Supplementary Material

Essential factors involved in the precise targeting and insertion of telomere-specific non-LTR retrotransposon, SART 1Bm

Authors:

Narisu Nichuguti* and Haruhiko Fujiwara*¹

Affiliations:

*Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Kashiwa, Chiba, 277-8562, Japan.

¹Corresponding author: Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Bioscience Bldg., Kashiwanoha 5-1-5, Kashiwa, Chiba, 277-8562, Japan.

Tel: +81-4-7136-3659; Fax: +81-4-7136-3660;

Email: haruh@edu.k.u-tokyo.ac.jp
Supplementary Figure S1 Full-length gels regarding data shown in Fig. 1b
Supplementary Table 1. Primers used in this study

| Name                          | Sequence (5’ to 3’)                                      |
|-------------------------------|--------------------------------------------------------|
| **SART1Bm target plasmid construction** |                                                        |
| E-TTAGG1-S-E-s                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG1-S-E-a                 | AATTCGACTAGTGATTTCTAAACCTAAATCG                        |
| E-TTAGG2-S-E-s                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG3-S-E-s                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG3-S-E-a                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG5-S-E-s                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG6-S-E-a                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG8-S-E-s                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG8-S-E-a                 | AATTCGATTTTAGGAATCACTAGTG                               |
| E-TTAGG13-S-E-s                | AATTCGATTTTAGGAATCACTAGTG                               |
| (TTAGG)6                       | TTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTTAGGGTACAGGTG |
| (CCTAA)6                       | CCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTAACCTA
| (TTTAGGG)3 | TTTAGGGTTAGGGTTAGGG |
| (CCCTAAA)3 | CCCTAAACCCTAAACCCTAAA |

**EGFP/3’UTR series plasmid**

BamHI-EGFP-S96

| BamHI-EGFP-S96 | AAGGATCCATGGTGAGCAAGGGCGAGG |

EcoRI-EGFP-A813

| EcoRI-EGFP-A813 | AAGAATTCCTTACATCATAGGATTCATG |

SART1-S6221-EcoRI-Takahashi

| SART1-S6221-EcoRI-Takahashi | TTTTTGAATTCGGACCGTCGGGCGTC |

**TTT-XbaI-PolyA(0)+SART1 3’UTR**

| TTT-XbaI-PolyA(0)+SART1 3’UTR | TTTctagaGGTATCGATGGGGAATCCC |

**TTT-XbaI-PolyA(5)+SART1 3’UTR**

| TTT-XbaI-PolyA(5)+SART1 3’UTR | TTTctagaTTTTTTGTATCGATGGGGAATCCC |

**TTT-XbaI-PolyA(10)+SART1 3’UTR**

| TTT-XbaI-PolyA(10)+SART1 3’UTR | TTTctagaTTTTTTTTTTGTATCGATGGGGAATCCC |

**TTT-XbaI-PolyA(18)+SART1 3’UTR**

| TTT-XbaI-PolyA(18)+SART1 3’UTR | TTTctagaTTTTTTTTTTTTTTTTGTATCGATGGGGAATCCC |

**TTT-XbaI-PolyC(18)+SART1 3’UTR**

| TTT-XbaI-PolyC(18)+SART1 3’UTR | TTTctagaGGGGGGGGGGGGGGGGGTATCGATGGGGAATCCC |

**TTT-XbaI-PolyA(0)AGG+SART1 3’UTR**

| TTT-XbaI-PolyA(0)AGG+SART1 3’UTR | TTTctagaCCTGATCGATGGGGAATCCC |

**Ex vivo retrotransposition assay**

S16131

| S16131 | AGAAAGAGAGTGCGACCCAAACTCAGTT |

A878T

| A878T | GCGTGCAGAATTCACTAGGTTT |

**Amp-F1**

| Amp-F1 | GAACGTTTTTCCAATGATGAGCATACTTT |

**Amp-R1**

| Amp-R1 | CCAATGCTTAATCATGAGGATAC |

**Trans-in vivo retrotransposition assay**

pEGFP1-S688

| pEGFP1-S688 | GCAACCAACTACCTGAGGAC |

**CCTAA6**

| CCTAA6 | CCTAACCCTAACCCTAA CCTAACCCTAA |
### Supplementary Table 2. Artifactual baculovirus sequences detected in *ex vivo* assay.

| Type | Query Length (bp) | Sbjct (bp) | Identities | Synthetic baculovirus AcMNPV-WIV-Syn1, complete sequence (ID: KY792989.1) | Baculovirus p10 and p74 genes’ junction region | Reverse Primer: AE/TF GCC/CAG/ACT/TTG/ATT/GTT | Forward primer: S16131
|------|------------------|------------|------------|--------------------------------------------------------------------------|-------------------------------------------------|-----------------------------------------------|-----------------------------------------------|
| 1    | 398 (8-398)      | 36785 to 37175 | 391/391(100%) | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGTTTNTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGTNTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 2    | 455 (1-392)      | 36774 to 37163 | 36785 to 37175 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 3    | 378 (1-375)      | 36800 to 37174 | 36785 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 4    | 439 (1-364)      | 36774 to 37163 | 36785 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 5    | 441 (59-441)     | 36785 to 37167 | 36785 to 37175 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 6    | 367 (6-333)      | 36785 to 37174 | 36800 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 7    | 442 (1-384)      | 36785 to 37167 | 36785 to 37175 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 8    | 456 (1-330)      | 36785 to 37175 | 36800 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 9    | 426 (1-332)      | 36785 to 37175 | 36800 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 10   | 391 (39-391)     | 36785 to 37175 | 36785 to 37175 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 11   | 405 (1-332)      | 36785 to 37175 | 36800 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
| 12   | 412 (1-391)      | 36785 to 37175 | 36800 to 37174 | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | AGAAAGAGAGTGCGACCCAAACTCAGTTAAACGGGCTGGAAGAATCATTCCAGCTTTTGGACGGTTTGCCCGCTCAATTG | GCC/CAG/ACT/TTG/ATT/GTT                          | S16131                                         |
Supplementary Table 3. Nucleotide sequence of the 3′ junction of ex vivo retrotransposed SART1Bm and point mutated (TTAGG) tracts

| Lane | Target plasmid | SART1Bm 3′ UTR/(A)n | Telomeric repeats | Vector sequence | Clones |
|------|----------------|----------------------|-------------------|-----------------|--------|
| 2    | (CTAGG)_{42}  | ---CATCGATACC(A)_{30} | AGG(CTAGG)_{22}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{34} | AGG(CTAGG)_{9}    | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{29} | AGG(CTAGG)_{29}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{37} | AGG(CTAGG)_{32}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{29} | AGG(CTAGG)_{32}   | AATCACTAGT---   | 1      |
| 3    | (TCAGG)_{23}  | ---CATCGATACC(A)_{36} | AGG(TCAGG)_{19}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{10} | AGG(TCAGG)_{11}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{23} | AGG(TCAGG)_{3}    | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{25} | AGG(TCAGG)_{9}    | AATCACTAGT---   | 1      |
| 4    | (TTCGG)_{35}  |                       |                   |                 | 0      |
| 5    | (TTAGG)_{30}  | ---CATCGATACC(A)_{47} | ACG(TTAGC)_{9}    | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{18} | ACG(TTAGC)_{19}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{15} | ACG(TTAGC)_{14}   | AATCACTAGT---   | 1      |
| 6    | (TTAGG)_{30}  | ---CATCGATACC(A)_{20} | ACG(TTAGC)_{12}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{36} | ACG(TTAGC)_{12}   | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)_{31} | ACG(TTAGC)_{3}    | AATCACTAGT---   | 1      |
Supplementary Table 4. Nucleotide sequence of the 3′ junction of *ex vivo* retrotransposed SART1Bm and telomeric repeats of other species

| Lane | Target plasmid | SART1Bm 3′ UTR/(A)$_m$ | Telomeric repeats | Vector sequence | Clones |
|------|----------------|------------------------|-------------------|-----------------|--------|
| 2    | (TTAGG)$_{34}$ | ---CATCGATACC(A)$_{29}$ | AGGG(TTAGGG)$_{31}$ | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)$_{18}$ | AGGG(TTAGGG)$_{17}$ | AATCACTAGT---   | 1      |
| 3    | (TTAGG)$_{33}$ | ---CATCGATACC(A)$_{14}$ | AGGC(TTAGGC)$_{12}$ | AATCACTAGT---   | 1      |
| 4    | (TTAGG)$_{21}$ | ---CATCGATACC(A)$_{29}$ | AGGG(TTTAGGG)$_{3}$ | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)$_{13}$ | AGGG(TTTAGGG)$_{2}$ | AATCACTAGT---   | 1      |
|      |                | ---CATCGATACC(A)$_{27}$ | AGGG(TTTAGGG)$_{11}$ | AATCACTAGT---   | 1      |
Supplementary Table 5. Nucleotide sequence of 3' junction clones obtained from the EGFP/ SART1 3' UTR/(A)ₙ construct.

| Reporter | EGFP/EcoRI | SART1Bm 3' UTR/(A)ₙ | Telomeric repeats | 6 |
|----------|------------|---------------------|-------------------|---|
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₁₈ | AGG(TTAGG)₁₂ | 1 |
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₂₉ | AGG(TTAGG)₃₂ | 1 |
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₁₈ | AGG(TTAGG)₅₇ | 1 |
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₇ | AGG(TTAGG)₄ | 1 |
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₁₂ | AGG(TTAGG)₂₇ | 1 |
| b-TACAAGTAgaattc | GGACCGT---GATACC(A)₁₄ | AGG(TTAGG)₁₂ | 1 |
**Supplementary Table 6. Nucleotide sequence of 3′ junction clones obtained from the A-0 and A-5 donor constructs**

| Reporter       | EGFP/EcoRI | SART1Bm 3′ UTR | Telomeric repeats | Clone |
|----------------|------------|----------------|-------------------|-------|
| **A-0**        |            |                |                   |       |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(5\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(6\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(12\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(38\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(41\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TT\(^{398}\) | AGG(TTAGG)\(8\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{16}\) | AGG(TTAGG)\(8\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{18}\) | AGG(TTAGG)\(9\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{17}\) | AGG(TTAGG)\(10\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{13}\) | AGG(TTAGG)\(10\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{27}\) | AGG(TTAGG)\(15\) | 1 |
| **A-5**        |            |                |                   |       |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(6\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---TGAGG\(^{394}\) | AGG (TTAGG)\(9\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{17}\) | AGG(TTAGG)\(1\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{24}\) | AGG(TTAGG)\(6\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{20}\) | AGG(TTAGG)\(6\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{15}\) | AGG(TTAGG)\(8\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{20}\) | AGG(TTAGG)\(9\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{24}\) | AGG(TTAGG)\(11\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{29}\) | AGG(TTAGG)\(12\) | 1 |
| ---TACAAGTAAgaattc | GGACCGT---GATACC(A)\(_{24}\) | AGG(TTAGG)\(16\) | 1 |

*Nucleotide positions are indicated with the junction of cleavage target site AGG and 3′UTR end, defined as 0. The telomeric repeat-like sequences AGG in the end of inserted copies are underlined and indicated at upstream -394.
### Supplementary Table 7. Nucleotide sequence of 3’ junction clones obtained from the A-10 and A-18 donor constructs

| Reporter | EGFP/EcoRI | SART1Bm 3’ UTR | Telomeric repeats | Clone |
|----------|------------|----------------|------------------|-------|
| A-10     | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{26} | AGG(TTAGG)\textsubscript{1} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{28} | AGG(TTAGG)\textsubscript{1} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{20} | AGG(TTAGG)\textsubscript{3} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{28} | AGG(TTAGG)\textsubscript{5} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{10} | AGG(TTAGG)\textsubscript{12} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{29} | AGG(TTAGG)\textsubscript{21} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{29} | AGG(TTAGG)\textsubscript{25} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{20} | AGG(TTAGG)\textsubscript{30} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{11} | AGG(TTAGG)\textsubscript{55} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{29} | AGG(TTAGG)\textsubscript{65} | 1     |
| A-18     | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{16} | AGG(TTAGG)\textsubscript{7} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{17} | AGG(TTAGG)\textsubscript{8} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{17} | AGG(TTAGG)\textsubscript{8} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{29} | AGG(TTAGG)\textsubscript{9} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{14} | AGG(TTAGG)\textsubscript{12} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{9} | AGG(TTAGG)\textsubscript{22} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{12} | AGG(TTAGG)\textsubscript{22} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{21} | AGG(TTAGG)\textsubscript{27} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{29} | AGG(TTAGG)\textsubscript{32} | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc(A)\textsubscript{18} | AGG(TTAGG)\textsubscript{57} | 1     |
Supplementary Table 8. Nucleotide sequence of 3' junction clones obtained from the A-AGG and C-18 donor constructs

| Reporter | EGFP/EcoRI | SART1Bm 3' UTR | Telomeric repeats | Clone |
|----------|------------|----------------|-------------------|-------|
| AGG      | ---TACAAGTAgaattc | GGACCGT---TGAGG<sup>394<sup>1</sup> | AGG (TTAGG)<sub>3</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---TGAGG<sup>394<sup>1</sup> | AGG (TTAGG)<sub>9</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>15</sub> | AGG (TTAGG)<sub>5</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>17</sub> | AGG (TTAGG)<sub>6</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>16</sub> | AGG (TTAGG)<sub>3</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>24</sub> | AGG (TTAGG)<sub>12</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>23</sub> | AGG (TTAGG)<sub>12</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>16</sub> | AGG (TTAGG)<sub>54</sub> | 1     |
| C-18     | ---TACAAGTAgaattc | GGACCGT---ATAGG<sup>230<sup>1</sup> | AGG (TTAGG)<sub>8</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>14</sub> | AGG (TTAGG)<sub>4</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>22</sub> | AGG (TTAGG)<sub>8</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>29</sub> | AGG (TTAGG)<sub>6</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>16</sub> | AGG (TTAGG)<sub>6</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>8</sub> | AGG (TTAGG)<sub>18</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>19</sub> | AGG (TTAGG)<sub>23</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>12</sub> | AGG (TTAGG)<sub>26</sub> | 1     |
|          | ---TACAAGTAgaattc | GGACCGT---GATAcc (A)<sub>40</sub> | AGG (TTAGG)<sub>40</sub> | 1     |

* Nucleotide positions are indicated with the junction of cleavage target site AGG and 3' UTR end, defined as 0. The telomeric repeat-like sequences AGG in the end of inserted copies are underlined and indicated at upstream -394 and -233.