

Additional file 1:

Targeted gene knockin in zebrafish using the 28S rDNA-specific non-LTR-retrotransposon R2OI

Azusa Kuroki-Kami^{1,3}, Narisu Nichuguti^{1,3}, Haruka Yatabe¹, Sayaka Mizuno¹, Shoji Kawamura², and Haruhiko Fujiwara^{1,*}

¹ Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Bioscience Bldg. 501, Kashiwanoha 5-1-5, Kashiwa, Chiba, 277-8562, Japan.

² Department of Integrated Biosciences, Graduate School of Frontier Sciences, The University of Tokyo, Bioscience Bldg. 502, Kashiwanoha 5-1-5, Kashiwa, Chiba, 277-8562, Japan.

³ These authors contributed equally to this work.

* To whom correspondence should be addressed. Tel: +81-4-7136-3659; Fax: +81-4-7136-3660; Email: haruh@edu.k.u-tokyo.ac.jp

Additional file 1:

Figure S1. Comparison of R2 element insertion site.

Figure S2. Nine clones of R2OI obtained from medaka fish genomic DNA.

Figure S3. A scheme for transgenesis in zebrafish using R2OI.

Figure S4. F1 progeny analysis of two founder lines, R2f-9 and R2f-10.

Figure S5. Experimental approach for analysis of hsp-GAL4:UAS-EGFP expression induced by heat shock.

	-100		-31
Medaka	GTTGACGCGATGTGATTCTGCCCAGTGCTCTGAATGTCAAAGTGAAGAAATTC AATGAAGCGCGGGTAA		
Zebrafish	GTTGACGCGATGTGATTCTGCCCAGTGCTCTGAATGTCAAAGTGAAGAAATTC AATGAAGCGCGGGTAA		
Human	GTTGACGCGATGTGATTCTGCCCAGTGCTCTGAATGTCAAAGTGAAGAAATTC AATGAAGCGCGGGTAA		
Mouse	GTTGACGCGATGTGATTCTGCCCAGTGCTCTGAATGTCAAAGTGAAGAAATTC AATGAAGCGCGGGTAA		
	-30	-1+1	+40
Medaka	ACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAA		
Zebrafish	ACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAA		
Human	ACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAA		
Mouse	ACGGCGGGAGTAACTATGACTCTCTTAAGGTAGCCAAATGCCTCGTCATCTAATTAGTGACGCGCATGAA		
		↑	
	+41	R2	+100
Medaka	TGGATGAACGAGATTCCCACTGTCCCTACCTCTATCTAGCGAAACCACAGCCAAGGGAA		
Zebrafish	TGGATGAACGAGATTCCCACTGTCCCTACCTCTATCTAGCGAAACCACAGCCAAGGGAA		
Human	TGGATGAACGAGATTCCCACTGTCCCTACCTACTATCCAGCGAAACCACAGCCAAGGGAA		
Mouse	TGGATGAACGAGATTCCCACTGTCCCTACCTACTATCCAGCGAAACCACAGCCAAGGGAA		

Figure S1. Comparison of R2 element insertion site. R2 insertion site sequence alignment of four vertebrates, medaka fish(CP020779.1), zebrafish(BX537263.17), human (M11167.1) and mouse (X00525.1). An approximately 200 base pair (bp) region around the R2 insertion site in 28S rDNA is strongly conserved except for differences at two nucleotides (shaded in yellow). The insertion site (red) of R2 is between -1 (G) and +1(T).

	CloneA	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneB	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneC	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneD	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneE	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneF	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneG	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneI	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	CloneJ	MGTDTVYVGQ	DYP8GLSKRV	PARGLVAGPML	RERSCHAHVF	RAGHMNNWRT	SLP8GRWDQP	ALEKSRVLTR	SVATATDPEI	TSPYGKSYST	STQVQEEDWC
	101										
	CloneA	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneB	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneC	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneD	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneE	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneF	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneG	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneI	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	CloneJ	SRESGWISPG	LAPEEPSVVTS	EITASMVATM	RVATEEEVILE	PQEPQVVITL	PHGRNVPPFG	LAEQDTASPI	EVSVLLFDLA	ENCNFCGVPS	GGLRLLGKHFI
	201										
	CloneA	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneB	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneC	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneD	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneE	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneF	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneG	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneI	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	CloneJ	AVRHAGVPVT	YECKRCANRS	PNSHSISCHV	PKCRGHARVP	SGDPGIACDL	CEARFATEVG	VAQHRRHVHP	VEWNKVRLER	RGARGGGIKA	TKLWSVAEVE
	301										
	CloneA	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneB	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneC	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneD	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneE	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneF	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI
	CloneG	TLIRLIREHG	DSGATYQLIA	DELGRGKTAE	QVRSKKRLLR	IDTASNSPDD	AEVEERLES	LAVRSSSSRP	PSLVATRvre	AVARGESEGG	EETRAIAALI

601													700
CloneA	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneB	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneC	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneD	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneE	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneF	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneG	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneI	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
CloneJ	GWRPVTIGSM	VTRLFSRLT	MRLTRACPIN	PRQRGFLASS	SGCAENLLIF	DEIVRRSRDD	GGFLAVVFPD	FARAFDSISH	EHILCVLEEG	GLDRHVIGLI			
701													800
CloneA	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneB	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneC	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneD	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneE	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneF	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneG	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneI	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
CloneJ	RNSYVDCVTR	VGCGEGMTTP	IQMKVGVKQG	DPMSPLLFLNL	AMDLPIHKLE	TAGTGLKWGD	LSIATLAFAD	DLVLVSDSEE	GMGRSLGILE	KFCQLTGLRV			
801													900
CloneA	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneB	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneC	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneD	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneE	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneF	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneG	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneI	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
CloneJ	QPRKCHGFEM	DKGVVNGCGT	WEICGSPIMH	IPPGESVRYL	GVQVGPGRGV	MEPDLITVTH	TWIERISEAP	LKPSQRMVRL	NSFALPRIY	QADLGKVTVT			
901													1000
CloneA	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			
CloneB	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			
CloneC	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			
CloneD	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			
CloneE	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			
CloneF	KLAQIDGIVR	KAVKKWLHLS	PSTCNGLLYS	RNRDGLGLL	KLERLIPSVR	TRRIYMSRS	PDITWRMTS	HSVSKSDWEM	LWQVAGGERG	SAPVMGAVEA			

Figure S2. Nine clones of R2OI obtained from medaka fish genomic DNA. **A** Amino acid alignment of nine clones. Different amino acids are shaded in light blue. **B** Schematic diagram of nine R2OI clones. Arrows indicate amino acid exchanges from all consensus sequences. Green arrows, conserved region exchanges; Black arrow, similar polarity exchanges; Red arrows, other exchanges. **C** Nucleic acid sequences of clone F, used in this study. Green, ORF; Black UTRs. Accession number: LC349444.

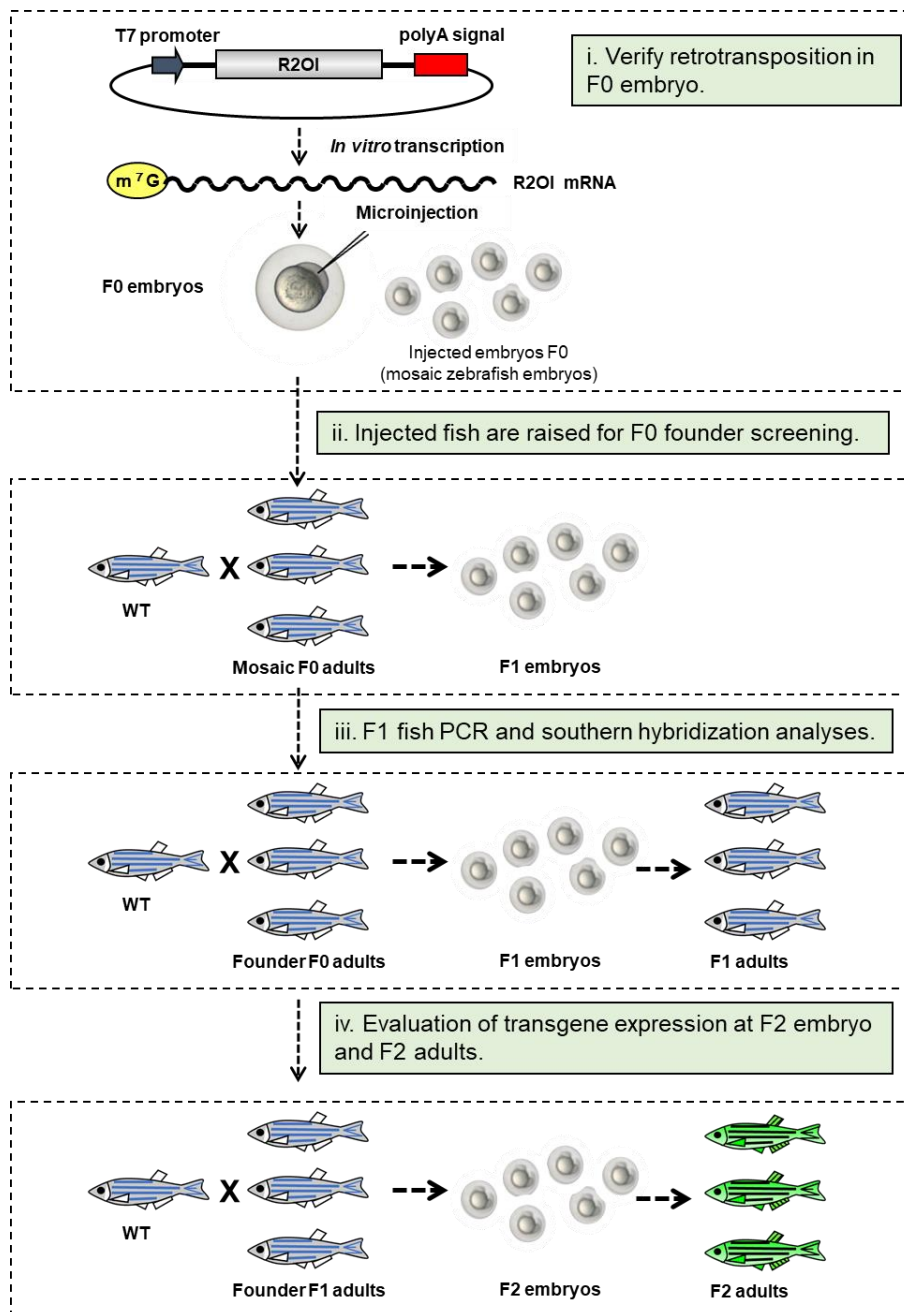


Figure S3. A scheme for transgenesis in zebrafish using R2OI.

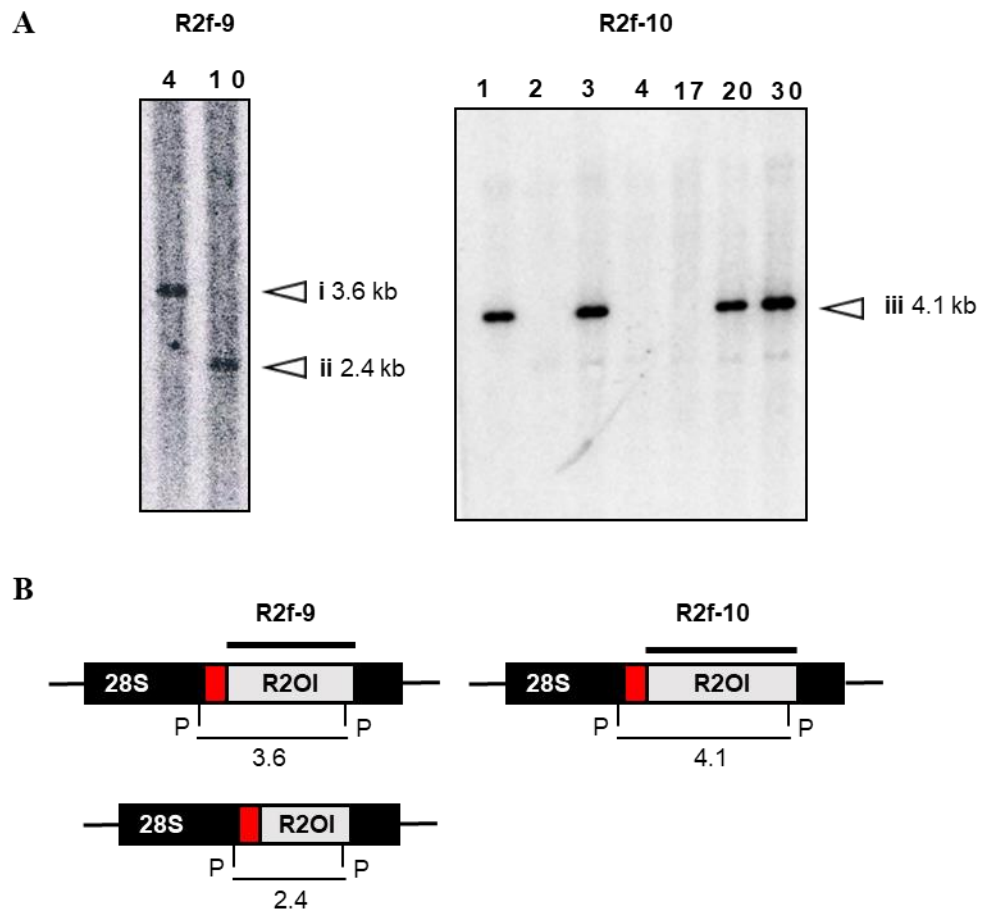


Figure S4. F1 progeny analysis of two founder lines, R2f-9 and R2f-10. **A** F1 fish Southern blot hybridization of representative founders R2f-9 (Left panel) and R2f-10 (Right panel). Genomic DNA were digested with *PvuII*. **B** Schematic diagram of each R2OI insertions. P, *PvuII* site; thick line, probe; thin lines and numbers, signal lengths; Red boxes indicate vector or non-templated sequence.

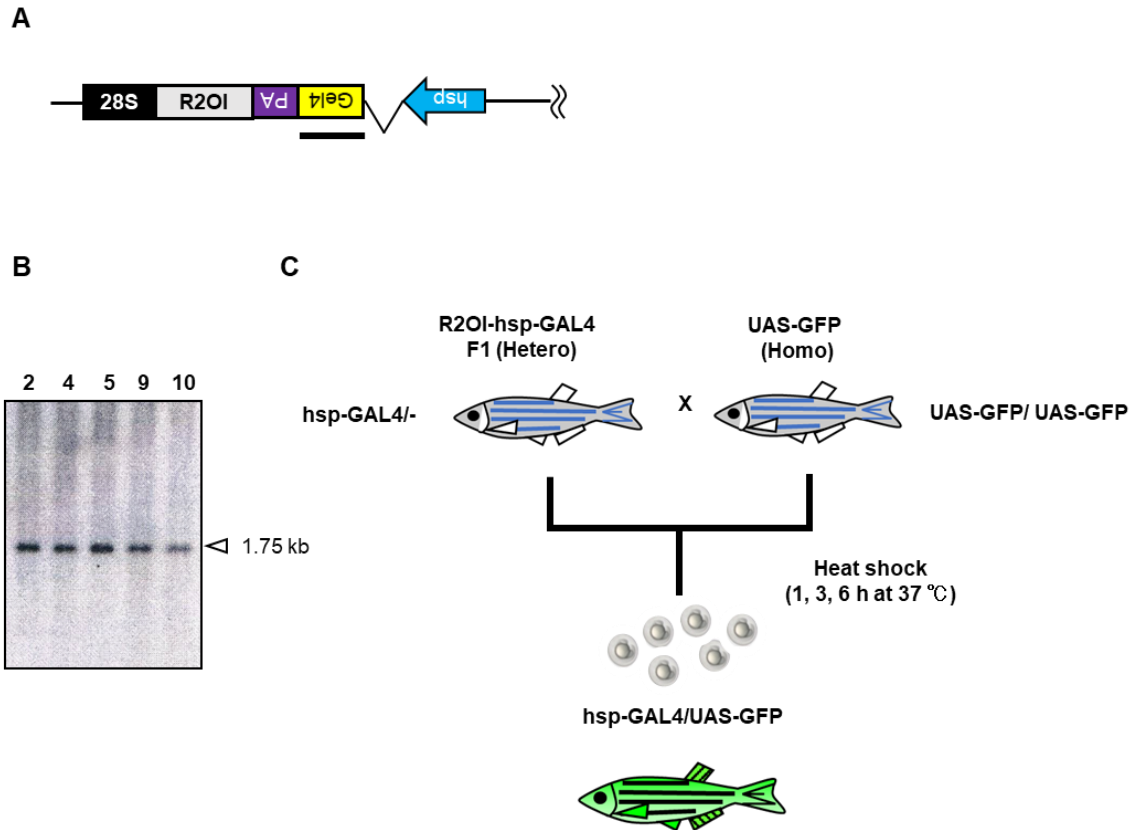


Figure S5. Experimental approach for analysis of hsp-GAL4:UAS-EGFP expression induced by heat shock. **A** Schematic diagram of hsp-GAL4 construct. thick line, probe. **B** Transgenic progeny analysis from R2OI-hsp-GAL4 founder. **C** Two lines R2OI-hsp-GAL4 and UAS-GFP were crossed and yielded embryos were heat shocked at 37 °C for 1, 3, 6, hours.