SUPPORTING INFORMATION

Efficient and reproducible multigene expression after single-step transfection using improved BAC transgenesis and engineering toolkit

Binhui Zhao[1,†,#], Pankaj Chaturvedi[1,†], David L. Zimmerman[1,‡] and Andrew S. Belmont[1,§,*]

[1] Department of Cell and Developmental Biology, University of Illinois, Urbana, Illinois, 61801, United States of America

[#] Present Address: Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.

[‡] Present Address: Biology Department, College of the Ozarks, Point Lookout, Missouri, 65726, United States of America

[§] Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

* Email: asbel@illinois.edu. Tel: +1-217-244-2311

Supplementary Table S1: List of primers used in this study.

Dual-reporter promoter assay

Primer ID	Primer sequence	Description	PCR product size
BZ#127	TTAAGCGGCCGCTCGTCATCACTGAGGTGGAG	NotI-hEEF1a promoter forward primer	1.2kb
BZ#204	TTAAGGCGCGCCCACGACACCTGAAATGGAAG	AscI-hEEF1a promoter reverse primer	
BZ#210	TTAAGCGGCCGCATGACAGCAGAGATCCAGTTT	NotI-hUBC promoter (pUGG as template) forward primer	1.4kb
BZ#211	TTCCACCGGTGAAGCTACTGTACACCAACCT	AgeI-hUBC promoter (pUGG as template) reverse primer	
BZ#151	TTAAGCGGCCGCAAGCAAAACAAATCAAGAGCA	NotI-hRPL32 promoter forward primer	2.2kb
BZ#152	GGCCACCGGTGATGCCTTTTGGGGAAGA	AgeI-hRPL32 promoter reverse primer	
BZ#147	TTAAGGCGCGCCTGCTCTACGTTTTAGGATGGAG	AscI-hPPIA promoter forward primer	2.7kb
BZ#148	GGCCACCGGTGGCTAATAGTACACGGTTTTCCTC	AgeI-PPIA promoter reverse primer	
BZ#141	TTAAGCGGCCGCATGGCGGCACCTTATTTATG	NotI-hB2M promoter forward primer	1.1kb
BZ#142	GGCCACCGGTTGCTGTCAGCTTCAGGAATG	AgeI-hB2M promoter reverse primer	
BZ#135	TTAAGCGGCCGCTGGGTGTAGTGGGAGGTAGG	NotI-hRPS3A promoter forward primer	1.4kb
BZ#136	GGCCACCGGTCTGGTCAGAGAGCCAAAAGG	AgeI-hRPS3A promoter reverse primer	
BZ#153	TTAAGGCGCGCCAGAAGGTGGATTTGGTGAGC	AscI-hGUSB promoter forward primer	2.3kb
BZ#154	GGCCACCGGTCCATCTTGGTTGAGGACGAG	AgeI-hGUSB promoter reverse primer	
HB2-RZ-check 5' NEW f	TTG CTC CCC TGT AAA CTG CT	forward primer for checking 5' junction of mRFP expression ca	2-4kb, depending on the size of the promoter in front of mRFP
HB2-RZ-check 5' NEW r	CTT GGC CAT GTA GGT GGT CT	reverse primer for checking 5' junction of mRFP expression ca	
HB2-RZ-check 3' f	CTA TGA AAG GTT GGG CTT CG	forward primer for checking 3' junction of mRFP expression ca	2kb
HB2-RZ-check 3' r	AGT AGC TGG GGC TGG GTA AT	reverse primer for checking 3' junction of mRFP expression ca	

Construction of BACs containing the UBC-GFP-Zeocin cassette

For amplifying the UBC-GFP-Zeocin-FRT-Galk-FRT cassette with flanking homologous arms

Primer Id	Sequence
ForBAC-UBBrec	AGCAAAGTGCTCCAAATATGCTCAGTCATGGAGTGGAGACAGATAAGTAAACAGACGCTGGGAATAGAGTGGAGacagcagagatccagt
RevBAC-UBBrec	AATGCCAATGCAGCACCTTGCTTGACTCTGGCCTGGCTCTGTATCACGACATATCCCCAGCATCTGTTTATTTATtgttggctagtgcgt
ForBAC-DHFRrec	ATAAGTTCATGGGGTCACAGTTTAGCCTGTAACCCATGAATGCACATCTGTACATGCATTATTCATTGTTCTATGacagcagagatccagt
RevBAC-DHFRrec	TAATGGTTTTCTCCAGAAATCCTAATTCTGATATTATACCAAGAGCAACTTCAGAATAAGTTTCCTAGAATTCtgttggctagtgcgt
ForBAC2207K13rec	CCAGGGCCTTTTTAGTGGATGTTAATGAGAACAAAGAAACCTCCTCCTAACATGGCGAAAGGTCTTGTCATCAGacagcagagatccagt
RevBAC2207K13rev	TGTTCATCCAAGATCAAAACAATGGGTCAATTTCATCAGCTCCTTTAAATAATGCATCTAGAAAGTGAAAAATAGtgttggctagtgcgt
ForROSArec	AGATTGGACCCTATTCTGAAGCCCAGGCTGTCCAACTCTAAACGAGTCTGTCTTACCTACACCAACAAGTATTCacagcagagatccagt
RevROSArec	TAGAGTAAAAGTTGACCATAGTGGGACCTGGTCAAGCAGACTACAATAGCAGAATCAGCGGGTATACCTGTAATCtgttggctagtgcgt
ForHBBrec	TCTTTCCGGATCATGCTTACAGATCTGATTACACCCAAGAAGCCTAAATTGTGTGTCAGCCAACTAGATGTGGACacagcagagatccagt
RevHBBrec	AACCACCCTATCACAGATTCTGCTTCAAAGGCACTCATCTTGGTGAATGAATATCTAGTCCCAGAAGGAGGCTAtgttggctagtgcgt

For screening of correct insertions

Primer Id	Sequence
ForUBB-PCR	GCCTGCTGAGTAAATAAACATGTTG
RevUBB-PCR	TTCACACCTTTTCCTGCAGGAA
ForDHFR-PCR	CTTAGGTATGCTAGGCTATGGA
RevDHFR-PCR	CTCAGGTTATATGCCTCGAGT
ForBAC2207K13PCR	CAGGCAGATGTACTTCTGAGTGA

RevBAC2207K13PCR	AGCTTATGACTGCTCCTGCTGT
ForROSA-PCR	ACCATCTGGTTCTGTTTGGAGT
RevROSA-PCR	TCTGGTCTGTCTGAAGACAGC
ForHBB-PCR	AGGCCTTCTGAAAGCTTACGAT
RevHBB-PCR	CACACACATGGGCATACATATC
GalKORFforw	GTGAACACACCGACTACAACGA
GalKORFrev	TCGTTGTAGTCGGTGTGTTCAC

Transgene copy number estimation

Primer Id	Sequence	Description
SGKforw	TGCGTGGTCAGTACCCTAGC	mouse *Sgk1* , endogenous control, forward
SGKrev	ATCCGGTAGCCCAAAGGAG	mouse *Sgk1* , endogenous control, reverse
mHPRT1for	ACGAGGAGTCCTGTTGATGT	mouse *Hprt1* , endogenous control, forward
mHPRT1rev	GGACGCAGCAACTGACATT	mouse *Hprt1* , endogenous control, reverse
Zeo-GFP2for	CCAACTTCAGGGATGCCAGT	UBC-GFP-ZeoR copy number, forward
Zeo-GFP2rev	TGGCCAGCTAGCTTTAGTCC	UBC-GFP-ZeoR copy number, reverse

Primers used for multi-reporter bac construction

DH2-4rev	tactatgtgacagccaatgtgagt
DH3-1for	atcaagactgcatcttcagctca
DH3-4rev	agctgacagcacacgtgtgta
DHM2-Seq2	TCACGTACCCACCGTCTAGT
DHM4F2-Fw	gcagtactgccctgcactgtactatagaatCCTCAGTGATAAAACTTGTTATCTGAT
DHM4F2-R	TACCACCTGTGGCATTTGGAATAG
DHM4F3-R	AAGGAAATGTGTAGGGAAATGGGT
ForCerFib	ggggtaccATG GTG AGC AAG GGC GAG GAG
Fw-M4F2-BamHI	AGCGAGAGCGGATCCTGGATGGTCTGTAAGTTTGATTGGG
GA-hRPL32-fwd	atgttctttcctgcgttatcAAGCAAAACAAATCAAGAGC
GA-hRPL32-rev	ccattcgcgcaagatcGATGCCTTTTGGGGAAGAAG
GA-RM01-Spec-For	atttccccgaaaagtgccacctgggtccttTAACAAATAAGCCACCTAAAG
GA-RM01-Spec-Rev	tttaaaatacctcgcgagtggcaacactgaACTATGCATGTGTCTTCTC
GA-RM02-Spec-For	atttccccgaaaagtgccacctgggtccttCTGTAGATGTTTATTTAGGTTG
GA-RM02-Spec-Rev	tttaaaatacctcgcgagtggcaacactgaGACAAAGATGAAATACATTTCAG
GA-RM03-Spec-For	atttccccgaaaagtgccacctgggtccttGTGCTCTTTCAAACTTAATG
GA-RM03-Spec-Rev	tttaaaatacctcgcgagtggcaacactgaATGTTCTAAGAAATATATATGGAAAC
GA-SNAP-fwd	aaaaggcatcgatcttgcgcgaATGGACAAAGACTGCGAAATGAAGC
GA-SNAP-rev	gccatggatctgagtccggaGGATCCGCCTGCAGGACC
M1F4-AgeIrev	AGCGAGAGCACCGGTgagtctccattacatgcaccatc
M1F4-BamHIfor	AGCGAGAGCGGATCCGTTTAACATAATGATATTCAGGTGCATCCA
M2F1-pRS413	GCATACGATATATATACATGTGTATATATGTATACCTATGAATGTCAGTAAGTATGTATACGAACAGTATGATACTGAAGATGACAAGGTAATGCATCATTCTATACGTG CATTCTGAACGAGGAGCGAGAGCACCGGTtgcagtgcctttcttcctgcctggctctgttcacttatatgccacccagaggtctgcctatgttcttgcttgtaacacagtttccctttttctctcatatggctaatatt
M2F12-AgeIFor	AGCGAGAGCACCGGTtgcagtgcctttcttcctgcc

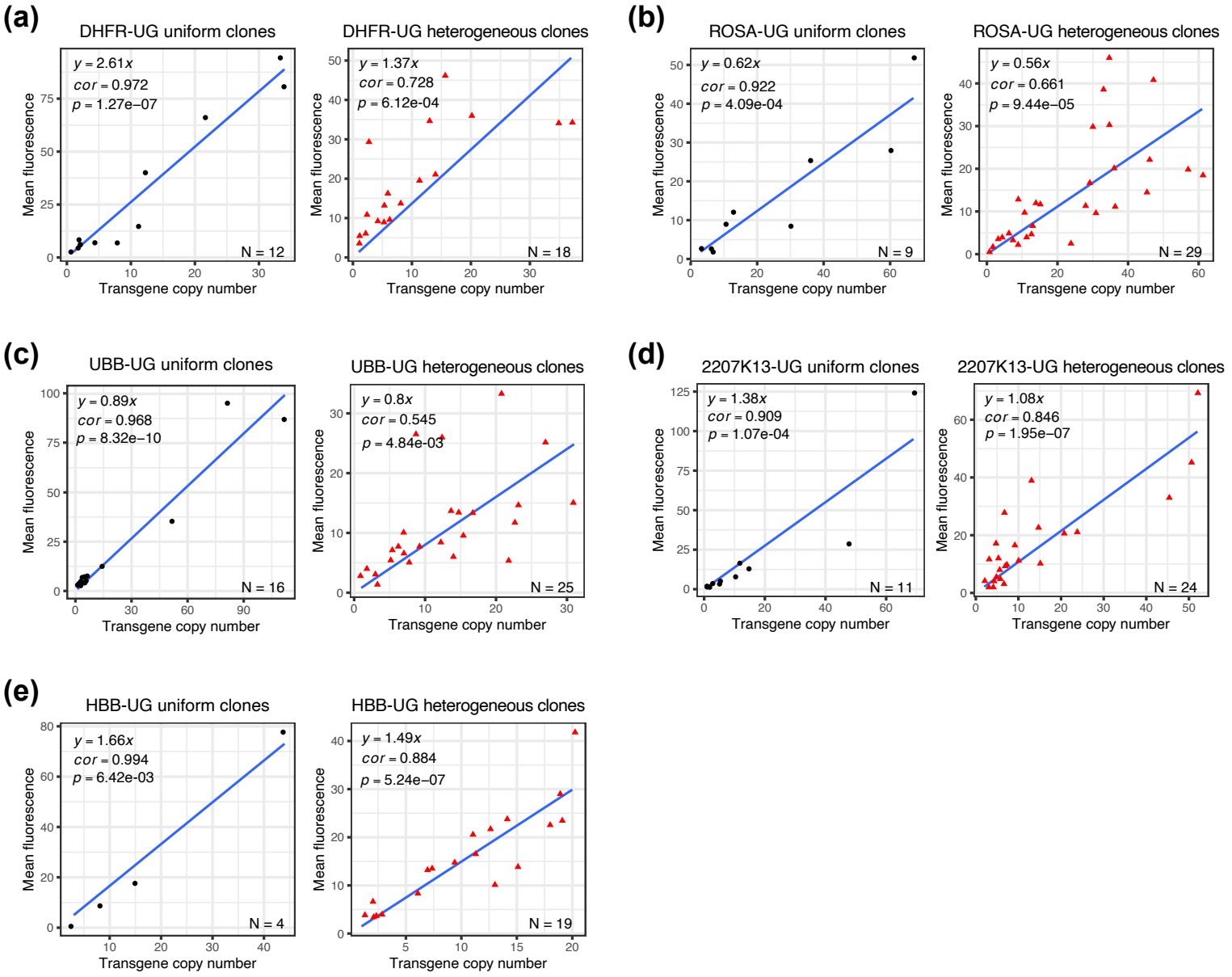
M2F12-PshRev	AGCGAGAGCGACGTCTGTCTGCTCACAAAATCAAGCCTTGAAGAC
M2F1rev	AGCGAGAGCACCGGTGGACCAGAAGCCTCTAAGACCT
M2F4-AgeIRev	AGCGAGAGCACCGGTACAGCGTGTGGCACAGAGAG
M2F4-BamHIfor	AGCGAGAGCGGATCCggttgaataaaccagcacttacagac
M2F4-pRS413	tgcacttttgttcctgtcagtgcactaggttctgtcagcattagttgcagtgttggcaaggtggaaaggcacttcctgtcctgtcaggacaaacttggcctctctgtgccacacgctgtACCGGTGCTCTCGCTCGGTAATACGGTT ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
M3-F3For	TACGCTGCGTCTCCtgacacatgatccttcctccacccagt
M3F1-AgeIFor	AGCGAGAGCACCGGTAGAGTACTTACTATCAGCCCTATCC
M3F1-pRS413	GCATACGATATATATACATGTGTATATATGTATACCTATGAATGTCAGTAAGTATGTATACGAACAGTATGATACTGAAGATGACAAGGTAATGCATCATTCTATACGTG CATTCTGAACGAGGAGCGAGAGCACCGGTagagtacttactatcagccctatccctgtctcctgttcaaggcttactgtcatttccatcgcctctaaatcccgtttagtgaatgtcagagactgtatcctgcatctga
M3F1-PshRev	AGCGAGAGCGACGTCTGTCCAAGGCGCAGATCACAGGCT
M3F3-BamHIrev	AGTCACACAGGTCCATGACATTC
M3F4-AgeIRev	AGCGAGAGCACCGGTccaaccaaggccaacaacagat
M3F4-BamHIfor	AGCGAGAGCGGATCCCTGCATGAGCTCTGGCTATGC
M3F4-PCR-Fw	ctgcatgagctctggctatgca
M3F4-pRS413	tgaagaggacagaaaggctacacacgtgtgctgtcagcttgtaaggttgtcaccctctgattgttaattagttatcaccgtacgaaaggcagaataaatctgttgttggccttggttggACCGGTGCTCTCGCTCGGTAATACGG TTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATA
M4F1-AgeIFor	AGCGAGAGCACCGGTGGGAGTCTTCTAGAAGCAAAGAG
M4F1-PCR-Rev	ttctcctaagagctgcttatctacc
M4F1-pRS413	GCATACGATATATATACATGTGTATATATGTATACCTATGAATGTCAGTAAGTATGTATACGAACAGTATGATACTGAAGATGACAAGGTAATGCATCATTCTATACGTG CATTCTGAACGAGGAGCGAGAGCACCGGTgggagtcttctagaagcaaagaggtcccttcaagagccccactgccttgggtactgattgtacgttctctatatttctgggtgagggcttgtttctatccatggcttgg
M4F1-PshRev	AGCGAGAGCGACGTCTGTCCGTGTTTCTCCTAAGAGCTGCTTA
M4F5-pRS413	cccctcacagggccaggaacatggggtccttcctctccttggttacccagctcagtgcctggggagatgcaggccaacggtgatgtcccagcatcaaaggcagcgttcactgcagcagcaACGCGTGCTCTCGCTCGGTAA TACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCAT
mCherry-H3-Magoh-Rev	atgctacgAAGCTTccttgtacagctcgtccatgc
mCherry-NheI-Magoh-For	AttagtatGCTAGCGATCCCGCCACCATGGTG
MMorCF-AgeIfor	AGCGAGAGCACCGGTtgctctacgttttaggatggagaga
newPCFAgeIrev	AGCGAGAGCACCGGTCCTTAAGATACATTGATGAGTTTGGACA
PPIACerFibFor	AGCGAGAGCACGCGTtgctctacgttttaggatggagaga
PPIACerFibRev	AGCGAGAGGCTAGCTGGCTAATAGTACACGGTTTTCCTC
PPIA-Magohfor	AGCGAGAGCCATATGtgctctacgttttaggatggagaga
PPIA-MagohRev	AGCGAGAGCGCTAGCGGCTAATAGTACACGGTTTTCCTC
PSF-AgeI-For	ATTCACCGGTAAATGGCCCGCCTGGCAT
PSF-AgeI-Rev	atgcACCGGTCAACACTCAACCCTATCTCGGTCTATTCT
R32CerLBAgeIfor	AGCGAGAGCACCGGTCCGCaagcaaaacaaatcaagagcat
Rev-M4F5-MluI	AGCGAGAGCACGCGTtgctgctgcagtgaacgctg
RevCerFIb	TGCTATTGCTTTATTTGTAACCATTATAAGCTGC
RS413-Fw	CGGTAATACGGTTATCCACAG
RS413-Rev	CCTCGTTCAGAATGACACGT
Snap-HpaI-Fib-Rev	atctactGTTAACttaTGCAGGACCCAGCCCAGG
Snap-XmaI-For	atgctgCCCGGGcACCATGGACAAAGACTGCGAAAT
SpecFor	TCAGTGTTGCCACTCGCGA
SpecRev	AAGGACCCAGGTGGCACTT
ZeoMluIFor	TATACGCGTCGCTAGCTCGAGGGTGTGGA
ZeoMluIRev	GGCACGCGTAGACATGATAAGATACATTGATGAGTTTGGACAA

UG insertion site DNA FISH probe		
BZ#369	GGT CAG TCT CCT TGC TGG AC	HBB BAC UG insertion site DNA FISH 1, 3.6 kb

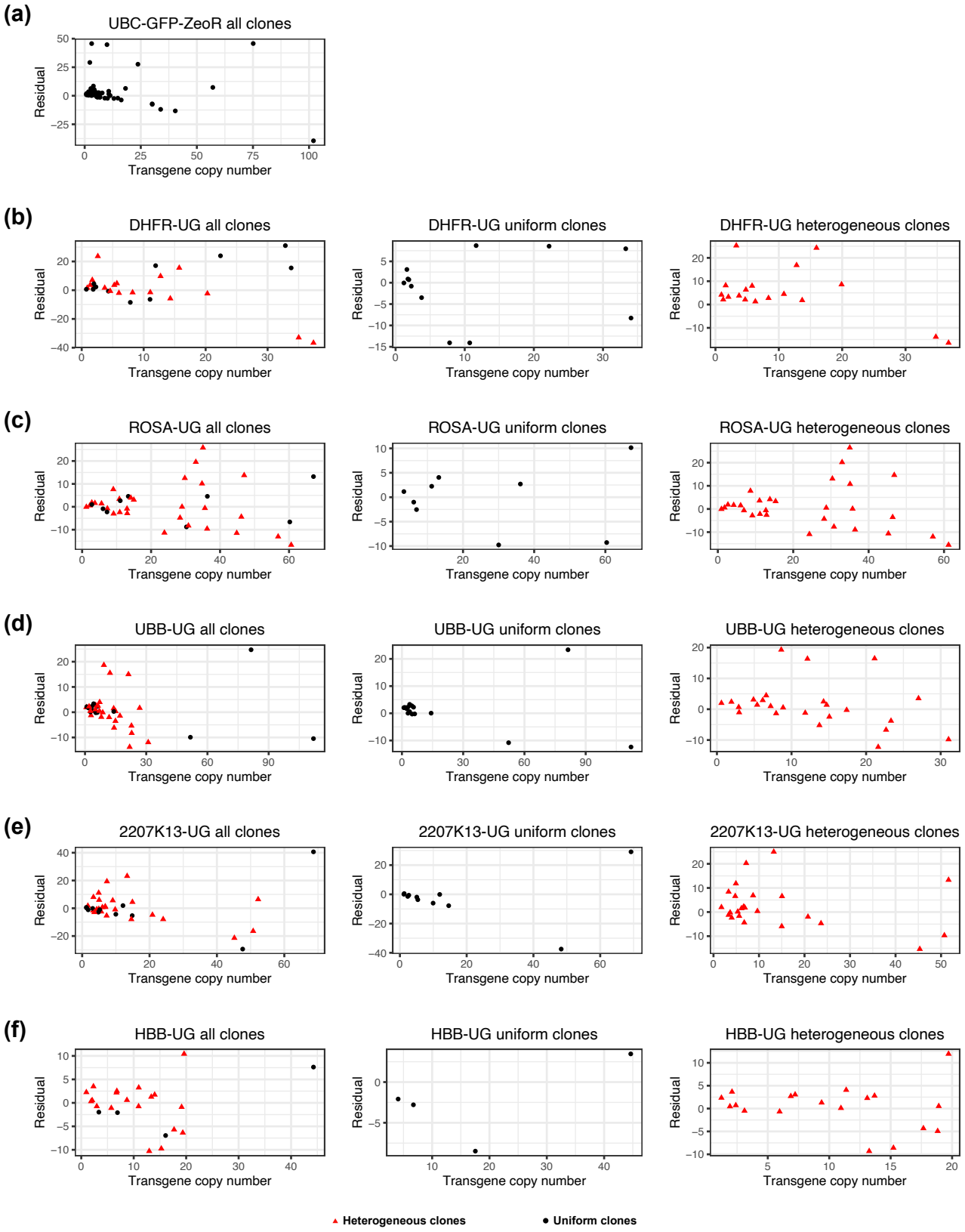
BZ#370	TTCTGAGTGAGCCTGGGTCT	HBB BAC UG insertion site DNA FISH 1, 3.6 kb
BZ#371	GGTGCTGTGTGGCTTAGAGT	HBB BAC UG insertion site DNA FISH 2, 2.96 kb
BZ#372	ACTTATTGGCTGCGTACCCT	HBB BAC UG insertion site DNA FISH 2, 2.96 kb
BZ#373	ATCGATCAATATAGACTTCCAGTGA	HBB BAC UG insertion site DNA FISH 3, 2.1 kb
BZ#374	CCCGGATCTAGGGTCTTTCA	HBB BAC UG insertion site DNA FISH 3, 2.1 kb
BZ#375	CCTGCCATGTCATGCTGAGT	HBB BAC UG insertion site DNA FISH 4, 2.6 kb
BZ#376	GCCGTAGGGAGGAGTTTCTG	HBB BAC UG insertion site DNA FISH 4, 2.6 kb
BZ #210	TTAAGCGGCCGCATGACAGCAGAGATCCAGTTT	UBC-GFP-ZeoR reporter gene, 2.7 kb
BZ #351	GCTCCAATTCCATACCACATTTG	UBC-GFP-ZeoR reporter gene, 2.7 kb
BZ#379	GGTTGTCTCCACCTCACTGG	DHFR BAC UG insertion site DNA FISH 2, 2.7 kb
BZ#380	GGTGAGAGCTGACTGCACTT	DHFR BAC UG insertion site DNA FISH 2, 2.7 kb
BZ#381	CCTGAGTGAAATGGCCTGGT	DHFR BAC UG insertion site DNA FISH 3, 2.7 kb
BZ#382	GCCACGGGTACTGACTGAAA	DHFR BAC UG insertion site DNA FISH 3, 2.7 kb
BZ#383	CCCACCCATCTGTGTCTGTC	DHFR BAC UG insertion site DNA FISH 4, 3.1 kb
BZ#384	AGACCCATAGCGAATGCCTG	DHFR BAC UG insertion site DNA FISH 4, 3.1 kb

Supplementary Table S2. Sequence of a synthetic DNA fragment "RCS" containing multiple rare restriction sites

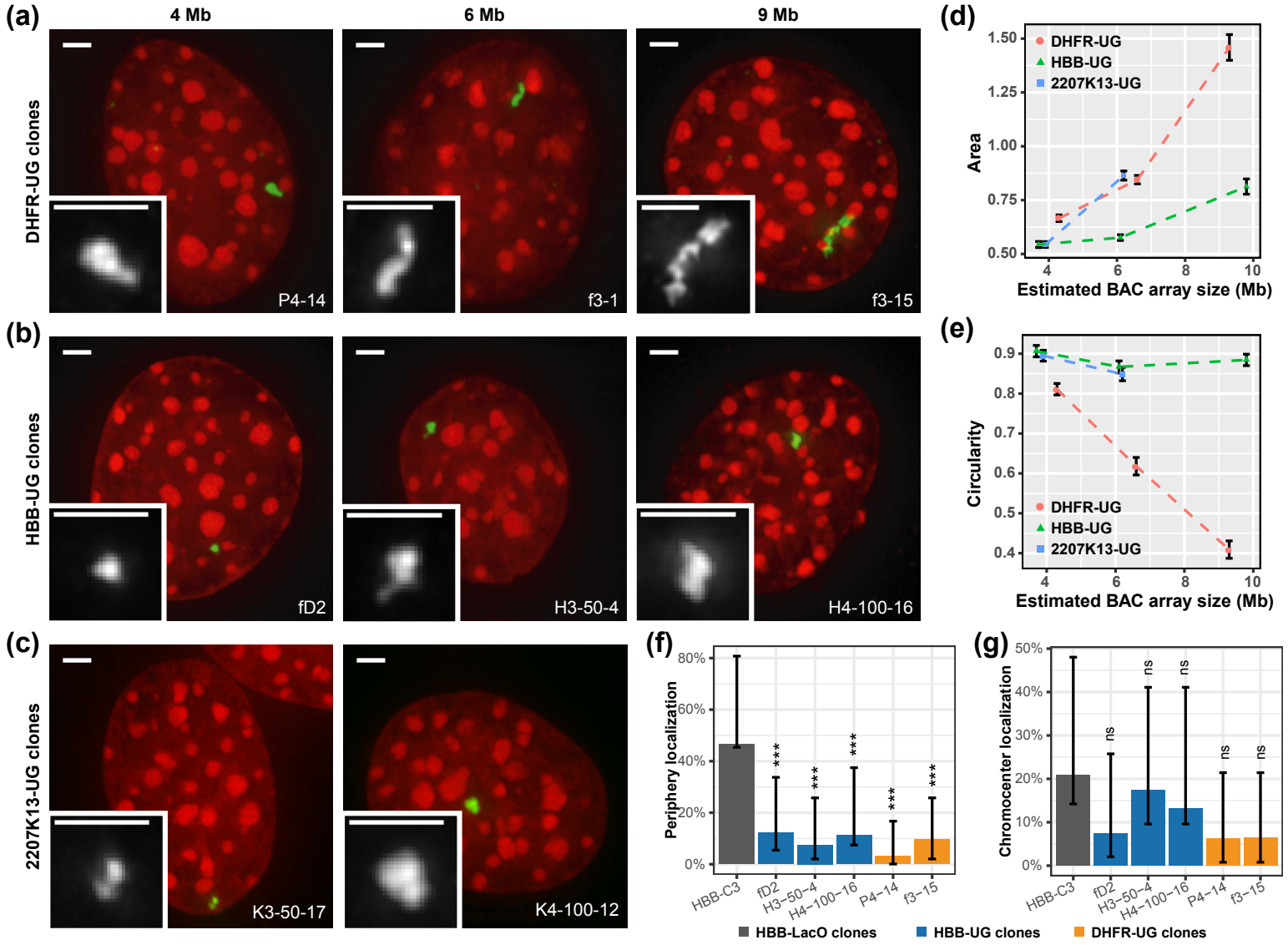
Name	Sequence	Notes
RCS	5′ GGCCGCGGCGCGCCTTAATTAACCGGTG 3′ 3′ CGCCGCGCGGAATTAATTGGCCACGATC 5′	NotI overhang NheI overhang



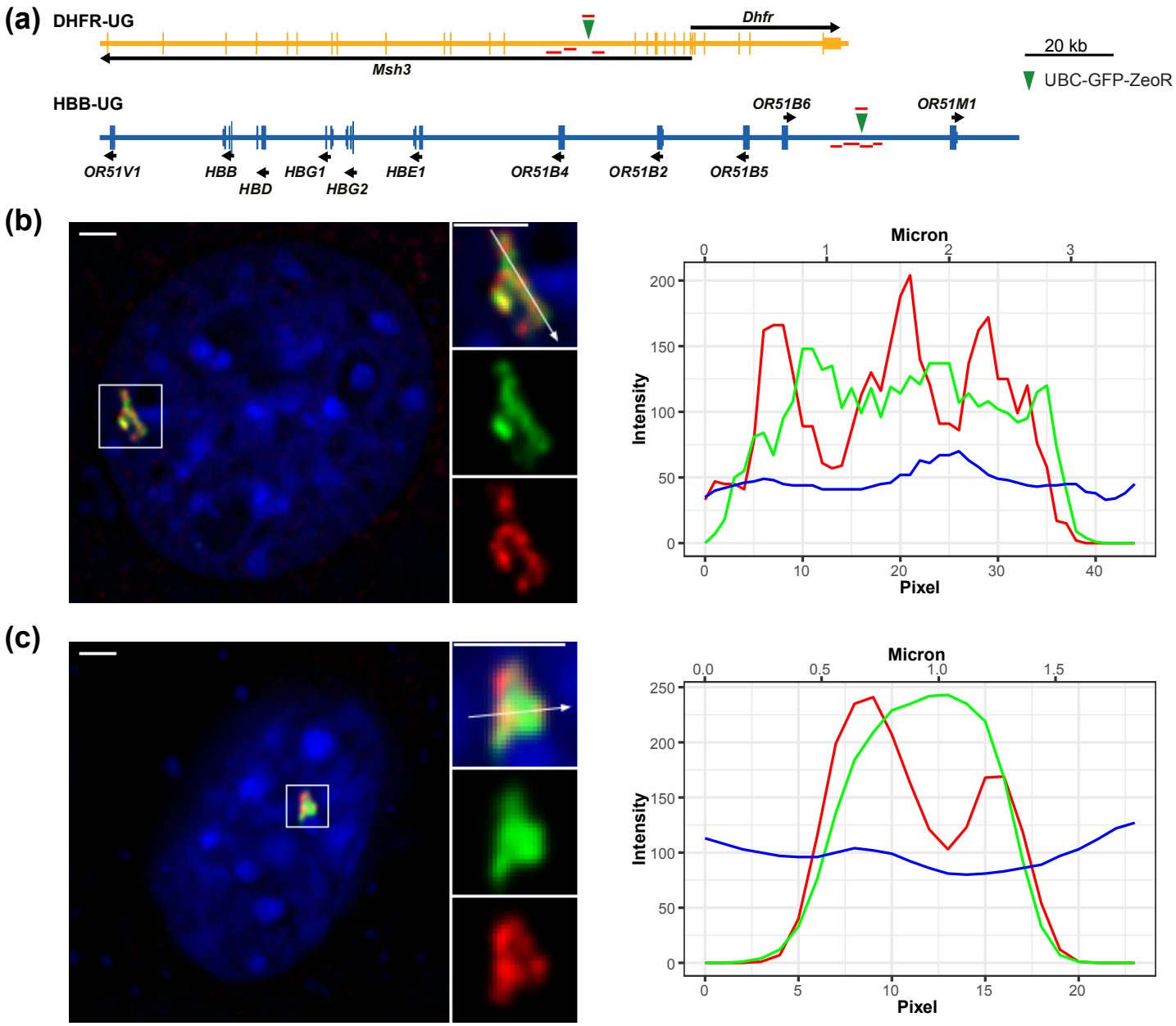
Supplementary Figure S1. Linear correlation of transgene expression level vs copy number among uniform clones or heterogeneous clones. (a-e) Scatter plots of mean normalized cellular GFP fluorescence (y-axis) vs reporter gene copy number (x-axis) for clonal populations transfected with different BAC scaffolds. Linear regression fits (blue lines, y-intercepts set to 0) are shown with corresponding Pearson correlation coefficients (cor), p-values of the correlation coefficients (p), and equations. Bottom right of plots: Number of clones analyzed. Left panels- clones showing uniform GFP expression; Right panels- clones showing heterogeneous GFP expression.



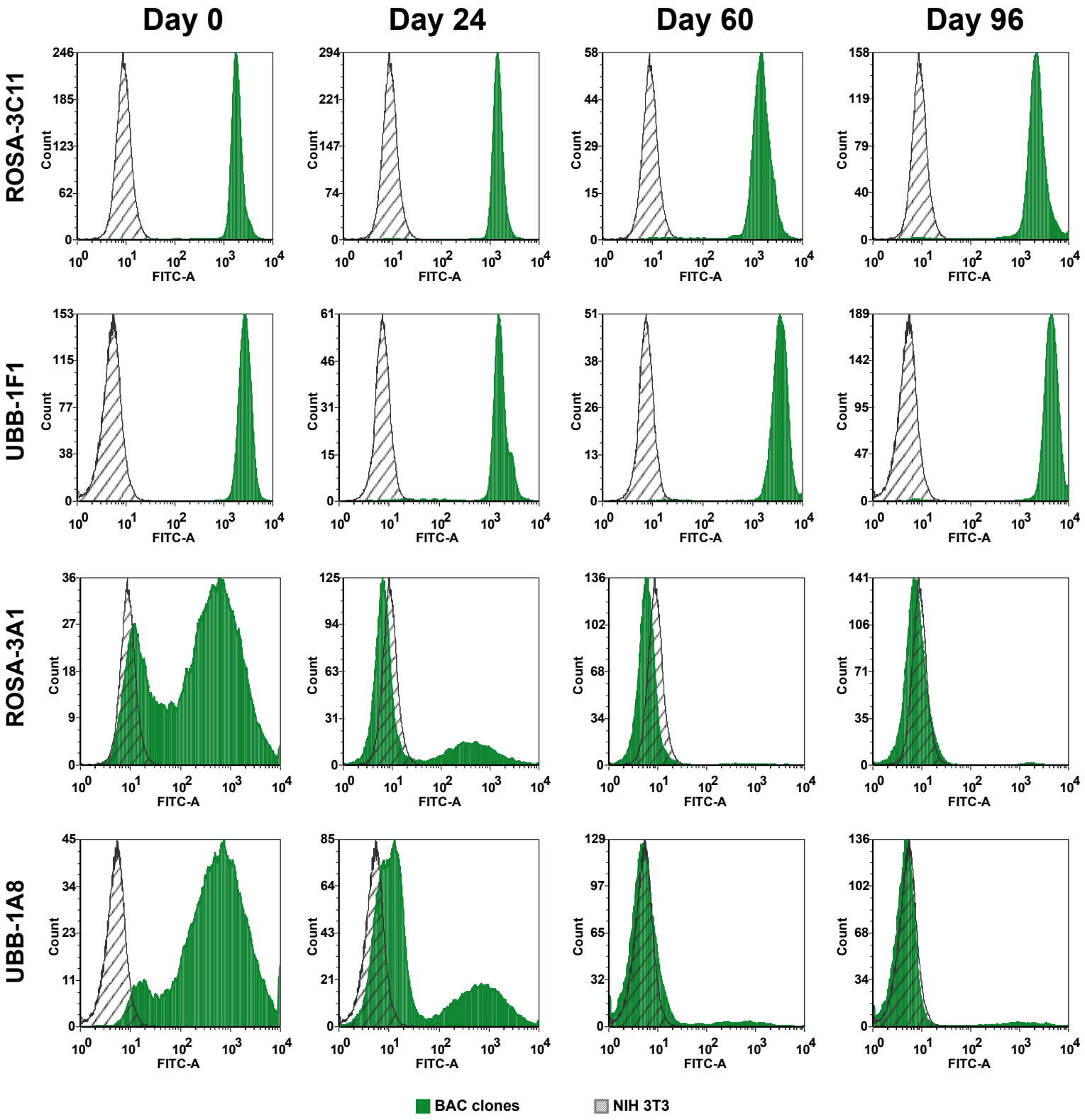
Supplementary Figure S2. Residual plots of linear regression fits for mean normalized cellular GFP fluorescence vs reporter gene copy number. (a-f) Scatter plots of residuals from linear regression fits (y-axis) vs reporter gene copy number (x-axis) for clonal populations transfected with the UBC-GFP-ZeoR cassette alone (a), or with different BAC scaffolds carrying the UBC-GFP-ZeoR reporter gene (b-f). Left panels: linear regression among clones showing uniform or heterogeneous GFP expression; Middle panels- linear regression among clones showing uniform GFP expression; Right panels- linear regression among clones showing heterogeneous GFP expression; Red triangles- heterogeneous clones; Black circles- uniform clones.



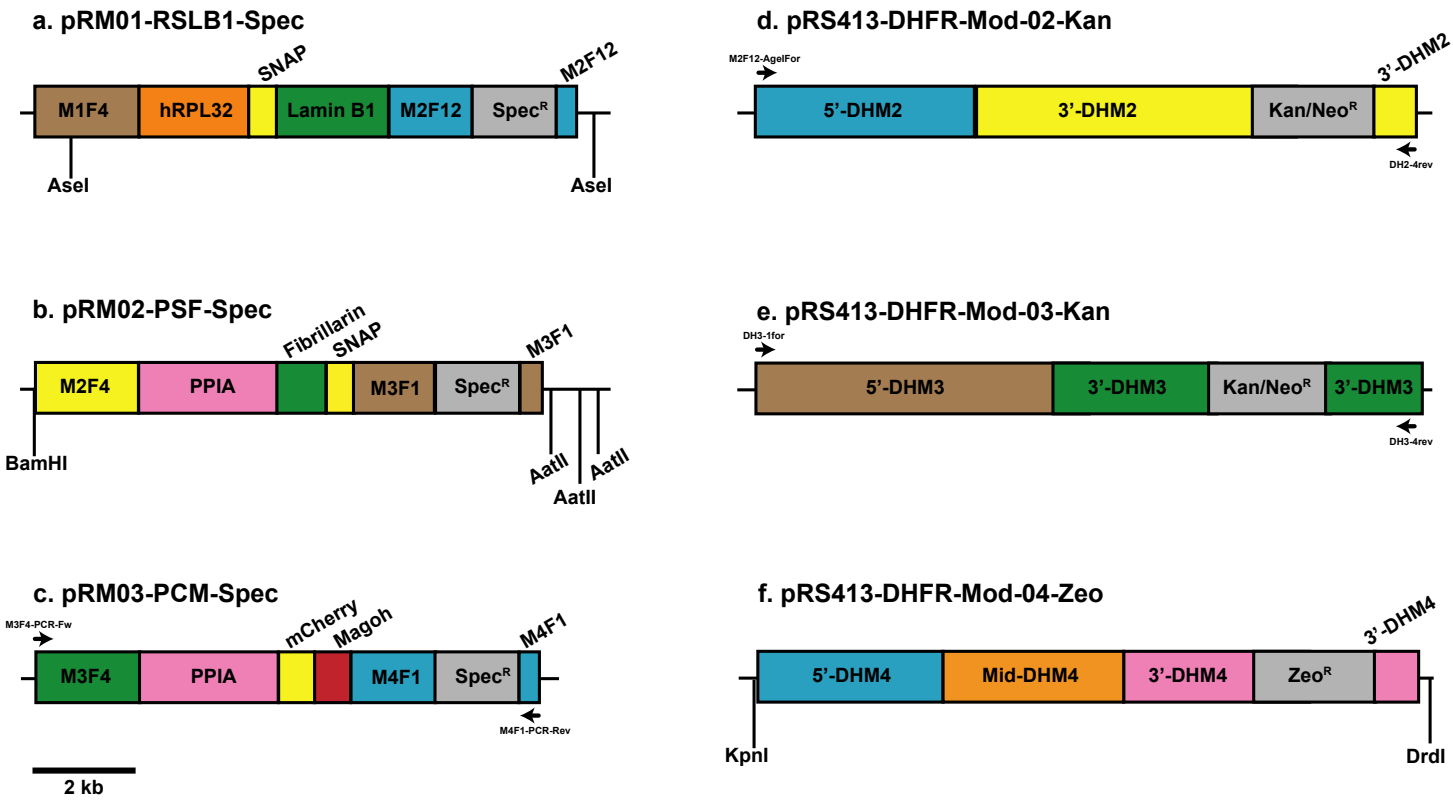
Supplementary Figure S3. Repressive BAC transgene arrays with UBC-GFP-ZeoR reporter gene form highly condensed chromatin but have altered nuclear localization. (a-c) 3D DNA FISH with BAC probes showing distinctive chromatin conformation formed by the DHFR-UG, HBB-UG, and 2207K13-UG BAC arrays. Maximum intensity z-projection images of three DHFR-UG clones (a), three HBB-UG clones (b), and two 2207K13-UG clones (c) are shown. Insets are enlarged FISH signals. Clone names are indicated at the bottom right of each image. Estimated sizes of the BAC transgene arrays are shown on the top. Red- DNA DAPI staining; Green- BAC FISH signal; Scale bars = 2 µm. (d-e) Quantitation of FISH signals. Median area (d, y-axis, unit = µm2) or median circularity (e, y-axis) of the z-projected FISH signals are plotted against qPCR estimated BAC transgene array size (Mb, x-axis). Error bars represent standard error; Red- DHFR-UG clones; Green- HBB-UG clones; Blue- 2207K13-UG clones; Number of analyzed images for each clone = 56 ~81. (f-g) Percentage localization of different BAC transgenes visualized by 3D DNA FISH at nuclear periphery (f) and chromocenter (g). x-axis- clone names; Gray- HBB-C3 clone carrying integrated array of HBB BAC without UBC-GFP-Zeo cassette inserted; Blue- clones with integrated HBB-UG BAC arrays; Yellow- clones with integrated DHFR-UG BAC arrays; Number of images analyzed- 43 for HBB-C3, 41 for fD2, 40 for H3-50-4, 53 for H4-100-16, 32 for P4-14, 31 for f3-15; Error bars represent 95% confidence intervals based on binomial distribution; p-values are calculated using two-sided Fisher's Exact Test comparing individual clones to clone HBB-C3 (***: p-value < 0.001; ns: p-value > 0.1).



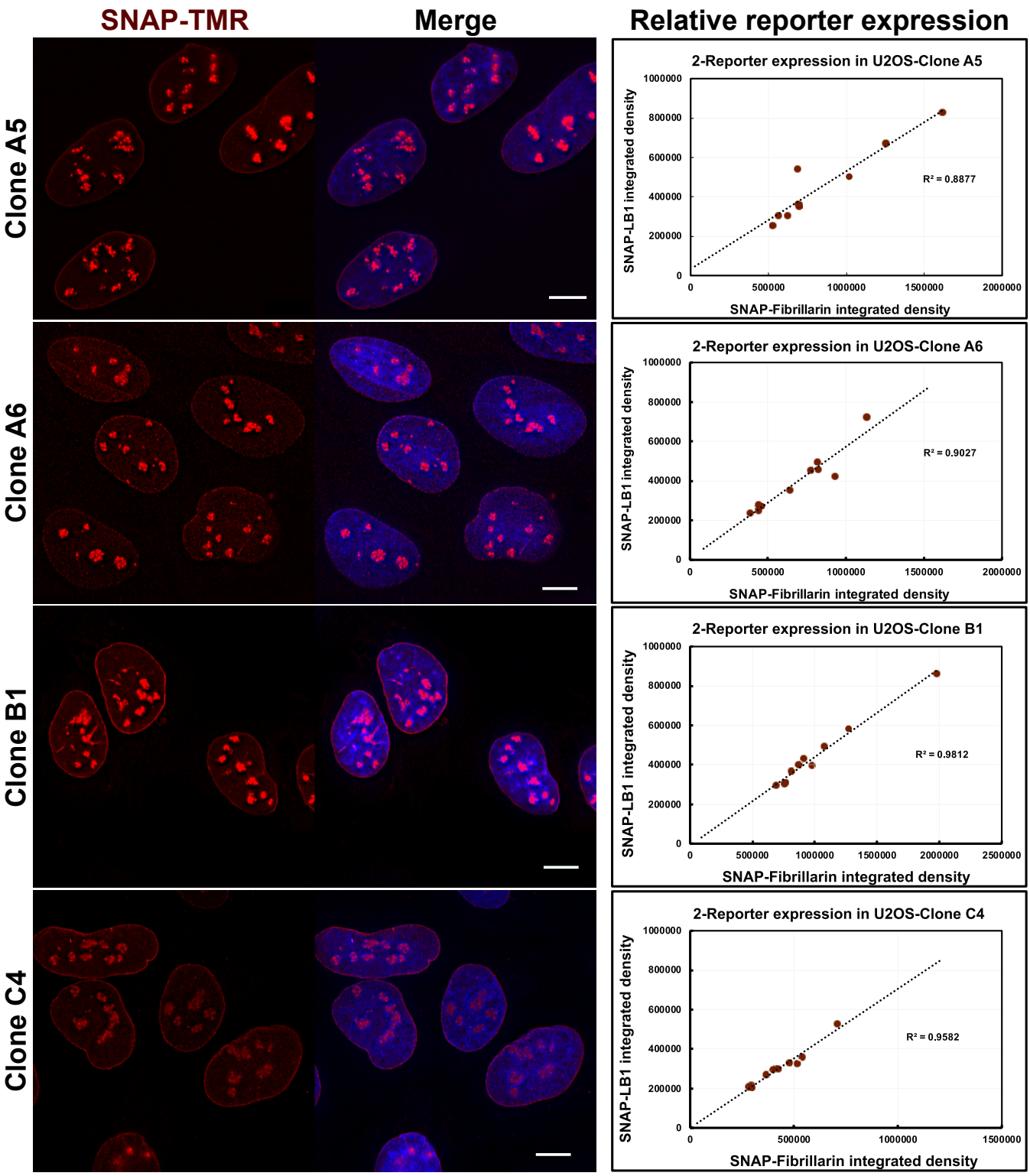
Supplementary Figure S4. Differential localization of the UBC-GFP-ZeoR reporter gene sequence in DHFR vs HBB BAC transgene territories. (a) DNA FISH probe regions (red lines) for visualizing the UBC-GFP-ZeoR reporter gene and ~12-15 kb regions around its insertion sites on the DHFR and HBB BAC. Longer vertical bars- exons; shorter vertical bars- UTRs; black arrows or arrowheads- direction of transcription; green arrow heads- UBC-GFP-ZeoR insertion site. (b-c) Two-color DNA FISH with BAC (green) and UBC-GFP-ZeoR insertion region (red) probes over DHFR-UG clone f3-15 (b) and HBB-UG clone H4-100-16 (c). Sections with FISH signals in focus are shown. Insets are enlarged regions within the white squares. Intensity line scans along arrows are plotted in the right panels. Blue- DNA DAPI staining; Scale bars = 2 μm.



Supplementary Figure S5. GFP fluorescence histogram of representative "uniform" and "heterogeneous" expressing NIH 3T3 clones at day 0, 24, 60 and 96 without selection obtained by flow-cytometry. Gray- autofluorescence of untransfected cells; Green- GFP fluorescence of the indicated clones; x-axis- fluorescence; y-axis- cell number.



Supplementary Figure S6. Maps of Reporter and DHFR modules used for BAC-MAGIC. (a-c) reporter expression cassettes subcloned in the respective reporter recipient modules harboring SpecR selection marker (gray). (d-e) Schematics of the intervening DHFR modules harboring Kan/NeoR or ZeoR selection markers (gray). (a-e) Longer vertical bars represent the indicated restriction endonucleases used to generate recombineering fragments and arrows show the binding sites of primers used for amplification of recombineering fragments. See Methods for details of terminal regions corresponding to DHFR BAC homology regions. Scale bar = 2 kb.



Supplementary Figure S7. Expression of dual-reporter BAC in U2OS cells. Representative images (maximum intensity projections of 3-4 optical sections) from the four independent cell clones (Clone A5, A6, B1 and C4) showing expression of the two reporter genes. Nuclear lamina is labeled with SNAP-tagged Lamin B1 (red), nucleoli with SNAP-tagged Fibrillarin (red), and nucleus is counterstained with DAPI (blue). Scale bars = 5 µm.