- **Supplementary Online Data** 1 2 Opossum APOBEC1 is a DNA mutator with retrovirus and retroelement 3 restriction activity 4 5 Terumasa Ikeda<sup>1,2,3,4,\*</sup>, Mayuko Shimoda<sup>5,6,7</sup>, Diako Ebrahimi<sup>2,3,4</sup>, John L. 6 VandeBerg<sup>8</sup>, Reuben S. Harris<sup>2,3,4,9</sup>, Atsushi Koito<sup>1</sup>, Kazuhiko Maeda<sup>6,7,\*</sup> 7 8 <sup>1</sup>Department of Retrovirology and Self-Defense, Faculty of Life Science, 9 Kumamoto University, Kumamoto 860-8556, Japan 10 <sup>2</sup>Department of Biochemistry, Molecular Biology, and Biophysics, 11 <sup>3</sup>Institute for Molecular Virology, <sup>4</sup>Masonic Cancer Center, University of 12 Minnesota, Minneapolis, Minnesota 55455, USA 13 <sup>5</sup>Department of Immunology, Graduate School of Medical Science, 14 Kumamoto University, Kumamoto 860-8556, Japan 15 <sup>6</sup>Laboratory of Host Defense, Research Institute for Microbial Diseases, 16 <sup>7</sup>WPI Immunology Frontier Research Center, Osaka University, Osaka 17 565-0871, Japan 18 <sup>8</sup>South Texas Diabetes and Obesity Institute, School of Medicine, The 19 University of Texas Rio Grande Valley, Brownsville/Harlingen/Edinburg, 20 21 Texas 78520, USA <sup>9</sup>Howard Hughes Medical Institute, University of Minnesota, Minneapolis, 22 Minnesota 55455, USA 23 24 \*Correspondence should be addressed to either of these authors 25 (e-mails: tikeda@umn.edu or kazmaeda@biken.osaka-u.ac.jp). 26 27 28 29 30
- 31
- 32



33

34 Figure S1. No A3 gene exists in monotremes and marsupials. (a) Phylogenetic 35 tree of selected monotreme (platypus), marsupial (opossum), and eutherian mammals (human, rabbit and mouse). (b) Genomic organization of A1, AID, and A3 genes. A 36 37 homology BLAST search was performed for all seven human A3 genes (A3A, A3B, 38 A3C, A3D, A3F, A3G, and A3H) in the genomes of platypus and opossum. Although all species have A1 and AID, non-eutherian mammals did not show any evidence for 39 40 homologous A3 genes between the conserved CBX6 and CBX7 genes. We also 41 found no evidence to suggest that a gene similar to one of human A3 genes exists in a 42 different locus in the opossum and platypus genomes. Green, orange, and blue colors 43 indicate Z1, Z2, and Z3 domains of A3 enzymes. For rabbit, two A3-like genes (shown 44 in glowing blue color) have been annotated. Based on the available sequences from Ensemble, the rabbit *A3G*-like gene (ensemble ID: ENSOCUG0000023532) appears 45 46 to encode a single Z1 domain, and the rabbit A3H-like gene (ensemble ID: ENSOCUG0000023913) has signatures of Z2 and Z3 domains. However, the 47 48 database sequences may not be complete or correctly annotated. 49

	10	20	30	40	50	60	70	80	90	100
			.					.		
Opossum	MNSKTGPSVGD	ATLRRRIKPWEF	VAFFNPQELRKI	TCLLYEIKW	IG-NQNIWRH	SNQNTSQHAEI	NFMEKFTAE	RHFNSSVRCSI	TWFLSWSPC	WECSKAI
Human	.T.EKT	PE	DV.YD.R	.A	.MSRK S	.GKTN.V.V	IKS.	.D.HP.MS		<b>.</b> .Q
Rabbit	.A.EKNK.	Y	EVD	.A	.ASSKTS	.GKTN.V.V	LL.S.	GRLGP.TC		М
Rat	.S.EVAV.	P	EVD.R	N .	. GRHS	TSNK.V.V	IT.	.Y.CPNT		GR
Mouse	.S.EVAV.	P	EVD.R	N .	. GRHSV	TSN.V.V	LT.	.Y.RPNT		GR
	Catalytic domain									
	110	120	130	140	150	160	170	180	190	200
			.			.		.		.
Opossum	RKFLDHYPNVT	LAIFISRLYWHM	DQQHRQGLKEL	/HSGVTIQIM	ISYSEYHYCW	RNFVDYPQGEE	DYWPKYPYLI	WIMLYVLELHC	IILGLPPCL	KISGSHS
Human	.ESRH.G	.V.YVAF	NRD.	N	RAYH	NP.D.	AHQP.	.MA	s	RRWQ
Rabbit	.E.SQH.G	.IVAFQ	. RRN D	TVRV.	.vc	ENP.KA	AQ R PR	.MLM.AY.		RR.Q
Rat	TESRH	.F.Y.AH.A	. PRN RD . 1	[S	TEQ.SG	N.SPSN.	AHRH.	.VRY.		N. LRRKQ
Mouse	TESRH.Y	.F.Y.AH.T		[ <b>S</b>	TEQC	NPSN.	ARH.	. VK Y.		LRRKQ
	210	220	230							
			.							
Opossum	NQLALFSLDLQ	DCHYQKIPYNVL	VATGLVQPFVTV	<b>VR</b> 235 aa						
Human	.H.TF.R.H	NTPHI.	LIH.S.A	236 aa	(67%)					
Rabbit	KTFTP.	YKMPYI.	LLS.P	236 aa	(64%)					
Rat	P TF. TIA	SRL.PHI.	WK	229 aa	(64%)					
Mouse	PTF.TIT	TR. PHL.	WK	<b></b> 229 aa	(66%)					

50

51 Figure S2. Alignment of amino acid sequences of mammalian A1s. Sequences 52 of the predicted full-length mammalian A1 proteins were aligned with Clustal W 53 software. The numbers above each amino acid sequences are the positions of amino 54 acid residues. The identical amino acids to opossum A1 are indicated by a single dot. 55 The amino acid positions that were not aligned are shown with a bar. The amino acid 56 residues involved in the catalytic reaction are shown in red. The amino acid length of 57 opossum A1 is 235 aa. The percent of amino acid identity of each eutherian A1 with 58 opossum A1 is indicated in parentheses at the end of each sequence. 59



Figure S3. Subcellular localisation of opossum A1 in 293T cells. (a) Images
show HA-tagged A1s stained with FITC (green) and nuclei stained with DAPI (blue).
(b) Nuclear and cytoplasmic distribution levels of A1s. The percentage nuclear and
cytoplasmic localisation is indicated as the graphs (n = 20, average +/- SD). *P* values
represent comparisons with rabbit A1.