263657  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Thema 4981379  CDA_Bacsu_80258  APOBEC3_Musmu_26340722  APOBEC3 Crilo 48474310  APOBEC3G_Macni_48476259  APOBEC3G_Bart_48476269  APOBEC3G_Part_48476269  APOBEC3G_Ponpy_48476299  APOBEC3G_Ponpy_48476299  APOBEC3C_Homsa_9294747  APOBEC3C_Homsa_9294747  APOBEC3D_Homsa_9294747  APOBEC3D_Homsa_9297041  APOBEC3D_Homsa_92907041  APOBEC3D_Homsa_92907041  APOBEC3D_Homsa_92907041	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTFGPV EEEEEQVTITCYLTWQVTITCYLTWQUTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISW	TMPENKQI ISPCGA SPCGA HH -SPCPN SPCPN SPCFS SPCFS SPCFS SPCFS SPCFS SPCFB SPCYD	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEHVSLCI CAQEVAEFLSNKKHVSLCI CAQEVAEFLOENPHVNLHI CAQEVAEFLOENPHVNLHI CAQEVAEFLOENPHVNLHI CAQEVAEFLARHSNVSLCI CAGEVAEFLGRHNNCHVSLCI CAGEVAEFLGRHNNCHVSLCI CAGEVAEFLGRHNNCHVSLCI CAGEVAEFLGRHNNCHVSLCI CAGEVAEFLGRHNNCHVSLCI CAKKLIAFVNDHPHISLRI CAGEVAEFLARHSNVNLTI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLHI CAGCVAFTLGRHNNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI	YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGG- SAARIYDYRGG- FAARIYDDQGGR- SAARIYDYRGG- FAARIYDYWRG- FAARIYYYWRG- FTARLYYFWRQ- FTARLYYFQYD- FAARIYYFQD- SAARLYYYRDR- SAARLYYYRDR-		SILFEDSGYLI HHHHH a1  POGTYIVMTVQ GSRSKVMTMG ROGREEIVKLK QGQIKEMTVE GEEEE GGILVDVMDLP ROGAVEIMGYK GGAKISIMTYS GGAKISIMTYS GGAKISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYD GGARVKIMDDE GGARVKIMDTE	DALI EE  166 131 144 18 44 59 39 433 433 433 433 431 114 1278 118 56
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_FLOMsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Sagla_48476319 APOBEC3G_Pantr_55661948 AROBEC3C_Homsa_9294747 APOBEC3C_Homsa_9294747 APOBEC3D_Homsa_22907041 AID_Canfa_50979250 AID_Homsa_22297288	G N H A S N N V V		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTITCYLWQVTRYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISWQTYRVTCFISWCTYRVTCFTSWCKFQIIWFVSWTKYQVTWYTSWCIYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSW	TMPENKQI ISPCGA SPCGA HH	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEWAEFISNKHVSLCI CAQEWAEFISNKHVSLCI CAQEWAEFISNNCHVNLRI CAQEVAEFLRHSNVNLTI CAKLIAFVNDHPHISLRL CAGEVAEFLAHBHPNVTLTI CAKKLIAFVNDHPHISLRL CAGGURAFLARHSNVNLTI CAGGURAFLARHSNVNLTI CAGGURAFLARHSNVNLTI CAGGURAFLARHSNVNLTI CAGGURAFLARHSNVNLTI CARCHVADFLEGYPNLSLRI CARHVADFLEGYPNLSLRI CARHVADFLEGYPNLSLRI	TYTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYREG- FAARIYDYWER- SAARLYYWER- FFASRLYFHWRQ- FTARLYYFQYP- FAARIYYFQYP- SAARLYYYCDR- FAARIYYCDR- FFARLYFCEDR- FTARLYFCEDR-		SILFEDSGYL HHHHH  a1  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGIKLMTWHOL GGAVEIMMYN GGAKIAMMNYS GGAKIAMMYS GGAKISIMTYS AGARVISIMTYS GARVISIMTYS GARVISIMTYS GARVISIMTYD GGARVISIMTYD GGARVISIMTYD GGARVIMTYD GGARVIMTY	DALI EE 166 133 144 188 444 599 399 433 433 433 433 433 278 1188 566 566
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Flomsa_24416443 APOBEC3G_Grag_50254066 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476309 APOBEC3G_Pantr_55661948 APOBEC3G_Pantr_55661948 APOBEC3G_Homsa_9294747 APOBEC3G_Homsa_9294747 APOBEC3G_Homsa_9294747 APOBEC3G_Homsa_9294747 APOBEC3G_Homsa_9294741 APOBEC3G_HOMSA_929474	G N A K N N G G G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEQVTITCYLIWQVTITCYLIWQVTYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISWCTYRVTCFISWCYFRVTCFISWKFQITWFVSWKFQITWFVSWCKFEICYVTW	TMPENKQI  ISPCGA SPCGVSPCGA	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYRFG- FAARIYDYWBR- FAARIYFWRQ- FTARLYYFWRQ- FTARLYYFQYP- FAARIYDYDB- FAARIYYPRDR- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK LQGQIKEMTVE EEEEE GGILVDVMDLP CGASVEIMGYK KGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGAFISMMKYS LEAKISIMTYD KGAFISMKYS KGYQIAIMTFK KGYQIAIMTFK KGYQIAIMTFK KGYQIAIMTFK	DALI EE 166 133 144 188 445 59 39 433 433 433 433 433 433 566 566 566
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Flomsa 24416443 APOBEC3G Macni 48476259 APOBEC3G Pantr 48476269 APOBEC3G Pantr 48476269 APOBEC3G Pantr 48476269 APOBEC3G Lagla 48476319 APOBEC3G Ponpy 48476299 APOBEC3G Ponpy 48476299 APOBEC3G Homsa 2994747 APOBEC3D Homsa 2997041 APOBEC3D Homsa 22997041 APOBEC3D Homsa 22997041 APOBEC3D Homsa 22997041 APOBEC3D Homsa 22997288 APOBEC3D Homsa 22997288 AND Mac 2797288 AND Mac 2797289 AND Mac 2797288 AND Mac 2797288 AND Mac 2797288 AND Mac 27972859	G H H Y N G G G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTFGPV EEEEEQVTITCYLTWQVTITCYLTWQUTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSW	TMPENKQI ISPCGA SPCGA HHSPCPN SPCPN SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFB SPCFB SPCFB SPCFB SPCPD SPCPD SPCYD SPCYD SPCYD	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHH EEEE GAWQLAAFKRDRPDLILHI CAWRLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEHVSLCI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNEHVSLCI CAQEWAEFLOENPHVNLHI CAQEVAEFLHENPHVNLHI CAQEVAEFLHENPHVNLHI CAQEVAEFLARHSNVNLTIT CAGCVAEFLARHSNVNLTIT CAGCVAFTLOENTHVRLEI CAGEVAEFLOENTHVRLEI CAGEVAEFLARHSNVNLTIT CAGREVAEFLARHSNVNLTIT CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGNPNLSLEI CARHVADFLERNNPNLSLEI CARHVADFLERNPNLSLEI CARH	YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGG- FAARIYDDQGG- FAARIYDDQGG- FAARIYDYRGG- FAARIYDYWBC- FAARIYYFWRQ- FTARLYFWRQ- FTARLYYFQYD- FAARIYYFQYD- FAARIYYFQD- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR-		SILFEDSGYLI HHHHH a1  POGTYIVMTVQ GSRSKVMTMG ROGREEIVKLK QGQIKEMTVE EEEEE GGILVDVMDLP GGASVEIMGYK GGAKISIMTYS GGAKISIMTYS GGAKISIMTYS GGARISIMTYS GGARISIMTYS GGARISIMTYS GGPLAWMYS GPLAWMYS GGPLAWMYS GGPLAWM	DALI EE 16 13 14 18 44 59 39 43 43 43 43 43 118 143 278 118 56 56 56
CDA Homsa 263657  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 F. Homsa 24416443  APOBEC3G Macni 48474310  APOBEC3G Macni 48476259  APOBEC3G Pantr 48476269  APOBEC3G Lagla 48476319  APOBEC3G Lagla 48476319  APOBEC3G Pantr 55661948  AROBEC3 Canfa 57093121  APOBEC3C Homsa 9294747  APOBEC3D Homsa 22907041  APOBEC3D Homsa 22907041  APOBEC3D Homsa 22297288	G N H A Y S V G G G G G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTTCYLTWQVTRYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISWQTYRVTCFISWCTYRVTCFISWKCFQITWFVSWKCFQITWFVSWTKYQVTWFTSWREYGITWFVSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSW	TMPENKQI ISPCGA SPCGA SPCGA SPCPN SPCPE SPCFS SPCFS SPCFS SPCFR SPCPT SPCPD SPCPD SPCYD	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNKHVSLCI CAQEWAKFISNKHVSLCI CAQEWAKFISNKHVSLCI CAQEWAKFISNNKHVSLCI CAQEWAFFISNNCHVSLCI CAQEVAEFLLENPHVNLHI CAQEVAEFLLENPHVNLHI CACKLIAFVNDHPHISIRL CAGGVRAFLAEHPNVTLTI CAGKLIAFVNDHPHISIRL CAGGVRAFLGENTHVILKI CARVADFLEGNPNLSLEI CARHVADFLEGNPNLSLEI CARHVADFLEGNPNLSLEI CARHVADFLEGNPNLSLEI CARHVADFLERNPNLTLEI CARHVADFLERNPMLTLEI CARHVADFLERNPMLTLEI CARHVADFLERNPMCHNLTLEI CARHVADFLERNPMCMPNLTLEI CARHVADFLER	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTYWER- FAARIYTHWRQ- FTARLYYFWER- FAARIYTOPL- SAARLYYFOP- FAARIYFCEDR- FTARLYFCEDR-		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG ROGNEETVKLK AQQIKEMTY EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGARVSIMTYS AGARVSIMTYS AGARVSIMTYD AGARVKIMDDE GIPLAVMSYL GGARVKIMDDE AGARVKIMDDE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVAIMTYK AGVOIGIMTFK AGVOIGIMTFK AGVOIGIMTFK AGVOIGIMTFK AGVOIGIMTFK AGAQUSIMTY	DALI EE 166 133 144 188 445 593 433 433 433 433 433 118 143 278 118 566 566 566 566
CDA Homsa 263657  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 Finmsa 24416443  APOBEC3 Forsa 24416443  APOBEC3G Macni 48476259  APOBEC3G Pantr 48476269  APOBEC3G Pantr 48476309  APOBEC3G Pantr 55661948  APOBEC3B Canfa 57093121  APOBEC3B Canfa 57093121  APOBEC3C Homsa 9294747  APOBEC3C Homsa 9294747  APOBEC3C Homsa 9294747  APOBEC3D Homsa 2297041  APOBEC3D Momsa 2297041  APOBEC3D Momsa 2297081  APOBEC3D Momsa 2297288  APOBEC3D Homsa 64876309  APOBEC3D Homsa 648763661  APOBEC3D Logida 650904041  APOBEC3D Logida 650		0 aa.0 a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEQVTITCYLIWQVTYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTRVTCFTSWQTRVTCFISWQTRVTCFISWCYFRVTCFISWCYFRVTCFISWCKFEITCYVTW	TMPENKQI  ISPCGA SPCGASPCGASPCPNSPCPNSPCFSSPCFSSPCFSSPCFSSPCFSSPCPDSPCPDSPCPDSPCYDSPCXN	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- YTSRLYFHWKR- FTARLYFHWKR- FTARLYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDPL- FAARIYFWRD- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDE- FVSRLYFCDEE-		SILFEDSGYL HHHHH  a1  POGTYIVMTVQ PGSRSKVMTMG ROGRIFEIVKLK LQGQIKEMTVE EEEEE GGILVDVMDLP RGASVEIMGYK RGAKISIMTYS RGAKISIMTYS RGAKISIMTYS RGAFISMMKYS LEAKISIMTYD RGAPVKIMDDE RGAPVSIMTYD RGAPVKIMDDE RGAVSIMTYD RGAPVKIMDDE RGAVSIMTYD RGAPVKIMDYE RGVQIAIMTFK RGVQIAIMTK RGVQISVMTYK	DALI EE 16 13 144 18 44 59 43 43 43 43 43 43 148 118 56 56 56 56 55 55
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Homsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3D_Homsa_22907041 APOBEC3D_HOMSA_2907041 APOBEC3D_HOMSA_29	G H H K Y N G G G G G G G G G G G	.0 aa 0 aa 1 a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMAVAADTFGPV EEEEEQVTITCYLTWQVTITCYLTWQUTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQKFEITCYVTWTKYQVTWYTSWREFQITWFVSWREFQITWFVSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRUTWFTSWRCYRUTWFTSWRCYRUTWFTSWRCYRUTWFTSW	TMPENKQI ISPCGA SPCGA HHSPCPN SPCPN SPCPS SPCFS SPCFS SPCFS SPCFS SPCFB SPCPD SPCPD SPCPD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYN	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNEHVSLCI CAQEWAKFISNNCHVSLCI CAQEWAKFISNNCHVSLCI CAQEWAKFISNNCHVSLCI CAQEWAEFLARHSNVNLTI CAQEVAEFLHENPHVNLHI CAQEVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CARHVADFLGGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLTLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- SAARIYDYRGG- FAARIYDDQGR- SAARLYYYWRR- FAARLYYYWR- FAARLYYFQYP- FAARLYYFQYP- FAARLYYFCDP- FTARLYFCEDR- FVSRLYFCDEE- FVSRLYFCDEE-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSSSKVMTMG ROGREEVEL GGGLVDVMDLP GGASVEIMGYR GGAVEIMGYR GGAKISIMTYS GGARISIMTYS GGARISIMTYS GGAPISMMKYS GGPLSMMKYS GGPLSMMKYS GGPLSMMYS GGVGISMMYS GGVGISMMYS GGVGISMMYS GGVGISMMYS GGVGISMMYS GGVGISMMYS GGVGISMMSYK GGVGISMMSYK	DALI EE  166 133 144 18 44 59 39 433 433 433 433 433 635 566 566 556 555
CDA Homsa 263657  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 F. Homsa 24416443  APOBEC3 Macni 48474310  APOBEC3 Macni 48476259  APOBEC3G Fantr 48476269  APOBEC3G Sagla 48476319  APOBEC3G Lagla 48476319  APOBEC3G Pantr 55661948  AROBEC3C Homsa 9294747  APOBEC3D Homsa 22907041  APOBEC3D Homsa 22907041  APOBEC3D Homsa 22297288  AID Homsa 22297288  AID Galga 50729359  AID Jactu 408486161  AID Janre 46487636  AID Jare 46487636  AID Tetni 47021672	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTTCYLTWQVTTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISWCTYRVTCFISWKCFQITWFVSWKCFQITWFVSWTKYQVTWFTSWCIYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRYRVTWFTSWRYRVTWFTSWRYRVTWFTSWRYRVTWFTSWRYRVTWFTSWRYRVTWFTSWRYRVTWFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCPN SPCPE SPCPS SPCFS SPCFS SPCFS SPCFR SPCPT SPCPD SPCFB SPCFB SPCPD SPCYD SPCSN SPCSN SPCSN SPCSN	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLSELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEMAKFISNKHVSLCI CAQEWAEFLENPHVNLHI CAQEVAEFLLENPHVNLHI CAQEVAEFLLENPHVNLHI CAGEVAEFLLENPHVNLHI CAGEVAEFLAHSNVNLTI CAKKLTAFVNDHPHISTRL CAGEVAEFLAHSNVNLTI CAGEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEROPNLISLEI CAGUCHAFTSGTPNLTREI CSIQLCQFLENTPNLRLEI CSIQLCQFLENTPNLRLEI	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTYMER- FAARIYDTYMER- FAARIYFTQYP- FAARIYFTQYP- FAARIYFTQP- FAARIYFTCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FVSRLYFCDLE- FVSRLYFCDLE- FVSRLYFCDLE-		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG ROGNEFEIVKLK AQQIKEMTY EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGARVKIMDDE GGIPLAWMSYL GGAPVEMMKYS AGARVKIMDDE AGARVKIMDDE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVKIMDTE AGARVAIMTYK AGVQIGIMTFK AGAQUSIMTFK AGAQUSIMTFK AGAQUSIMTYK AGAQUSIWTYK AGAQUSIWTYK AGAQUSIWTYK AGAQUSIWTYK AGVATISWMSYK AGAGTFSIAGAN	DALI EE 16 13 144 18 44 599 399 433 433 433 433 455 56 56 55 55 531
CDA Homsa 263657  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Dacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Homsa 24416443  APOBEC3G_Macni_48476259  APOBEC3G_Barni_48476269  APOBEC3G_Sarla_48476319  APOBEC3G_Sarla_48476319  APOBEC3G_Ponpy_48476299  APOBEC3G_Ponpy_48476299  APOBEC3G_Homsa 9294747  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661364  APOBEC3B_Pantr_55661364  APOBEC3B_Pantr_55661364  APOBEC3B_Pantr_55661364  APOBEC3B_Pantr_55661364  APOBEC3B_CHOMSa_2297288  AID_Canfa_50979250  AID_Homsa_22297288  AID_Galga_50729359  AID_Galga_50729359  AID_Takfu_41016736  AID_Takfu_41016736  AID_Takfu_41016736  AID_Tartni_47221672  APOBECC2_Danre_61651784		0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEE	TMPENKQI  ISPCGA SPCGVSPCGASPCPNSPCPSSPCFSSPCFSSPCFSSPCFSSPCFSSPCFSSPCPDSPCPDSPCPDSPCPDSPCYDSPCXNSPCXNSPCXNSPCXNSPCXN	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYOPLFOARIYFGP- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FVSRLYFCDE-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ PGSRSKVMTMG ROGRIFEIVKLK LQGQIKEMTVE EEEEE GILVDVMDLP RGASVEIMGYK RGAKISIMTYS RGAKISIMTYS RGAKISIMTYS RGAKISIMTYS RGAFISMMKYS LEAKISIMTYD RGAPVKIMDDE RGAVVIMDDE RGAVVIMDDE RGAVVIMDYE RGAVIMTYD RGAVVIMTYD RGAVVIMTYD RGAVVIMTYK RGVOIAIMTFK RGVOIAIMTFK RGVOIAIMTFK RGVOIAIMTFK RGVQIAIMTFK RGVQIAIMTFK RGVQVIVMTYK RGVRISVMSYK RGRISVMSYK RGRISVMS	DALI EE  166 133 144 18 44 59 39 433 433 433 433 431 118 566 566 554 555 551 339
CDA Homsa 263657  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Dacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Homsa 24416443  APOBEC3G_Macni_48476259  APOBEC3G_Macni_48476259  APOBEC3G_Pantr_48476269  APOBEC3G_Lagla_48476319  APOBEC3G_Denpy_48476299  APOBEC3G_Ponpy_48476299  APOBEC3G_Ponpy_48476299  APOBEC3D_Homsa 29297041  APOBEC3D_Homsa 22907041  APOBEC3D_Homsa 22907041  APOBEC3D_Homsa 22907041  APOBEC3D_Homsa 22907041  APOBEC3D_Homsa 22907041  APOBEC3D_Dense_61651784  APOBEC2D_Darre_46487636  AID_Tarkfu_41016736  AID_Tarkfu_41016736  AID_Tarti_47221672  APOBEC2_Danre_61651784  APOBEC2_Danre_61651784  APOBEC2_Danre_61651784  APOBEC2_Danre_61651784	N	.0 a a 0 a a 1 a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKT FQMLAVAADTPGPV EEEEEQVTITCYLTWQVTITCYLTWQUTYTCYTSWQDYRVTCFTSWQDYRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQTRVTCFTSWQKFEITCYVTWTKYQVTWYTSWREFQITWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRUTWFCSWCKYTITWFCSWCKYTITWFSSW	TMPENKQI  ISPCGA SPCGA HHSPCPN SPCPN SPCFS SPCFS SPCFS SPCFS SPCFS SPCFB SPCPD SPCPD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYD SPCYN SPCXN	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CROFINEFVVKDFFIVMLN CRQVLVEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE GAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEWAEFLDENPHVNLHI CAQEVAEFLHENPHVNLHI CAQEVAEFLARHSNVNLTI CAQEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGEVAEFLAEHPNVTLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CAGCVAEFLARHSNVNLTI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLTLRI CAGQUAHFLSQTPNLRLRI CAQQLAHFLSQTPNLRLRI CSIQLCGFLNNTPNLRLRI CSTCLSGFLRNTPNLRLRI CATKLAEILRSRKNIRLAI	TYSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- SAARIYDYWBC- FAARIYDYWBC- FAARIYYFQYP- FAARIYFCEDR- FTARLYFCEDR- FTSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDME- FSSRLFEWEEP- LVSRLFFWEEP- LVSRLFFWEEP-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK QGQIKEMTVE EEEEE GGILVDVMDLP GGASVEIMGYK GGAKISIMTYS KGAKISIMTYS KGARISIMTYS KGARISIMTYS KGAPISMMKYS KGAPISMMYS KGAPISMMYS KGAPISMMYS KGAPISMMYS KGAPISMYS KGARVKIMDYE KGAQVINTYFK KGVQIAIMTFK KGVQIAIMTFK KGVQUISMTYS KGVQISWTYNY KGVQUISWTYK KGVQVISWTYK KGVQVISWTYK KGVQVISWTYK KGVGKLRMMKPL KGCKLRMMKPL KGCKLRMMKPL	DALII EF  166 133 144 18 44 59 39 433 433 433 433 433 433 566 566 555 555 311 342
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Crilo 48474310 APOBEC3F Homsa 24416443 APOBEC3G Macni 48476259 APOBEC3G Fantr 48476269 APOBEC3G Fantr 48476269 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Pantr 55661948 AROBEC3C Homsa 9294747 APOBEC3L Homsa 22907041 APOBEC3D Homsa 22907041 APOBEC3D Homsa 22297288 AID Homsa 22297288 AID Homsa 22297288 AID Jactu 408487636 AID Tathi 47026624 APOBEC3C Homsa 67039312 APOBEC3C Homsa 67039312 APOBEC3D Homsa 22297288 AID Homsa 22297288 AID Homsa 22297288 AID Homsa 22297288 AID Tathi 47046736 AID Tetni 47221672 APOBEC2 Danre 61651784 APOBEC2 Canfa 57094914	G N H K Y NN G G (11 G (11 G (11 G (11 G (11 A A A		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAWADTPGPV- EEEEEEQVTITCYLIWQVTITCYLIWQVTYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISWCTYRVTCFISWKCFQITWFVSWKCFQITWFVSWKYQVTWFTSWRYGVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWBUYSYTWFCSWBUYSYTWFCSWBUSYSTWFCSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPS SPCFS SPCPD SPCYD SPCYN SPCAN	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTAGGE- FAARIYDTAGGE- FAARIYDTAGGE- FAARIYDTAGGE- FAARIYDTAGGE- FAARIYTHWRO- FTARLYFFOED- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDE- FTARLYFCEDE- FVSRLYFCDE- FVSR		SILFEDSGYL HHHHH a1  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQQIKEMTY EEEEE GILVDVMDLP GGASVEIMGYK GGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYS KGAKISHMTYD KGARVKHMDDE KGARVKHMDDE KGARVKHMDDE KGAQVSIMTYD KGAQVISHMTYK KGVQIISHMTYK KGVGIISHMTYK KGVGIISHMTYK KGCKLRMMKPL KGCKLRIMKPQ KGCKLRIMKPQ	DALII EE  166 133 144 18 444 599 433 433 433 433 433 433 566 566 556 555 311 399 42 42
CDA Homsa 263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3 Crilo_48474310 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Sagla_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Homsa_9294747 APOBEC3C_Homsa_9294747 APOBEC3D_Homsa_9297041 AID_Canfa_50979250 AID_Homsa_22297288 AID_Galga_50729359 AID_Ictpu_40949661 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Takfu_41016736 AID_Teni_47221672 APOBEC2_Danre_61651784 APOBEC2_Canfa_57094914 APOBEC2_Canfa_57094914 APOBEC2_Canfa_57094914 APOBEC2_Canfa_57094914 APOBEC2_Canfa_57094914 APOBEC2_Calga_50760475	G H H NN A NN G NN G NN NN A NA		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLTWQVTYCFTSW QUTYCFTSW QUTYCFTSW QUTYCFTSW QTYVTCFTSWQTYVTCFTSWQTYVTCFTSWCYPKVTCFTSWCYPKVTCFTSWCKFQITWFVSWCKFQITWFVSWCKFTITYWFUSWREFQITWFUSWREFQITWFUSWREFQITWFUSWREFQITWFUSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWCKYTTWFTSWCKYTTWFCSWB) LSYSUTWFCSWCKYTTTWYMSSLKYNVTWYVSSLKYNVTWYVSS	TMPENKQI  ISPCGA SPCGVSPCGASPCPNSPCPSSPCFSSPCFSSPCFSSPCFSSPCFSSPCPDSPCPDSPCPDSPCYD	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYTOPL- FAARRIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FVSRLYFCDE-		SILFEDSGYL HHHHH a1  POGTYIVMTVQ PGSRSKVMTMG ROBEREIVELK AQQIKEMTVE EEEEE GILVDVMDLP ROBAVEIMGYK GGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGAKISIMTYS AGARISIMTYS AGAPISMMKYS AGARISIMTYD AGAPVKIMDDE ROGAVKIMDDE ROGAVKIMDE ROGAVKIMTYB ROGAVIMTYB ROVOIGHTFK ROGCKLRMKPQ ROGCKLRMKPQ ROCKLRIMKPQ ROCKLRIMKPQ ROCKLRIMKPQ ROCKLRIMKPQ ROCKLRIMKPQ ROCKLRIMKPQ ROCKLRIMKPQ	DALI EE  16 13 144 18 44 59 43 43 43 43 43 43 43 65 56 56 54 55 55 55 55 55 55 54 42 43
CDA Homsa 263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Dacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_CTilo_48474310 APOBEC3F_Homsa 24416443 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476269 APOBEC3G_Darlt_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Darlt_48476319 APOBEC3G_Darlt_48476399 APOBEC3G_Darlt_48476299 APOBEC3G_Darlt_57661948 AROBEC3C_Momsa 9294747 APOBEC3C_Homsa 9294747 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 22907041 APOBEC3D_Homsa 2297288 APOBEC4D_APOBEC6B_CTCAPCAPCAPCAPCAPCAPCAPCAPCAPCAPCAPCAPCAPC			EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTITCYLTWQVTYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWCTYRVTCFISWCTYRVTCFISWCKFEITCYVTWTKYQVTWYTSWRKFQITWFVSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRUTWFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCFR SPCFB SPCPD SPCYD SPCXN	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTARE FAARIYDTARE FAARIYDTARE FAARIYDTARE FAARIYTHWRO- FTARLYFFUEDR- FAARIYTYFUEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- FVSRLYF		SILFEDSGYL HHHHH al  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK AQQQIKEMTVE EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGARISIMTYS KGAQUSIMTYK KGAQUSIMTYK KGAQISIMTYK KGAQISIMTYK KGAQISIMTYK KGAQISIMTYK KGAQISIMTYK KGARISIMTYS KGARISIMTYK KGCKLRIMKPQ	DALII EE  16 13 144 18 44 59 43 43 43 43 43 43 56 56 56 55 55 55 55 53 42 42 43 39
CDA Homsa 263657 CDA Sacce 6323274 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Crilo 48474310 APOBEC3 Homsa 24416443 APOBEC3G Macni 48476259 APOBEC3G Pantr 48476269 APOBEC3G Pantr 48476269 APOBEC3G Pantr 55661948 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Lagla 78476319 APOBEC3G Lagla 78476311 APOBEC3G Lagla 78794741 APOBEC3G Lagla 78794941 APOBEC5G Lagla 50769475 APOBEC2 Lagla 50760475 APOBEC2 Canfa 57094914 APOBEC2 Tetni 47228640 APOBEC2 Tetni 47228640 APOBEC2 Tetni 47228640	G	.0 aa 1 aa 1 a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC: -FVAIAIASDSPDKT FQMLAVAADTPGPV EEEEEEQVTITCYLTWQVTRYCFTSW QDYRVTCFTSW QDYRVTCFTSW QDYRVTCFTSW QTRYCFISWTYPEVTFISWTKYQVTWFTSWCKFITCYVTWTKYQVTWFTSWREFQITWFVSWREFQITWFVSWREFQITWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSW	TMPENKQI  ISPCGA SPCGVSPCGA	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARRIYDDQGR- FAARRIYTQPD- FAARRIYTGEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTSRLYFCDE- FVSRLYFCDE- FVSRL		SILFEDSGYL HHHHH a1  POGTYIVMTVQ PGSRSKVMTMG ROBEREIVELK LQGQIKEMTVE EEEEE GILVDVMDLP RGASVEIMGYK RGARISIMTYS RGARISIMTYS RGARISIMTYS RGARISIMTYS RGAPISMMKYS LEAKISIMTYS RGAPISMMKYS LEAKISIMTYD RGAPVKIMDDE RGAPVKIMDDE RGAPVKIMDDE RGAPVKIMDDE RGAPVKIMDDE RGAPVKIMDDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMDE RGAPVKIMTYD RGARVKIMDE RGAPVKIMTYD RGARVKIMDE RGAPVKIMTYD RGARVKIMTYD RGARVKIMTYD RGARVKIMTYD RGVOISIMTYK RGVOISIMTYK RGVOISIMTYK RGVOISIMTYK RGVOISIMTYK RGVOISIMTYK RGVRISIMSYK RGVRISIMSYK RGCKLRMKPQ RGCKLRMKPQ RGCKLRMKPQ RGCKLRMKPQ RGCKLRMKPQ RGCKLRMKPA RGCCVRMMRRA RGCCVRMMRRA	DALI EE 16 13 144 18 44 18 49 39 43 43 43 43 43 45 56 56 55 55 55 31 39 24 24 33 39 53
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Criio 48474310 APOBEC3 Macni 48476259 APOBEC3 Gargo 50254066 APOBEC3G Pantr 48476269 APOBEC3G Pantr 48476309 APOBEC3G Pantr 55661948 AROBEC3G Lagla 48476309 APOBEC3G Sagla 48476309 APOBEC3G Sagla 7093121 APOBEC3G Sagla 7093121 APOBEC3G Homsa 9294747 APOBEC3D Homsa 9294747 APOBEC5D Homsa 9294744 APOBEC5D HOMSA 929474 APOBEC5D HOMSA 98474 APOBEC5D HOMSA 929474 APOBEC5D HOMSA 929474 APOBEC5D HOMSA 9	G		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLIWQVTITCYLTWQVTRYCFTSWQDYRVTCFTSWQDYRVTCFTSWQTYRVTCFISWQTYRVTCFISWQTYRVTCFISWQTYRVTCFISWQKFEITCYVTWKCFQITWFVSWCYRVTWFTSWCYRVTWFTSWRCYRUTWFTSWRCYRUTWFTSWRCYRUTWFTSWRCYRUTWFTSWRCYRUTWFTSW	TMPENKQI  ISPCGA SPCGA	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEWAEFILARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFISNNHWSLCI CAQEWAEFISNNHWSLCI CAQEWAEFISNNHWSLCI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGEVAEFLARHSNVNLTI CAKKLIAFVNDHPHISLRI CAGGVRAFLEGNPHVLSLRI CAGGVRAFLEGNPNLSLRI CARHVADFLEGNPNLSLRI CARHVADFLEGNPNLSLRI CARHVADFLEGNPNLSLRI CARHVADFLEGNPNLSLRI CARHVADFLEGNPNLSLRI CAGULHFLSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTPNLRLRI CAGGURAFLTSGTRNLRLI CAGRILLTSKTKNLRLI CADRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CANKLASILQCLKRNKTVRIQL CAASIAQCLRRNKTVRIQL CASAIAQCLRRNKTVRIQL CSASIAKFLDHYPNVTLAI	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYTGER- FAARIYFCDR- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- FVSRLFWEEP- LVSRLFWEEP- LVGRLFWEEP- AVARLFQWEEP- AVARLFQWEEP- FISKLYWHMOQ-		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGIKLMTVMLP GGSKSKVMTMG GGIKLMTVMLP GGAVEIMMYS GGAKISHMTYS AGAKISHMTYS AGAKISHMTYS AGAKISHMYS AGARVKIMDYE AGARVKIMDYE AGARVKIMDYE AGARVKIMDYE AGAVGIAHTFK AGVQIAHTFK AGVARIMKPQ AGCKLRMKPQ AGCCVYMMRGA AGCVYTQMRMRGA AGVTIQMSY	DALII EE  16 13 144 18 44 43 43 43 43 43 43 43 43 43 43 43 43
CDA Homsa 263657  CDA Sacce 6323274  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 Crilo 48474310  APOBEC3 Macni 48476239  APOBEC3G Macni 48476269  APOBEC3G Pantr 48476269  APOBEC3G Pantr 55661948  APOBEC3G Pantr 55661948  APOBEC3G Anfa 57093121  APOBEC3B Pantr 55661948  APOBEC3C Homsa 9294747  APOBEC3 Anfa 57093121  APOBEC3B Pantr 55661948  APOBEC3C Homsa 9294747  APOBEC3C Homsa 9294741  APOBEC3C Homsa 9294747  APOBEC4 Confa 50979250  APOBEC5 APOBEC6 Homsa 9294741  APOBEC5 Confa 57094661  APOBEC6 Confa 50979250  APOBEC6 Confa 57094914  APOBEC6 Confa 57094914  APOBEC6 Confa 57094914  APOBEC6 Confa 57094914  APOBEC6 Tetni 4722863039  APOBEC6 Mondo 23396444  APOBEC1 Musmu 13624299	G N N N N N N N G G G G G G G G G G G G S S S S S S		EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEQVTITCYLIWQVTITCYLTWQVTYCFTSW QDYRVTCFTSW QDYRVTCFTSW QTYRVTCFISWCYFRVTCFISWCYFRVTCFISWCYFRVTCFISWCKFEITCYVTWTKYQVTWYTSWRYPQVTWFTSWRYPQVTWFTSWRYPVTWFTSWRYPVTWFTSWRYPVTWFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA HH SPCPN SPCPS SPCFS SPCYD	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTARE FAARIYDTARE FAARIYDTARE FAARIYDTARE FAARIYFHWRO- FTARLYFFUEDE- FAARIYTCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- F		SILFEDSGYL HHHHH al  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK AQQQIKEMTVE EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGAKISIMTYS KGARISIMTYS KGARISIMTYK KGAQUSIMTYK KGAQUSIMTYK KGAQISIMTYK KGAQISIMTYK KGAQISIMTYK KGARISIMTYS KGARISIMTYK KGARISIMTY KGARI	DALI EE  166 134 148 44 599 399 433 433 433 433 433 556 566 554 555 531 399 422 433 598 82
CDA Homsa 263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema 4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3 Criio 48474310 APOBEC3G_Macni_48476259 APOBEC3G_Macni_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Lagla_48476319 APOBEC3G_Donyy_48476299 APOBEC3G_Donyy_48476299 APOBEC3C_Homsa_9294747 APOBEC3C_Homsa_9294747 APOBEC3C_Homsa_9294747 APOBEC3D_Homsa_22997041 APOBEC3D_Homsa_22997041 APOBEC3D_Homsa_22997041 APOBEC3D_Homsa_22997288 APOBEC3D_HOmsa_22997288 APOBEC3D_HOmsa_22997288 APOBEC4D_HOMSA_92958 APOBEC4D_APOBEC5C_APOBEC6C_APOBEC6C_APOBEC6C_APOBEC6C_APOBEC6C_APOBEC6C_Musmu_162286440 APOBEC5C_Ratno_27681627 APOBEC5C_Balga_50760475 APOBEC5C_Ratno_27681627 APOBEC5C_Ratno_27681627 APOBEC5C_Balga_50760475 APOBEC5C_Ratno_27681627 APOBEC5C_RAT			EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLTWQVTYCFTSW QUTYCFTSW QUTYCFTSW QUTYCFTSW QTYRVTCFTSW QTYRVTCFISWQTYRVTCFISWQTYRVTCFISWCKFICTYVTWKKFQITWFYSWKKFQITWFYSWKKFQITWFYSWREFQITWFYSWREFQITWFYSWREFQITWFYSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFSWALKYNTTWYWSSLKYNTTWYWSSLRYNTTWYSSLRYNTTWYSSLRYNTTWYSS	TMPENKQI  ISPCGA SPCGVSPCGA	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CROFINEFVVKDFFIVMLN CRQVLVEF-SDDLDVIMA- GRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHHH EEEE CAWQLAAFKRDRPDLILHI CAGEVAAFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNHVSLCI CARENTENDHPHISIRL CAGEVAEFLARHSNVNLTI CARENTAPHONHPHISIRL CARHVADFLEGYPNLSLRI CARHVADFLEGYPNLSLRI CARHVADFLEGYPNLSLRI CAQULAHFLSGTPNLIRLI CAQULAHFLSGTPNLIRLI CAQULAHFLSGTPNLIRLI CAGILGETLRINKHNLLLI CADRIBETLRINKHNLLLI CADRIBETLRINKHNLLLI CADRIBETLRINKHNLLLI CADRIBETLRINKHNLLLI CARSTANGENTARI CAASIAQCLERNKTVRIQL CAASIAQCLERNKTVRIQL CSRAITEFLSGHPNVTLFI CSRAITEFLS	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYFHWKR- FTARLYFHWKR- FTARLYYFWDT- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDTARD- FTARLYFGP- FTARLYFCEDR- FVSRLYFCDE- FVSRLYFCD		SILFEDSGYL HHHHH  a1  POGTYIVMTVQ POSSSKVMTMG ROBEREIVELK LQGQIKEMTVE EEEEE GILVDVMDLP ROGASVEIMGYK GOAKISIMTYS GOAKISIMTYS GOAKISIMTYS GOAKISIMTYS GOAVEIMTYE ROGAVEIMTYE ROGAVEIMTY ROGAVEIMTYE ROGAVEIMTY	DALI EE  166 131 144 18 44 599 399 433 433 433 433 433 433 433 433 4
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 F.Homsa 24416443 APOBEC3F.Homsa 24416443 APOBEC3G Macni 48476259 APOBEC3G Jeant 48476269 APOBEC3G Sagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Ponpy 48476299 APOBEC3G Homsa 9294747 APOBEC3G Homsa 9294747 APOBEC3D Homsa 22907041 APOBEC3D Homsa 22907041 APOBEC3G Jeant 755661364 APOBEC3G Jeant 750979250 AID Homsa 22297288 AID Musmu 67573018 AID Galga 50792950 AID Jeant 40949661 AID Tarki 41016736 AID Tarki 41016	G	0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  FRAIAIASDMQDDF: -WKCMVICGDSEDQC' FVAIAIASDSPDKT- FQMLAWADTPGPV- EEEEEEQVTITCYLIWQVTTCYLTWQVTTCYTSWQVTTCFTSWQVTYCFTSWQTYRVTCFTSWQTYRVTCFTSWCTYRVTCFISWCTYRVTCFISWCKFQITWFVSWKCFQITWFVSWKCFQITWFVSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPE SPCPS SPCFS SPCFS SPCFS SPCFS SPCFS SPCPT SPCPD SPCYD SPCX	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLLENPHVNLHI CAGEVAEFLLENPHVNLHI CACKLLAFVNDHPHISIRL CAGEVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLTIRI CARHVADFLEGYPNLTIRI CARHVADFLEGYPNLTIRI CARHVADFLEGYPNLTIRI CARHVADFLEGYPNLTIRI CAGURAFLEGNPHNLTIRI CAGURAFLEGNPHNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNLTIRI CAGURAFLEGYPNTLFI CAGNIKASILQORKWRLCI CAASVAQCLRRNKTVRIQL CASAIRKFLBHPYVTLFI CSKAIRFFLSGHPVVTLFI CSKAIRFFLSGHPVVTLFI CSKAIRFFLSGHPVVTLFI CSKAIRFFLSGHPVVTLFI CSCAIRFFLSGHPVVTLFI CSCAIRFFLSGHPVTLFI CSCAIRFFLSGHPVVTLFI CSCAIRFFLSGHPVTLFI CSCAIRFTLSGHP	TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FTARLYFCEDR- FVSRLYFCDLE- FVSRLYFCLE- FVS		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQQIKEMTY EEEEE GILVDVMDLP GGASVEIMGYK GGAKISHTYS AGAKISHTYS AGAFUKHMDYE AGAFUKHMDYE AGAFUKHMDYE AGAFUKHMDYE AGAFUKHMTYA AGYOITAHTFK AGYOITAHTHE AGYOITAHTFK AGYOITAHTH AGYOITAHTFK AGYOITAHTFK AGYOITAHTFK AGYOITAHTFK AGYOITAHTFK AGYOITAHTFK AGYOITAHTFK AGYOITAHTH AGYOITAHTH AGYOITAHTH AGYOITAHT AGYOITAH	DALI EE 16 13 1 14 18 4 4 5 9 3 9 3 9 3 4 3 4 3 4 3 4 3 4 3 4 3 5 5 6 6 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Crilo 48474313 APOBEC3 Macni 48476259 APOBEC3 Macni 48476259 APOBEC3 Fantr 48476269 APOBEC3 Fantr 48476269 APOBEC3 Fantr 55661948 APOBEC3 Canfa 57093121 APOBEC3 Canfa 57093121 APOBEC3 Canfa 57093121 APOBEC3 Macni 55661948 APOBEC3 Macni 5709329 APOBEC Canfa 57094914 APOBEC Canfa 57094914 APOBEC2 Tetni 47228640 APOBEC2 Xenla 49256526 APOBEC1 Macni 12002871	G N	0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEE	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPS SPCFS SPCFSWG SPCYD	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- GRQFINEFVVKDFFIVMLN GRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHHH EEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI GAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFLENDHVNLHI CAQEWAEFLENDHVNLHI CAQEWAEFLENDHVNLHI CAQEWAEFLENDHVNLISLAN CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CAGGVAFLGENTPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLTLRI CARHVADFLRGYPNLTLRI CARHVADFLRGYPNLTLRI CARHVADFLRGYPNLTLRI CARLAFLSGTENDHRILRI CAGGLENTHYNLRI CAGGLENTHYNLRI CAGGLENTHYNLRI CAGGLENTHYNLRI CAGRICAGE CASVAGCLERNKTVRIQL CAASVAGCLERNKTVRIQL CAASVAGCLERNKTVRIQL CASKAIREFLSGHPNVTLFI CSKAIREFLSGHPNVTLFI CSCAAITEFLSGHPNVTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFLSGHPNTLFI CSCAAITEFL	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDVRFG- FAARIYDWRPG- FAARIYDWRPG- FAARIYFOED- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE-		SILFEDSGYL HHHHH  a1  POGTYIVMTVQ GSSSKVMTMG KDGNFEIVKLK  QQQIKEMTVE EEEEE GILVDVMDLP CGASVEIMGYK GGAVEIMGYK GGAKISIMTYS GGAKISIMTYS GGAKISIMTYS GGAKISIMTYS GGAKISIMTYS GGARVKIMDDE GGAVUMSYL GGAVIMTYD GGAVIMTYD GGAVIMTYC GGAVIMTYC GGAVIMTYC GGAVIMTYC GGAVIMTYC GGAVIMTYC GGAVIMTYC GGCKLRMKPL GGCKLRIMKPQ GCCYTUMTEQ GGVTIQIMTEQ GGVTIQIMSS	DALIE 163 144 18 4 4 5 9 9 4 3 3 9 3 4 3 3 9 3 1 1 8 8 3 9 2 5 5 3 3 9 2 4 2 3 9 3 5 5 3 3 8 9 2 8 8 1 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Crilo 48474310 APOBEC3 Macni 48476259 APOBEC3 Gorgo 50254066 APOBEC3 Pantr 48476269 APOBEC3 Pantr 48476269 APOBEC3 Fantr 48476269 APOBEC3 Fantr 55661948 AROBEC3 Canfa 57093121 APOBEC3 Confa 57093121 APOBEC3 Macni 55661948 APOBEC3 Macni 55661948 APOBEC3 Ponpy 48476299 APOBEC3 Ponpy 48476299 APOBEC3 Ponpy 5661948 APOBEC3 Ponpy 48476299 APOBEC3 Ponpy 5661948 APOBEC3 Ponpy 5661948 APOBEC3 Ponpy 5661948 APOBEC3 Ponpy 5661948 APOBEC3 Ponpy 67661948 APOBEC3 Ponpy 67661949 APOBEC Canfa 57094914 APOBEC CANTA 49523039 APOBEC CANTA 49526526 APOBEC Musmu 13624299 APOBEC Musmu 13624299 APOBEC Musmu 13624299 APOBEC Musmu 13624299 APOBEC L Mondo 23396444 APOBEC Musmu 13624299 APOBEC Musmu 13624299 APOBEC Musmu 13624299 APOBEC Homsa 2696116	G N A N N N N S S S S S S S S S S S S S	0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVIGGDSEDQC: -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEEEQVTITCYLTWQVTYCFTSW QUTYCFTSW QUTYCFTSW QUTYCFTSW QTYRVTCFTSW QTYRVTCFISWQTYRVTCFISWQTYRVTCFISWCKFEITCYVTWKFQITWFVSWKFQITWFVSWKFQITWFYSWRFYQITWFTSWRFYQITWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFCSWA) LSYSTTWFCSWLKYNTTWYWSSLKYNTTWYWSSLRYNTTWYWSSLRYNTTWYSSLRYNTTWYSSLRYNTTWYSSSVTVTCYVSSSVTVTCYVSSVRCSITWFLSWTRCSITWFLSWTCCSITWFLSW	TMPENKQI  ISPCGA SPCGV	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- CROFINEFVVKDFFIVMLN CRQVLVEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEE CAWQLAAFKRDRPDLILHI CAGEVAAFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNHWSLCI CARENTHENDHPHISIRL CAGEVAEFLARHSNVNLTI CAGEVAEFLARHSNVNLTI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CAQCLAHFLSGTPNLRILI CAQCLAHFLSGTPNLRILI CAQCLAHFLSGTPNLRILI CAQCLAHFLSGTPNLRILI CAQCLAHFLSGTPNLRILI CADRISTIRNHWNLLIT CANKLASILQORKWRLCI CAASVAQCLRRNKTVRIQL CAASVAQCLRRNKTVRIQL CAASIAQCLRRNKTVRIQL CSRAITEFLSGHPNVTLFI CSGAIREFLSGHPGVTLVI CSGMAIREFLSGHPGVTLVI CSGMAIREFLSGHPGVTLVI CSGMAIREFLSGHPGVTLVI CSGMAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSGAIREFLSGHPGVTLVI CSGAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSGAIREFLSGHPGVTLVI CSGAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSGAIREFLSGHPGVTLVI	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYFHWKR- FTARLYFHWKR- FTARLYYFWDT- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDDQGR- FAARRIYDTALYFQP- FAARRIYDTALYFQP- FAARRIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FVSRLYFCDE- F		SILFEDSGYL HHHHH  a1  POGTYIVMTVQ POSSSKWTMG ROBEREIVELK LQGQIKEMTVE EEEEE GILVDVMDLP ROGAVEIMGYR ROGAVEIMGYR ROGAVEIMGYR ROAKISIMTYS ROAPISMMYS ROAPISMMS ROAPISMS ROYULIMTFK ROVOISMTY ROYULIMTFK ROVOISMTY ROYULIMTS ROYULIMTS ROCULIMMS ROCULIMMS ROCULIMMS ROCULIMS ROYULIMTS	DALIE 163 144 18 4 4 5 9 9 4 3 3 9 3 4 3 3 9 3 1 1 8 8 3 9 2 5 5 3 3 9 2 4 2 3 9 3 5 5 3 3 8 9 2 8 8 1 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9 9
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 F.Homsa 24416443 APOBEC3F.Homsa 24416443 APOBEC3G Macni 48476259 APOBEC3G Sorgo 50254066 APOBEC3G Pantr 48476269 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Lagla 48476319 APOBEC3G Ponpy 48476299 APOBEC3G Homsa 9294747 APOBEC3G Homsa 9294747 APOBEC3G Homsa 92907041 APOBEC3G Homsa 92907041 APOBEC3G Homsa 92907041 APOBEC3G Homsa 9291709 APOBEC3G Homsa 92907041 APOBEC4G Homsa 920907041 APOBEC5G Homsa 9209070475 APOBEC5G Musmu 13624299 APOBEC1 Mosma 13624299 APOBEC1 Homsa 92696116 APOBEC1 Homsa 92696116 APOBEC1 Ratno 6978519	G		EFYEAFGFPYGP HHHEE  FRAIAIASDMQDDF: -WKCMVICGDSEDQC' FVAIAIASDSPDKT- FQMLAWADTPGPV- EEEEEEQVTITCYLIWQVTITCYLTWQVTITCYTSWQVTYCFTSWQVTYCFTSWQYRVTCFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPB SPCPE SPCPS SPCFS SPCFS SPCFS SPCPT SPCPD SPCYD SPCXB	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLSELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CACKLLAFVADHPHISIRL CAGEVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CAGVAEFLENPHVNLHI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLSLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CARHVADFLEGYPNLTLEI CAGNURAFLSGTPNLRLEI CAGDRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CADRILKTLSKTKNLRLI CANKLASILQORKWRLCI CAASVAQCLRRNKTVRIQL CAASIAGCLKRNKTVRIQL CASAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSMAIREFLSGHPGVTLVI CSGAIREFLSRYPHVTLFI	TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FTARLYFGTQF- FAARIYFGTQF- FAARIYFCEDR- FTARLYFCEDR- FUSRLYFCDE- FUSRLYFWEEP- MUGRLFWWEEP- MUGRLFWHDQ- TYARLYHHTDQ- TYARLYHHTDQ- TYARLFWHMDQ-		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG ROGREEIVELK LQGQIKEMTE EEEEE GILVDVMDLP GGASVEIMGYK GGARISMTYS GGARISMTYS GGARISMTYS GGARISMTYS LGARKISMTYS LGARKISMTYS LGARKISMTYS LGARKISMTYS LGARKISMTYS LGARKISMTYD LGARVKIMDDE LGARVKIMDDE LGARVKIMDDE LGARVKIMDDE LGARVKIMDTE LGARVKIMTYK LGAQUSIMTFK LGAQUSIMTFK LGAQUSIMTFK LGACURIMTRA LGAVRISMTY LGCKLRIMKPQ LGCKLRIMKPS LGCYTIQIMTEQ LG	DALIE 163343334433333334423114855555555555555555555555555555555555
CDA Homsa 263657  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Bacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Flomsa 24416443  APOBEC3G_Macni_48476259  APOBEC3G_Macni_48476259  APOBEC3G_Andri_48476269  APOBEC3G_Sagla_48476319  APOBEC3G_Ponpy_48476299  APOBEC3G_Ponpy_48476299  APOBEC3G_Homsa 9294747  APOBEC4G_GAGG_HOMSA 929474  APOBEC4G_GAGG_HOMSA 929474  APOBEC5G_ATA_TOMSA 929474  APOBEC5G_ATA_TOMSA 949414  APOBEC6G_ATA_TOMSA 949414  APOBEC6	G	.0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC' -FVAIAIASDSPDKTFQMLAVAADTPGPV- EEEEE	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPP SPCFS SPCFS SPCFS SPCFS SPCFS SPCFS SPCPD SPCPD SPCYD	LIFYEVKDFSGTNIQK EEEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- GRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHHH EEEE CAWQLAAFKRORPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEWSLCI CAQEMAKFISNNEWSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNCHVSLCI CACHUADFIRGNPNLSLRI CAGGVAFFLACHNTPNLSLRI CAGGVAFFLACHNTPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLSLRI CARHVADFIRGNPNLTLRI CAGGLAFFLSGTENNTPNLRIRI CATKLAEILCSTOLLSGTENNTPNLRIRI CATKLAEILCSTOLLSGTENNTPNLRIRI CADRIITTICKTKNLRLLI CADRIITTICKTKNLRLLI CADRIITTICKTKNLRLLI CANSKAJCLRNKTVRIQL CAASVAQCLRNKTVRIQL CASASIAQCLRNKTVRIQL CASASIACLERNKTVRIQL CASAIREFLSGHPBVTLUT CSQAIREFLSGHPBVTLUT CSGAIREFLSGHPBVTLUT CSGAIREFLSGHPGVTLUT CSGAIREFLSGHPGVTLUT CSGAIREFLSGHPGVTLUT CSGAIREFLSGHPGVTLUT CSGAIREFLSRYPHVTLFI CISKMTTFLAKYPDITLSV	TYTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDPL- FAARIYDPL- FAARIYTPGEDR- FTARLYYFCEDR- FTARLYFCEDR- FYSRLYFCOME- FYSRLYFCOME- FYSRLYFCOME- FYSRLYFCOME- TYSRLYFWEEP- LVGRLFWWEEP- LVGRLFWWEEP- KORLFWEEP- AVARLFQWEEP- AVARLFQWEEP- JIARLYHHTDQ- YARLYHHTDQ- YVARLFWHMDQ- FVARLFWHMDQ- YVARLFWHMDQ- YYARLYHHTDG- YYARLFWHMDQ- YYARLYHHTDG- FYSGLYHTEAEF		SILFEDSGYL HHHHH  a1  POGTYIVMTVQ PGSRSKVMTMG ROGRIFEIVKLK LQGQIKEMTVE EEEEE GILVDVMDLP ROGRIFEIVKLK LQGQIKEMTVE ROGRIFSMKYS LGAKISIMTYS LGAKISIMTYS LGAKISIMTYS LGAKISIMTYS LGAKISIMTYS LGARISIMTYS LGARISIMTYS LGARISIMTYS LGARISIMTYS LGARISIMTYS LGARISIMTYS LGARISIMTYD LGARVKIMDDE LGARVKIMDDE LGARVKIMDTE LGARVKIMDTE LGARVKIMDYE LGARVINTYK LGAQUISIMTYK LGAQUISIMTYK LGAVQISIMTYK LGAVQISIMTYK LGAVQISIMTYK LGAVQISIMTYK LGAVGISIMTKE LGAVGISIMTKE LGAVGISIMTKE LGAVGISIMTKE LGAVGISIMTKE LGAVGISIMTKE LGAVGISIMTYK LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY LGAVGISIMTY	DALIE 163 144 18 44 43 143 43 143 43 143 143 143 143 143
CDA Homsa 263657  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Sacce_6323274  CDA_Dacsu_80258  APOBEC3_Musmu_26340722  APOBEC3_Flomsa 24416443  APOBEC3G_Macni_48476259  APOBEC3G_Macni_48476259  APOBEC3G_Macni_48476269  APOBEC3G_Lagla_48476319  APOBEC3G_Lagla_48476319  APOBEC3G_Denpy_48476299  APOBEC3G_Denpy_48476299  APOBEC3B_Pantr_55661948  APOBEC3B_Pantr_55661948  APOBEC3B_Homsa 22997041  APOBEC3B_Homsa 22997041  APOBEC3B_Homsa 22997041  APOBEC3B_Homsa 22997041  APOBEC3B_Homsa 22997041  APOBEC3B_HOMSa 22997288  APOBEC3B_HOMSa 22997288  APOBEC4B_HOMSa 22997288  APOBEC5B_HOMSa 27681627  APOBEC5C_RATHO_27681627  APOBEC6C_MOMO_23396444  APOBEC6C_MOMO_23396444  APOBEC6C_MOMO_23396444  APOBEC6C_MOMO_23396444  APOBEC6C_MOMSA 23396444  APOBEC6C_	G	.0 a a a a a a a a a a a a a a a a a a a	EFYEAFGFPYGP HHHEE  FRAIAIASDMQDDF: -WKCMVICGDSEDQC' FVAIAIASDSPDKT- FQMIAVAADTPGPV- EEEEEEQVTITCYLIW QVTTCYLIW QVTTCYLTW QVTYCFTSW QVTYCFTSW QVTYCFTSW QYRVTCFTSW QYRVTCFISW QYRVTCFISW QYRVTCFISW QYRVTCFISW QYRVTCFISW QYRVTCFISW QYRVTCFISW QYRVTCFISW QKFEITCYVTW TKYQVTWYTSW QYRVTWFTSW RCYRVTWFTSW RCYRVTWFTSW RCYRVTWFTSW RCYRVTWFTSW RCYRVTWFTSW RCYRVTWFTSW A) LSYSITWFCSW a) LSYSITWFCSW a) LSYSITWFCSW a) LSYSITWFCSW LKYNVTWYVSS LKYNVTWYVSS LKYNVTYVSS LYNVTYVSS SVTVTCYVSS SVTVTCYVSS SVTVTCYVSS SVTVTCYVSS SVTVTCYVSS SVTVTCYVSS TRCSITWFLSW TRCSITW	TMPENKQI  ISPCGA SPCGA	LIFYEVRDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWPVYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVLYEF-SDDLDVIMA- CRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKRDRPDLILHI CAGEWAEFLARHSNVNLTI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGEVAEFLENPHVNLHI CAGGVAEFLENPHVLITI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLTLRI CAGNIRFTSSTRNRKNIRLAI CARHVADFLRGYPNLTLRI CAGNIRFTSSTRNRKNIRLAI CARHVADFLRGYPNLTLRI CSQALRFKNRTVRIQL CAASIACCLRRNKTVRIQL CSKAITEFLSGHPNVTLFI CSQAIREFLSGHPGVTLVI CSGAIRFLSGHPGVTLVI CSGAIRFLSGHPWTLFI CSGAIRFLSGHPGVTLVI CSGAIRFLSGHPGVTLVI CSGAIRFLSGHPWTLFI CSGAIRFLSGHPGVTLVI CSGAIRFLSGHPWTLVI CSGAIRFLSGHPGVTLVI CSGAIRFLSGHPWTLVI CISKMYTFLAKYPDITLSV CISKMYNFLMNYPEVUTLSV CISKMYNFLMNYPEVUTLSV	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYWBR- SAARLYYWBR- FASRLYFHWRQ- FTARLYFCEDR- FAARIYTODEL- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- FVSRLYFHEEP- LVGRLFWWEEP- AVARLFQWEEP- FISRLYWHMDQ- YVARLFWHMDQ- YVARLFWHMDQ- YVARLFWHMDQ- YVARLFWHMQ- YFSQLYHTEAE- FFFSQLYHTENQF		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGIKLDVMDLP GGASVEIMGYK AGAKISHMTYS AGAKISHMTYS AGAKISHMTYS AGAKISHMYS AGARVISHMYD AGARVKHMDYE AGAVUSHMYD AGARVKHMDYE AGVQIALMTFK AGVALMTFK AGVALM	DALLI 161 163 144 18 443 433 433 433 433 433 433 1188 556 555 555 555 555 555 555 555 555 5
CDA Homsa 263657  CDA Sacce 6323274  CDA Sacce 6323274  CDA Thema 4981379  CDA Bacsu 80258  APOBEC3 Musmu 26340722  APOBEC3 Musmu 26340722  APOBEC3 Macni 48474310  APOBEC3 Macni 48476343  APOBEC3 Macni 48476259  APOBEC3 Fantr 48476269  APOBEC3 Fantr 55661948  APOBEC3 Sagla 48476309  APOBEC3 Anfa 5703121  APOBEC3 Anfa 5703121  APOBEC3 Anfa 5703121  APOBEC3 Macni 55661948  APOBEC3 Macni 55661948  APOBEC3 Macni 55661948  APOBEC3 Macni 55661948  APOBEC3 Macni 5703121  APOBEC3 Macni 5703121  APOBEC3 Macni 50703121  APOBEC4 Macni 5070312  APOBEC5 Macni 5070312  APOBEC5 Macni 5070318  APOBEC6 Macni 5070318  APOBEC6 Macni 40703661  APOBEC6 Macni 40703661  APOBEC6 Macni 40703641  APOBEC6 Macni 40703641  APOBEC6 Macni 407036444  APOBEC7 Macni 407036444  APOBEC6 Macni 407036444  APOBEC6 Macni 407036444  APOBEC6 Macni 12002871  APOBEC6 Macni 26978519  APOBEC6 Macni Andro 207878519  APOBEC6 Macni Andro 44708E640  APOBEC6 Macni 447038641  APOBEC6 Macni 40708611  APOBEC6 Macni 407081  APOBEC6 Macni 407081  APOBEC6 Macni 407081  APOBEC6 Macni 407081  APOBEC6 Mac	G N S S N N N N S N	0 a a 1 a a	EFYEAFGFPYGP HHHEE  FRAIAIASDMQDDF: WKCMVICGDSEDQC: FVAIAIASDSPDKT FQMLAWADTPGPV EEEEEE QVTITCYLIW QVTITCYLIW QVTITCYLTW QVTYCFTSW QDYRVTCFTSW QDYRVTCFTSW QTYRVTCFISW CTYRVTCFISW CTYRVTCFISW CTYRVTCFISW CYFRVTCFISW KCFQITWFVSW KCFQITWFVSW CYRVTWFTSW CYRVTWFTSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPA SPCPS SPCFS SPCFS SPCFS TPCPD SPCPD SPCPD SPCYD SPCW	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  CRQVMREFGTNWP-VYMT- CRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- CAMPLANFKDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEWARFISNNKHVSLCI CAQEWAEFISNKHVSLCI CAQEWAEFISNKHVSLCI CAQEWAEFISNNKHVSLCI CAQEWAEFISNNCHVSLCI CAQEWAEFILAHNVNLTI CAQEVAEFLARHSNVNLTI CAGQUAFILAFVNDHPHISIRL CAGQUAFILAFVNDHPHISIRL CAGQUAFILAFVNDHPHISIRL CAGQUAFILAFVNDHPHISIRL CARHVADFILGYPNLSLRI CARHVADFILGYPNLSLRI CARHVADFILGYPNLSLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARHVADFILGYPNLTLRI CARLABILGSTKNIRLLI CARLABILGSTKNIRLLI CARLABILGSTKNIRLLI CADRILKTLSKTKNIRLLI CADRILKTLSKTKNIRLLI CANKLASILQORKWRTUCI CAASVAQCLRRNKTVRIQL CAASVAQCLRRNKTVRIQL CAASVAQCLRRNKTVRIQL CASAIREFLSQHPGVTLUI CSKAIRKFLSQHPGVTLUI CSKAIRKFLSQHPGVTLUI CSCAIREFLSGHPGVTLUI CSCAIREFLSGHPGVTLUI CSCAIREFLSGHPGVTLUI CSCAIREFLSGHPGVTLUI CSCAIREFLSGHPGVTLUI CSCAIREFLSRYPHVTLFI CSCAIREFLSGHPGVTLUI CSCAIREFLSRYPHVTLFI CSCAIREFLSRYPHVTLFI CSCAIREFLSRYPHVTLFI CSCAIREFLSRYPHVTLFI CISKMYTFLAKYPOITLSV CISKMYNFLMYPEUTLSV CISKMYNFLMYPEUTLSV CISKMYNFLMYPEUTLSV CISKMYNFLMYPEUTLSV CISKMYNFLMYPEUTLSV	TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQDPL- SAARLYYFGPR- FAARLYFGER- FAARLYFCER- FTARLYFCER- FTSRLYFCER- FVSRLYFCER- FVSRLYFCER- FVSRLYFCER- FVSRLYFCER- FVSRLYFCER- FVSRLYFCER- FVSRLYFCER- FTSRLYFWHEP- LVGRLFWWEEP- LVGRLFWWEEP- AVARLFQWEEP- AVARLFQWEEP- AVARLFQWEEP- TYARLFWHDQ- TYARLYHTDQ- TYARLFWHDQ- TYARLFWHDQ- TYARLFWHDQ- FVARLFWHDQ- FVARLFWHDQ- FVARLFWHDQ- FYSGLYHTEADF TFSGLYHTEADF		SILFEDSGYL HHHHH al  PDGTYIVMTVQ PGSRSKVMTMG ROGREFEIVELK LQGQIKEMTE EEEEE GILVDVMDLP GGASVEIMGYK GGARIKAMMYS GGARIKAMMYS KGARIKIMTYS KGYOLIAMTYK KGVQISHTYK KGVARIKIMKPQ KGCKLRMMKPS KGVTIQIMTS KGVTILIMTS KGVTIQIMTS KGVTIQ	DALLI 163 144 153 143 433 433 433 433 433 1188 433 1278 1566 1565 1565 1575 1571 1571 1571 1571
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Thema_4981379 CDA_Bacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Crilo_48474310 APOBEC3_Homsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Pantr_48476269 APOBEC3G_Sagla_48476319 APOBEC3G_Sagla_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Homsa_9294747 APOBEC3D_Homsa_2297041 APOBEC3D_Homsa_2297041 APOBEC3D_Homsa_2297041 APOBEC3D_Homsa_2297084 APOBEC3D_HOMSa_2297288 ATD_Musmu_6753018 ATD_Musmu_6753018 ATD_Musmu_6753018 ATD_Galga_50729359 ATD_Ictpu_40949661 ATD_Takfu_41016736 ATD_	G	0 aa 1 aa 1 a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC: -FVAIAIASDSPDKT FQMLAVAADTPGPV EEEEEQVTITCYLIWQVTYCFTSW QDYRVTCFTSW QDYRVTCFTSW QDYRVTCFTSW QTRYCFTSWQTRYCFTSWQTRYCFTSWQTRYCFTSWQKFEITCYVTW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPN SPCPS SPCFS SPCFS SPCFS SPCFS SPCFS SPCPD SPCPD SPCYD	LIFYEVKDFSGTNIQK EEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYFHWKR- FTARLYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYBE- FAARIYDYBE- FASRLYFHWRQ- FTARLYFCDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCDE- FVSRLYFCDE- FVSRLYFTDD- TVSRLYHHTDQ- YAARLYHHTDQ- YAARLYHHTDQ- YAARLYHHTDQ- YYARLYHHTDQ- YYARLYHHTDQ- YYARLYHHTDQ- YYFSQLYHTEMDF YFSQLYHTEMDF YFSQLYHTEMDF		SILFEDSGYL HHHHH al  POGTYIVMTVQ PGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKVMTMG ROGSRSKSIMTYS ROGAVEIMGYK ROGAVEIMGYK ROGAVEIMGYK ROGAVEIMTYS ROGAVEIMTYS ROGAVEIMTYS ROGAVEIMTYK ROGOLIMTFK ROGVUTVMTYK ROGULIMTFK ROGULIMTFC ROGULIMTEQ ROGULIMTEQ ROGULIMTEQ ROGULIMTEQ ROGULIMTEQ ROGULIMTEQ ROGULIMTS R	DALLE 163 144 18 44 459 433 433 433 433 1143 432 185 566 555 555 313 422 433 533 892 892 175 196 197 197 197 197 197 197 197 197 197 197
CDA_Homsa_263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Dacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Fomsa_24416443 APOBEC3_Macni_48476259 APOBEC3_Fomsa_24416443 APOBEC3G_Macni_48476259 APOBEC3G_Pantr_48476269 APOBEC3G_Lagla_48476319 APOBEC3G_Lagla_48476319 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3G_Ponpy_48476299 APOBEC3D_Homsa_22907041 APOBEC3D_Bomsa_22907041 APOBEC3D_Bomsa_22907041 APOBEC3D_Bomsa_22907041 APOBEC3D_Bomsa_22907041 APOBEC3D_Bomsa_22907041 APOBEC4_Bomsa_2290820 APOBEC4_Entil_47228640 APOBEC5C_Salva_5056264 APOBEC1_Mondo_23396444 APOBEC6_Tentil_47228640 APOBEC6_Tentil_47228640 APOBEC6_Musmu_13624299 APOBEC6_Tentil_47228640 APOBEC6_Tentil_4728640 APOBEC6_Tentil_472	G	.0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  FRAIAIASDMQDDF:	TMPENKQI  ISPCGA SPCGA SPCGA HH SPCPN SPCPE SPCFS SPCYD	LIFYEVRDFSGTNIQK EEEEEEE EEEE	TYTSRLYFHWKR- YTSRLYFHWKR- YTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQGR- FAARIYDTQCR- FAARIYTCEDR- FAARIYFCEDR- FAARIYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FVSRLYFCDE- FYSRLYFCDE- TYSRLYFWWEP- LVGRLFWWEP- LVGRLFWWEP- AVARLFQWEP- AVARLFQWEP- YIARLYHHTDQ- YYARLFWHMQQ- YYARLFQHMCQ- YYARLFWHMQQ- YYARLFWHMQQ- YYARLFWHMQQ- YYARLFWHMCQ- YYARLFWHCG- YYARLFWHMCQ- YYARLFWHMCQ- YYARLFWHMCQ- YYARLFWHMCQ- YYARLFWHCG- YYARLFWHMCQ- YYARLFWHCG- YYARLFWHMCQ- YYARLFWHMCQ- YYARLFWHCG- YYARLFWHMCQ- YYARLFWH		SILFEDSGYL HHHHH al  PDGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGSRSKVMTMG KDGNFEIVKLK AQGIKEMTY GGIKLDVMDLP GGASVEIMGYK GGAKIAMMYS KGAKISIMTYS KGAKISIMTYS KGARISIMTYS KGARISIMTYS KGARVKIMDYE KGARVKIMDYE KGARVKIMDYE KGARVKIMDYE KGARVKIMDYE KGARVKIMDYE KGAQUAINTFK KGVQIAINTFK KGVQVTUNTYK KGVARINKPQ KGCKLRMMKPA KGCKLRMMKPA KGCKLRMMKPA KGCVUTIQIMTA KGVTIQIMTA KGVTI	DALLE 163 144 18 449 339 433 433 4433 433 118 566 564 555 555 555 555 555 555 555 555
CDA Homsa 263657 CDA Sacce 6323274 CDA Thema 4981379 CDA Bacsu 80258  APOBEC3 Musmu 26340722 APOBEC3 Crilo 48474310 APOBEC3 Homsa 24416443 APOBEC3 Macni 48476259 APOBEC3G Macni 48476269 APOBEC3G Pantr 48476269 APOBEC3G Pantr 55661364 APOBEC3D Homsa 29297041 APOBEC3C LAG 5709325 APOBEC CANADA SA APOBEC APOBE	G N	0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC: -FVAIAIASDSPDKT FQMLAWADTPGPV EEEEEEQVTITCYLIWQVTITCYLIWQVTITCYLTWQVTYCFTSW QDYRVTCFTSW QDYRVTCFTSW QTYRVTCFISWCYTRVTCFISWCYTRVTCFISWCYTRVTCFISWCYTRVTCFISWCKFEITCYVTWTKYQVTWYTSWQKFEITCYVTWTKYQVTWYTSWQKFEITCYVTWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWLKYNVTWYVSSLKYNVTWYVSSLKYNTTWYVSSLEYDITWYVSSLEYDITWYVSSTRCSITWFLSWTRCSITWFLSWTCCSITWFLSWTCCSITWFLSWTCCSITWFLSWTCCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA HH SPCPN SPCPE SPCFS SPCYD SPCYN SPCXB	LIFYEVKDFSGTNIQK EEEEEEE EEEE b1 b2  GRQVMREFGTNWP-VYMT- GRQFINEFVVKDFFIVMLN CRQVLYEF-SDDLDVIMA- GRQVISELCTKDVIVVLT- HHHHHHHH EEEEE CAWQLAAFKEDRPDLILHI CAGEVAEFLARHSNVNLTI CAQEMAKFISNNEHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEMAKFISNNKHVSLCI CAQEWAFFISNNKHVSLCI CAQEWAFFISNNKHVSLCI CAQEWAFFISNNKHVSLCI CAQEWAFFISNNCHVNLLTI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAQEVAEFLENPHVNLHI CAGGVAFFLARHSNVNLTI CARHVADFLRGYPNLSLRI CARHVADFLRGYPNLTLRI CARHVADFLRGYPNLTLRI CARHVADFLRGYPNLTLRI CAGNIELKTLSKTKNLRLI CARLABILKSTKNLRLI CADRILKTLSKTKNLRLI CANKLASILQORKKVRLCI CANKLASILQORKKVRLCI CANKLASILQORKKVRLCI CANKLASILQORKKVRLCI CARSAIGCLRRNKTVRIQL CASAIRFFLSOHPOGVTLVI CSGAIRFFLSOHPOGVTLVI CSCAIRF	TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYYFWDT- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDYRG- FAARIYDYRG- FAARIYDYRG- FAARIYDYRG- FAARIYDYPB- FAARIYDYPB- FAARIYDYPB- FAARIYFOUP- FAARIYFOUP- FAARIYFOUP- FAARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDE- FVSRLYFCDE- TVSRLYFDDE- TVSRLYFWHED- TYARLYHHTDQ- TYARLYHHTDQ- TYARLYHHTDQ- TYARLYHHDQ- TYARLYHHDQ- TYARLYHHDQ- FVSRLYHTEMF TYFSQLYHTEMF TYFSQLYHTEMF TYFSQLYHTEMF TYFSQLYHTEMF TYFSQLYHTEMF		SILFEDSGYL HHHHH al  POGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK  QQQIKEMTVE EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS GGAVEIMGYK GGAKISIMTYS GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDYE GGARVKIMDYE GGAQVISIMTYD GGARVKIMDYE GGAQVISIMTYB GGAQUSIMTYB GGAQUSIMTYB GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVGIGIMTFK GGVTIQIMTS	DALLE 163 144 18 449 143 144 14 14 14 14 14 14 14 14 14 14 14 14
CDA Homsa 263657 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Sacce_6323274 CDA_Dacsu_80258  APOBEC3_Musmu_26340722 APOBEC3_Crilo 48474310 APOBEC3_Momi_48476259 APOBEC3_Gorgo_50254066 APOBEC3_Pantr_48476269 APOBEC3_Barl 48476319 APOBEC3_Lagla_48476319 APOBEC3_Dantr_55661948 AROBEC3_Dantr_55661948 AROBEC3_Dantr_55661948 AROBEC3_Dantr_55661364 APOBEC3_Dantr_55661364 APOBEC3_Dantr_55661364 APOBEC3_Dantr_55661364 APOBEC3_Dantr_55661364 APOBEC3_Dantr_646863 APOBEC3_Dantr_65636364 APOBEC3_Dantr_64686364 APOBEC3_Dantr_64686364 APOBEC3_Dantr_64686364 APOBEC3_Dantr_646876364 APOBEC3_Dantr_646876364 APOBEC3_Dantr_646876364 APOBEC3_Cantr_47228640 APOBEC3_Cantr_646876364 APOBEC3_Cantr_646876364 APOBEC3_Cantr_64786464 APOBEC3_Cantr_646876364 APOBEC4_Cantr_646876364 APOBEC4_Cantr_646876364 APOBEC5_Cantr_646876364 APOBEC5_Cantr_646876364 APOBEC6_Cantr_646876364 APOBEC6_Cantr_6678519 APOBEC6_Cantr_646876364 APOBEC6_Ratno_6978519 APOBEC6_Homsa_APOBEC6_Momsa_APOBEC6_Homsa_APOBEC6_Homsa_APOBEC6_Homsa_APOBEC6_Homsa_AP	G N	0 a a 0 a a 1 a a	EFYEAFGFPYGP HHHEE  -FRAIAIASDMQDDF: -WKCMVICGDSEDQC: -FVAIAIASDSPDKT FQMLAWADTPGPV EEEEEEQVTITCYLIWQVTITCYLIWQVTITCYLTWQVTYCFTSW QDYRVTCFTSW QDYRVTCFTSW QTYRVTCFISWCYTRVTCFISWCYTRVTCFISWCYTRVTCFISWCYTRVTCFISWCKFEITCYVTWTKYQVTWYTSWQKFEITCYVTWTKYQVTWYTSWQKFEITCYVTWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWRCYRVTWFTSWLKYNVTWYVSSLKYNVTWYVSSLKYNTTWYVSSLEYDITWYVSSLEYDITWYVSSTRCSITWFLSWTRCSITWFLSWTCCSITWFLSWTCCSITWFLSWTCCSITWFLSWTCCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSWTRCSITWFLSW	TMPENKQI  ISPCGA SPCGA SPCGA SPCGA SPCPN SPCPP SPCPS SPCFS SPCFS SPCFS SPCFS SPCFS SPCPD SPCPD SPCYD	LIFYEVRDFSGTNIQK EEEEEEE EEEE	TYTSRLYFHWKR- TYTSRLYFHWKR- FTARLYFWTD- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDDQGR- FAARIYDTARIYFGER- FASRLYFHWRQ- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FTARLYFCEDR- FYSRLYFCDE- FVSRLYFCDE- FVSRLYFTDD- FYSRLYFTENDF FFSQLYHTENDF		SILFEDSGYL HHHHH al  POGTYIVMTVQ GSRSKVMTMG KDGNFEIVKLK  QQQIKEMTVE EEEEE GILVDVMDLP GGASVEIMGYK GGAKISIMTYS GGAVEIMGYK GGAKISIMTYS GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDDE GGARVKIMDYE GGARVKIMDYE GGAQVISIMTYD GGARVKIMDYE GGAQVISIMTYB GGAQUSIMTYB GGAQUSIMTYB GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVQIGIMTFK GGVGIGIMTFK GGVTIQIMTS	DALLE 163 144 18 44 53 93 433 433 433 433 433 433 433 433 433

Figure 1 (previous page). Multiple alignment of the CDA/AID/APOBEC superfamily. Only the deaminase domain that is conserved in all proteins is shown. The proteins are designated by the corresponding species abbreviation appended by GI numbers. The secondary structure of the CDA proteins derived from the available crystal structures<sup>27</sup> is shown under the CDA sequences. The α-helices are denoted by "H"s and marked a1 and a2, and the β-strands are denoted by "E"s and marked b1-b5. The predicted secondary structure of the APOBEC4 subfamily is plotted below the alignment. Columns with 100% conserved residues are shaded gray. The species abbreviations are as follows: Bacsu, Bacillus subtilis; Bosta, Bos taurus; Canfa, Canis familiaris; Crilo, Cricetulus longicaudatus; Danre, Danio rerio; Galga, Gallus gallus; Gorgo, Gorilla gorilla; Homsa, Homo sapiens; Ictpu; Ictalurus punctatus; Lagla, Lagothrix lagotricha; Macfa, Macaca fascicularis; Macni, Macaca nigra; Mesau, Mesocricetus auratus; Mondo, Monodelphis domestica; Musmu, Mus musculus; Orycu, Oryctolagus cuniculus; Pantr, Pan troglodytes; Ponpy, Pongo pygmaeus; Ratno, Rattus norvegicus; Sacce, Saccharomyces cerevisiae; Sagla, Saguinus labiatus; Takru, Takifugu rubripes; Tetni, Tetraodon nigroviridis; Thema, Thermotoga maritima; Xenla, Xenopus laevis; Xentr, Xenopus tropicalis.

this subfamily.<sup>8</sup> The problems with the APOBEC3 subfamily are likely to be caused by long-branch attraction artifacts (long branches tend to corrupt the phylogenetic signal)<sup>28</sup> given the dramatic differences in the evolutionary rates between the fast-evolving APOBEC3 and other subfamilies (Table 2). We showed that APOBEC4 evolved at an intermediate rate, much lower than that of APOBEC3, similar to that of APOBEC1, but considerably greater than APOBEC2 and AID (Table 2).

To minimize long-branch attraction artifacts in tree reconstruction, we removed the fast-evolving APOBEC3 subfamily (see Table 2) from the present analysis. In the reconstructed tree, which can be rooted with the CDAs, APOBEC1 and APOBEC4 form distinct clades that are joined in a moderately supported cluster (Fig. 2). The APOBEC2 and AID subfamilies form a third clade with a high bootstrap support (Fig. 2). Interestingly, APOBEC4 is present in mammals, birds and amphibia (but so far not fishes) similarly to AID and APOBEC2 but unlike APOBEC1 which is so far restricted to mammals (Fig. 2 and Table 1). This suggests that the duplication leading to the distinct APOBEC1 and APOBEC4 genes might have occurred prior to the amphibia-reptile divergence, perhaps, with subsequent loss of APOBEC1 in some lineages. This interpretation contradicts the hypothesis that APOBEC1 is a mammalianspecific derivative of AID.8 However, phylogenetic reconstructions for families with a high rate variation among subfamilies (Table 2)

should be interpreted with utmost caution due to the substantial impact of long-branch attraction artifacts on tree topology.<sup>28</sup> In particular, it cannot be strictly ruled out that the APOBEC1-APOBEC4 clade is a result of such an artifact.

The deep internal branches of the tree, particularly, the branch connecting CDA and AID/APOBEC families, are very long compared to the branches within the CDA clade and within each clade of the AID/APOBEC family (Fig. 2). This suggests that the ancestor of AID/APOBEC family had an accelerated rate of evolution which still can be observed in the APOBEC3 subfamily (Table 2). Such acceleration of evolution characteristic of proteins involved in direct interactions with infectious agents (e.g., viruses)<sup>29</sup> which is compatible with the accelerated rate of evolution and the confirmed antiviral function of some APOBEC3 subfamily members.<sup>7,11,12,30,31</sup> Thus, suppression of infectious agents might be the original function of AID/APOBEC ancestors. Later in evolution, ancestors of AID, APOBEC1, APOBEC2 and APOBEC4 proteins gained functions different from the ancestral one and their evolution substantially slowed down (Table 2 and Fig. 2).

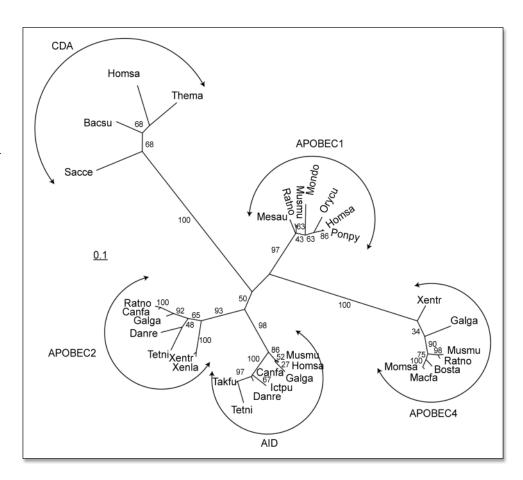


Figure 2. A maximum likelihood phylogenetic tree of the CDA/AID/APOBEC superfamily. Maximum likelihood RELL bootstrap support values are shown next to internal branches. The species abbreviations are as in Figure 1.

# **BIOLOGICAL IMPLICATIONS AND CONCLUSIONS**

The spectrum of biological functions of the editing enzymes of the AID/APOBEC family is expanding. In particular, it has been recently shown that these deaminases, in addition to mRNA editing, are required for innate immunity to retroviruses and humoral immunity.<sup>7,9-12</sup> Here we describe a new subfamily of AID/APOBEC homologs, APOBEC4, which is represented by readily identifiable orthologs in mammals, chicken, and frog, but not fishes. The Zn-coordinating motifs and the secondary structure of the APOBEC4 deaminase domain are evolutionarily conserved which suggests that APOBEC4 proteins possess the polynucleotide (deoxy)cytidine deamination activity. Examination of mouse expression arrays (Fig. 3) and Expressed Sequence Tag (EST) data for human, mouse, and rat showed that APOBEC4 is expressed primarily in testis (Table 1). Based on this observation, it is tempting to speculate that APOBEC4 is an editing enzyme for mRNAs involved in spermatogenesis.

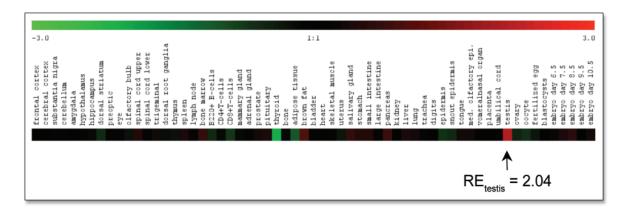


Figure 3. Preferential expression of mouse APOBEC4 in testis. The expression data are from the Novartis Gene Expression Atlas 2.<sup>32</sup> The figure shows a "thermal map" representation of the expression array data for APOBEC4 for the indicated mouse tissues. Colors are scaled to the relative expression level (REi) for each tissue (i). REi is equal to log2(Ei/M), where Ei is the signal intensity for each tissue and M is the median signal intensity for the gene. Tissue designations are given in the body of the figure. The >4x excess of APOBEC4 expression in testis over the median expression for all tissues was highly statistically significant according to one-sample t-test (p = 1.4e-36).

### References

- Cohen RM, Wolfenden R. Cytidine deaminase from *Escherichia coli*. Purification, properties and inhibition by the potential transition state analog 3,4,5,6-tetrahydrouridine. J Biol Chem 1971; 246:7561-5.
- Munch-Petersen A, Nygaard P, Hammer-Jespersen K, Fiil N. Mutants constitutive for nucleoside-catabolizing enzymes in *Escherichia coli* K12. Isolation, charactrization and mapping. Eur J Biochem 1972; 27:208-15.
- Carter Jr CW. The nucleoside deaminases for cytidine and adenosine: Structure, transition state stabilization, mechanism, and evolution. Biochimie 1995; 77:92-8.
- Navaratnam N, Morrison JR, Bhattacharya S, Patel D, Funahashi T, Giannoni F, Teng BB, Davidson NO, Scott J. The p27 catalytic subunit of the apolipoprotein B mRNA editing enzyme is a cytidine deaminase. J Biol Chem 1993; 268:20709-12.
- Teng B, Burant CF, Davidson NO. Molecular cloning of an apolipoprotein B messenger RNA editing protein. Science 1993; 260:1816-9.
- Durandy A. Activation-induced cytidine deaminase: A dual role in class-switch recombination and somatic hypermutation. Eur J Immunol 2003; 33:2069-73.
- Pham P, Bransteitter R, Goodman MF. Reward versus risk: DNA cytidine deaminases triggering immunity and disease. Biochemistry 2005; 44:2703-15.
- Conticello SG, Thomas CJ, Petersen-Mahrt SK, Neuberger MS. Evolution of the AID/APOBEC family of polynucleotide (deoxy)cytidine deaminases. Mol Biol Evol 2005; 22:367-77.
- Muramatsu M, Kinoshita K, Fagarasan S, Yamada S, Shinkai Y, Honjo T. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell 2000; 102:553-63.
- Revy P, Muto T, Levy Y, Geissmann F, Plebani A, Sanal O, Catalan N, Forveille M, Dufourcq-Labelouse R, Gennery A, Tezcan I, Ersoy F, Kayserili H, Ugazio AG, Brousse N, Muramatsu M, Notarangelo LD, Kinoshita K, Honjo T, Fischer A, Durandy A. Activation-induced cytidine deaminase (AID) deficiency causes the autosomal recessive form of the Hyper-IgM syndrome (HIGM2). Cell 2000; 102:565-75.
- Sheehy AM, Gaddis NC, Choi JD, Malim MH. Isolation of a human gene that inhibits HIV-1 infection and is suppressed by the viral Vif protein. Nature 2002; 418:646-50.
- Neuberger MS, Harris RS, Di Noia J, Petersen-Mahrt SK. Immunity through DNA deamination. Trends Biochem Sci 2003; 28:305-12.
- Sasada A, Takaori-Kondo A, Shirakawa K, Kobayashi M, Abudu A, Hishizawa M, Imada K, Tanaka Y, Uchiyama T. APOBEC3G targets human T-cell leukemia virus type 1. Retrovirology 2005; 2:32.
- Liao W, Hong SH, Chan BH, Rudolph FB, Clark SC, Chan L. APOBEC-2, a cardiac- and skeletal muscle-specific member of the cytidine deaminase supergene family. Biochem Biophys Res Commun 1999; 260:398-404.
- Anant S, Mukhopadhyay D, Sankaranand V, Kennedy S, Henderson JO, Davidson NO. ARCD-1, an apobec-1-related cytidine deaminase, exerts a dominant negative effect on C to U RNA editing. Am J Physiol Cell Physiol 2001; 281:C1904-16.
- Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucleic Acids Res 1997; 25:3389-402.
- Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004; 32:1792-7.
- Jones DT. Protein secondary structure prediction based on position-specific scoring matrices. J Mol Biol 1999; 292:195-202.
- Kumar S, Tamura K, Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform 2004; 5:150-63.

- Felsenstein J. Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods. Methods Enzymol 1996; 266:418-27.
- Hasegawa M, Kishino H, Saitou N. On the maximum likelihood method in molecular phylogenetics. J Mol Evol 1991; 32:443-5.
- Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003; 52:696-704.
- Hache G, Liddament MT, Harris RS. The retroviral hypermutation specificity of APOBEC3F and APOBEC3G is governed by the C-terminal DNA cytosine deaminase domain. J Biol Chem 2005; 280:10920-4.
- Dance GS, Beemiller P, Yang Y, Mater DV, Mian IS, Smith HC. Identification of the yeast cytidine deaminase CDD1 as an orphan C >U RNA editase. Nucleic Acids Res 2001; 29:1772-80.
- Scott J, Navaratnam N, Carter C. Molecular modelling and the biosynthesis of apolipoprotein B containing lipoproteins. Atherosclerosis 1998; 141(Suppl 1):S17-24.
- Navaratnam N, Fujino T, Bayliss J, Jarmuz A, How A, Richardson N, Somasekaram A, Bhattacharya S, Carter C, Scott J. *Escherichia coli* cytidine deaminase provides a molecular model for ApoB RNA editing and a mechanism for RNA substrate recognition. J Mol Biol 1998: 275:695-714.
- Huthoff H, Malim MH. Cytidine deamination and resistance to retroviral infection: Towards a structural understanding of the APOBEC proteins. Virology 2005; 334:147-53.
- Philippe H. Opinion: Long branch attraction and protist phylogeny. Protist 2000; 151:307-16.
- 29. Hughes AL. Adaptive evolution of genes and genomes. Oxford: Oxford Univ, 1999.
- Zhang J, Webb DM. Rapid evolution of primate antiviral enzyme APOBEC3G. Hum Mol Genet 2004; 13:1785-91.
- Sawyer SL, Emerman M, Malik HS. Ancient adaptive evolution of the primate antiviral DNA-editing enzyme APOBEC3G. PLoS Biol 2004; 2:E275.
- Su AI, Wiltshire T, Batalov S, Lapp H, Ching KA, Block D, Zhang J, Soden R, Hayakawa M, Kreiman G, Cooke MP, Walker JR, Hogenesch JB. A gene atlas of the mouse and human protein-encoding transcriptomes. Proc Natl Acad Sci USA 2004; 101:6062-7.