Correction notice
Nat. Methods **5**, 409–415 (2008).

BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals

Ina Poser, Mihail Sarov, James R A Hutchins, Jean-Karim Hériché, Yusuke Toyoda, Andrei Pozniakovsky, Daniela Weigl, Anja Nitzsche, Björn Hegemann, Alexander W Bird, Laurence Pelletier, Ralf Kittler, Sujun Hua, Ronald Naumann, Martina Augsburg, Martina M Sykora, Helmut Hofemeister, Youming Zhang, Kim Nasmyth, Kevin P White, Steffen Dietzel, Karl Mechtler, Richard Durbin, A Francis Stewart, Jan-Michael Peters, Frank Buchholz & Anthony A Hyman

In the version of this supplementary file originally posted online, the supplementary figure legends were missing. The error has been corrected in this file as of 30 July 2008.

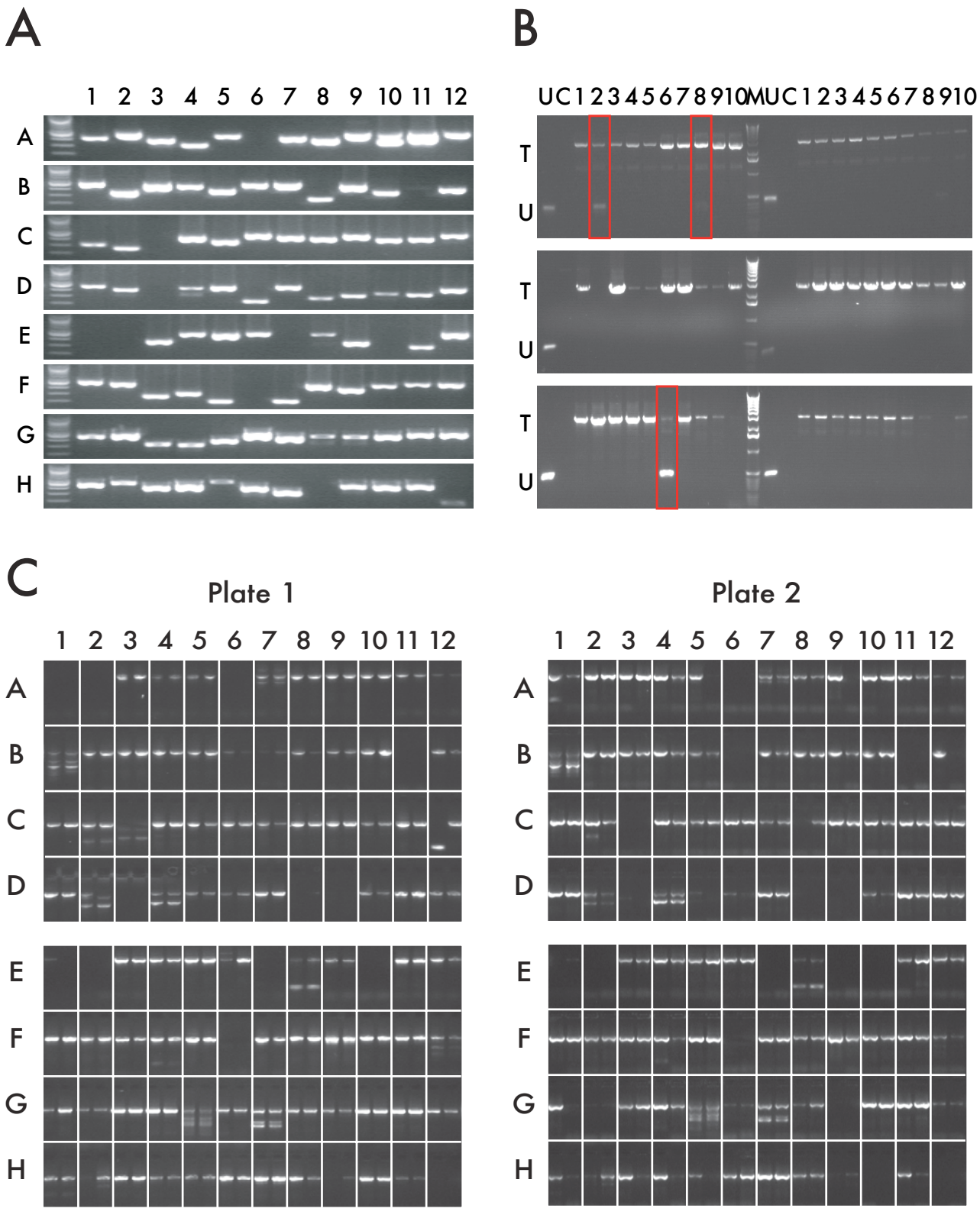
BAC TransgeneOmics: a high-throughput method for exploration of protein function in mammals

Ina Poser, Mihail Sarov, James R A Hutchins, Jean-Karim Hériché, Yusuke Toyoda, Andrei Pozniakovsky, Daniela Weigl, Anja Nitzsche, Björn Hegemann, Alex Bird, Laurence Pelletier, Ralf Kittler, Sujun Hua, Ronald Naumann, Martina Augsburg, Martina M. Sykora, Helmut Hofemeister, Youming Zhang, Kim Nasmyth, Kevin P White, Steffen Dietzel, Karl Mechtler, Richard Durbin, A Francis Stewart, Jan-Michael Peters, Frank Buchholz & Anthony A Hyman

Supplementary figures and text:

Note: Supplementary Table 3 is available on the Nature Methods website.

Supplementary Figure 1: 96 well recombineering efficiency and fidelity



Supplementary Figure 1: 96 well recombineering efficiency and fidelity

A. Clone selection efficiency. BAC clones for 96 genes were selected with the BACFinder tool and the presence of the targeted tag insertion point was evaluated by PCR. The majority of the clones 88/96 gave a band of the expected size. The PCR failure in the other 8 cases is most likely due to rearranged or wrongly assigned clones.

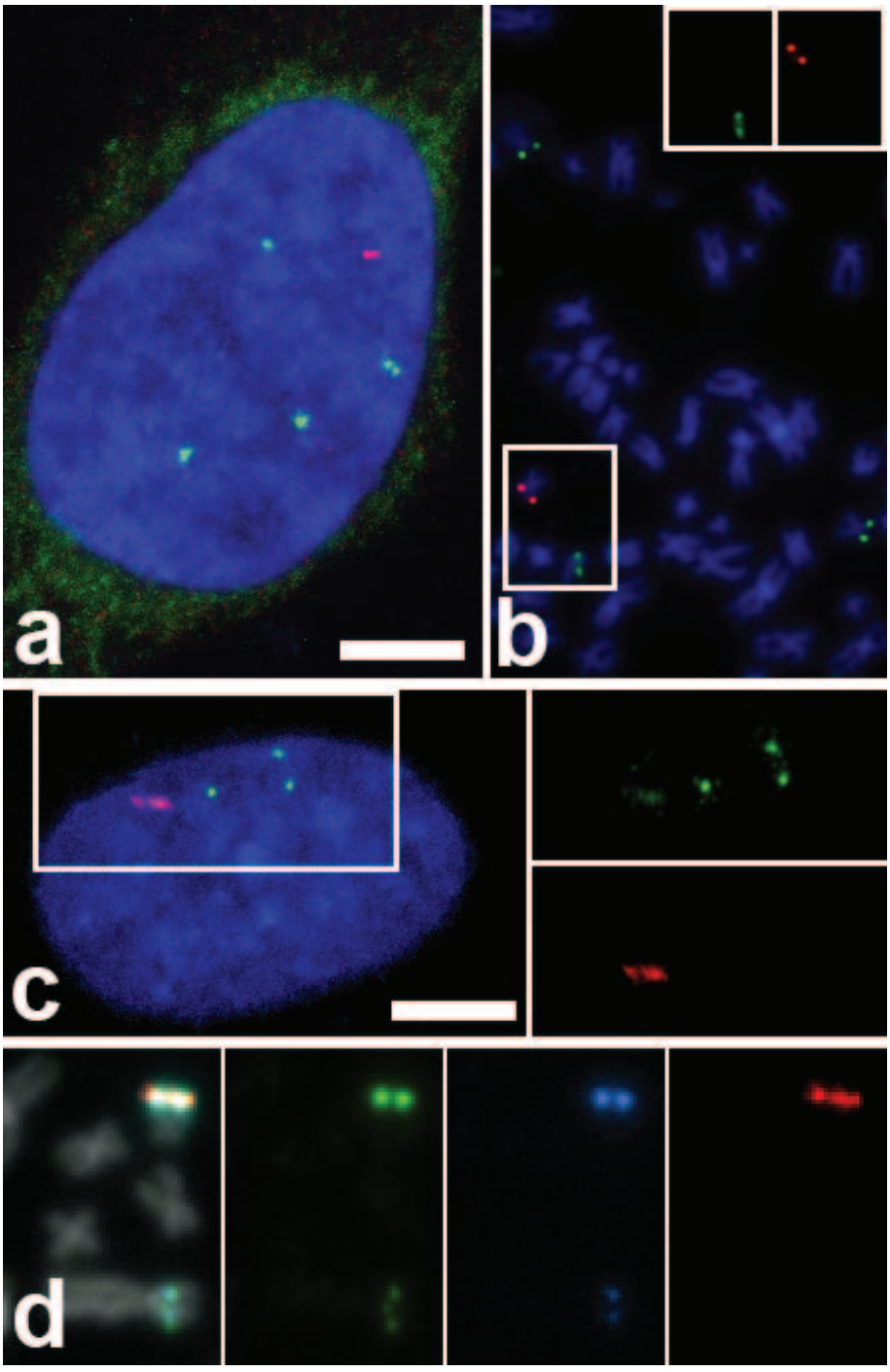
B. Recombineering fidelity:

To validate the fidelity of the method 6 clone pools were plated on selective agar and 10 individual clones of each pool were checked by PCR through the insertion point. The untagged BACs (U) resulted in small PCR bands. No product was obtained with a control unrelated BAC (C). For all tagged clones (T) the PCR product is approximately 2.5 kb larger than the untagged BAC due to the insertion of the cassette sequence. Three of the clones (outlined in red) showed carryover of the unmodified BAC clone.

C.Recombineering efficiency:

The 96 genes form panel A were tagged in 2 parallel experiments. The genes that gave no wild type band also failed to grow in selection. After plating each well on selective agar two individual clones from were checked by PCR through the tag insertion point. Of the 88 genes that grew in selection 78 (89%) of the genes a correct PCR product was obtained, which was approximately 2.5 kb larger than the that of an unmodified BAC. Further 7 of the genes gave the expected band but were excluded from further processing because of additional non-specific bands. Finally 3 of the genes that grew in selection were PCR negative in both repetitions.

Supplementary Figure 2:
Localization of BAC transgenes and homologous endogenous sequences by FISH



Supplementary Figure 2: Localization of BAC transgenes and homologous endogenous sequences by FISH

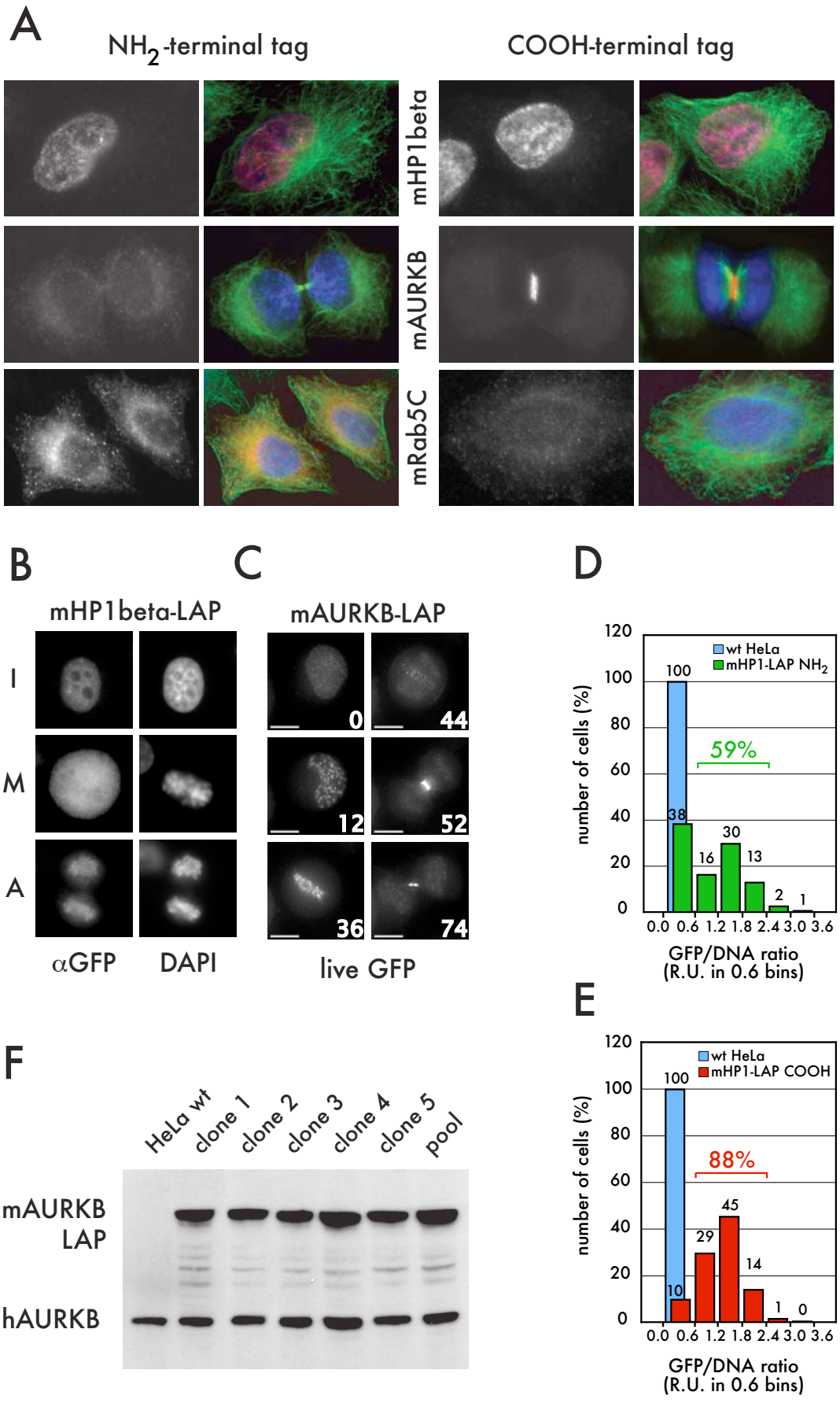
A. AURKB cell line interphase nucleus with four FISH signals from the human locus (green) and an independent FISH signal from the transgenic mouse BAC (red). Projection of confocal sections, counterstained DNA is in blue, scale bar: 5 µm. As most tumor cell lines, HeLa cells are highly aneuploid. They contain most chromosomes more than twice.

B. Part of a metaphase spread of the AURKB cell line showing three of the four chromosomes with endogenous signals, colors as in A. The boxed area at the bottom is shown again at the top right as single color channels. Note that there is no cross-talk or cross-hybridization between green and red channels.

C. KIAA1387 cell line interphase nucleus with three FISH signals from human endogenous regions (green) and a signal for the transgenic mouse BAC (red). The boxed area is shown as single color channels to the right. Note that a weaker signal from human sequences is also present at the site of the transgene. Projection of confocal sections, counterstained DNA in blue, scale bar: 5 µm.

D. Part of a metaphase spread of the KIAA1387 cell line as overlay (left) and single color channels showing the transgene integration site (top) and one endogenous human sequence (bottom). In this FISH experiment, the BACs for the mouse sequence (red), the homologous human sequence (green) and an additional, neighboring human BAC (blue) were cohybridized. The colocalization of the homologous mouse and human sequence as well as a neighboring human sequence and the large signal size suggests a complex integration event.

Supplementary Figure 3: LAP cassette evaluation



Supplementary figure 3: LAP cassette evaluation

A. NH_2 and COOH tagging cassettes performance.

HeLa cell lines stably transfected with mHP1beta, mAURKB and mRab5C tagged at either NH_2- or COOH-termini were fixed and stained for the tagged transgene (red), alpha-tubulin (green) and DNA (blue). The grayscale image shows the red channel only. mHP1beta localizes correctly to the nucleus when tagged at its NH_2- and COOH-terminus. For mAURKB, only the COOH-terminal tagged transgene correctly localized to the cleavage furrow during cytokinesis. mRab5C only localized correctly to early endosomes when tagged at its NH_2-terminus.

B,C. Cell cycle dynamics of mHP1beta and mAURKB.

B. HP1beta localizes to chromatin throughout interphase (I) and is rapidly lost at metaphase (M) during mitosis but "comes back" to the chromosomes later in anaphase (A). Representative images of fixed cells stained with an anti-GFP antibody (αGFP) for the LAP tag and DAPI for DNA (DAPI) are shown for each phase described.

C. Single frames (0 – 74 min) of a time-lapse movie showing GFP fluorescence of HeLa cells, stably expressing mAURKB-LAP at endogenous levels. Scale bars, 10µm.

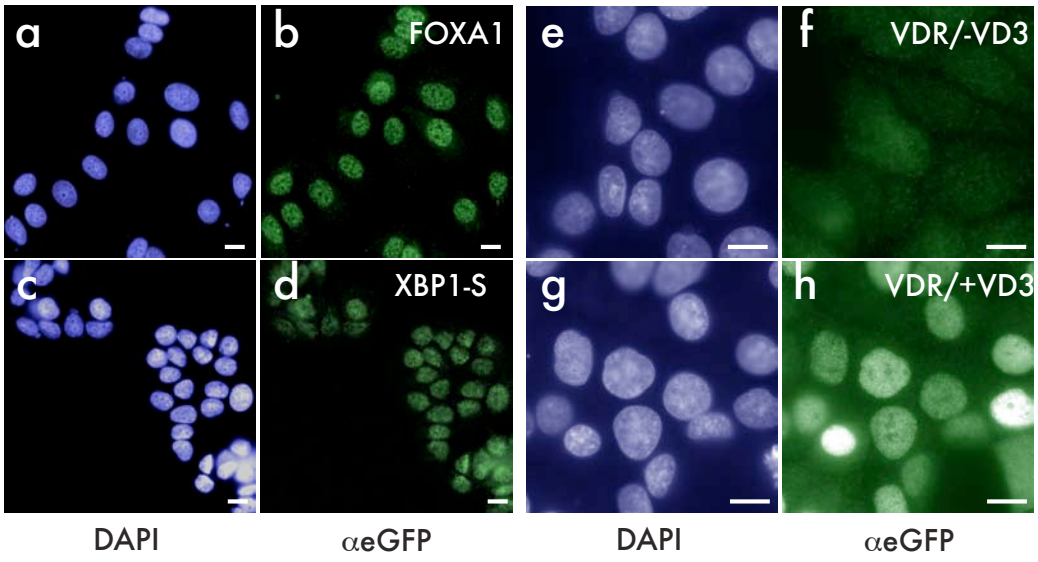
D,E. EGFP intensity distribution of mHP1beta-LAP pools.

GFP fluorescence of HeLa cell pools stably expressing mHP1beta-LAP tagged on its NH_2- (**D**) or COOH- (**E**) terminus were analyzed by FACS. Fluorescence intensities were normalized to DNA content (DRAQ5) yielding a GFP/DRAQ5 intensity ratio. It is notable, that the entire population of untransfected wild type (wt) HeLa cells resulted in a GFP/DRAQ5 ratio between 0.0 and 0.6 (blue column).

F. Stable BAC cell lines and cell pools express the tagged transgene at comparable levels to the endogenous protein.

The Western Blot shows protein levels of tagged mAURKB and endogenous hAURKB of 5 clonal HeLa cell lines and the unsorted HeLa cell pool they have been derived from. mAURKB was detected with an anti-AURKB antibody (Becton Dickinson) recognizing mouse and human AURKB likewise.

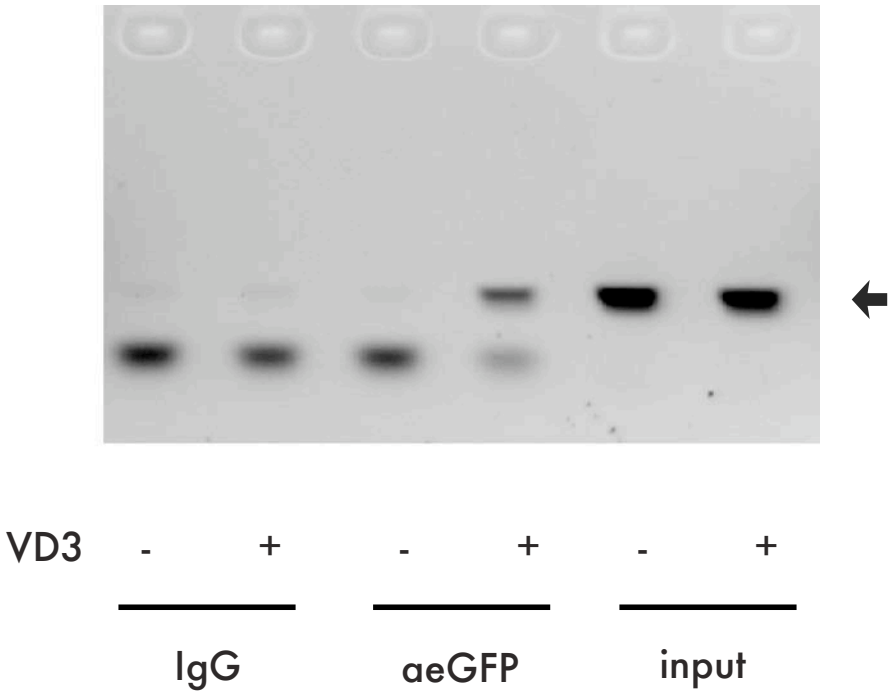
Supplementary Figure 4: Correct localization of LAP tagged transcription factors.



Supplementary Figure 4: Correct localization of LAP tagged transcription factors.

a-h. Localization of LAP-tagged transcription factors. a-d, MCF7 cells expressing LAP-tagged FOXA1, XBP1-S were fixed and immunostained with an antibody against eGFP (**b, d**) and DAPI (**a,c**). **e-h**, Ligand-dependent localization of tagged VDR. MCF7 cells expressing LAP-tagged VDR3 were grown in medium with charcoal-stripped serum for 48 hours. The ligand-depleted cells were then grown for two hours in the presence of 20 nM VD3 (VD3+) or vehicle (VD3-), fixed and immunostained with an antibody against eGFP (**f, h**) and DAPI (**e,g**). Bars are 10 μm.

Supplementary Figure 5:
ChIP-PCR analysis of ligand dependent binding of a known VDR target.



Supplementary Figure 5: ChIP-PCR analysis of ligand dependent binding of a known VDR target.

MCF7 cells expressing GFP-tagged VDR (Vitamin D receptor) were incubated for one hour with VDR ligand VD3 (20 nM) (+) or vehicle (-). ChIP was performed with an antibody against eGFP and IgG. The ChIP extracts and input were used for PCR of a fragment representing a known VDR binding site in the promoter region of the gene *PIT1* (indicated by the arrow).

Supplementary Table 1:
Western blot analysis of BAC transgenic cell lines

lane	gene name	species	human Ensembl ID	mouse ortholog	BAC ID	tagging site	predicted protein size (kDa)
1	Hela wildtype						
2	NRIP3	mouse	ENSG00000175352	ENSMUSG00000034825	RP24-392O24	C-terminus	27
3	FLT3	human	ENSG00000122025		RP11-136G6	C-terminus	113
4	AKAP12	mouse	ENSG00000131016	ENSMUSG00000038587	RP23-257K20	C-terminus	19, 182, 192
5	Bles03	mouse	ENSG00000175573	ENSMUSG00000047423	RP24-271A1	C-terminus	27
6	RGR	mouse	ENSG00000148604	ENSMUSG00000021804	RP24-292M6	C-terminus	13, 28, 32
7	HP1-alpha	mouse	ENSG00000094916	ENSMUSG00000009575	RP23-307M22	C-terminus	22
8	CNOT3	mouse	ENSG00000088038	ENSMUSG00000035632	RP23-146B24	C-terminus	59, 82
9	UBE1	mouse	ENSG00000130985	ENSMUSG00000001924	RP24-99E9	C-terminus	31, 57, 118
10	CDC2	mouse	ENSG00000170312	ENSMUSG00000019942	RP23-421E3	C-terminus	27, 34
11	DNCH1	mouse	ENSG00000197102	ENSMUSG00000018707	RP23-60K23	C-terminus	532
12	DNCI2	mouse	ENSG00000077380	ENSMUSG00000027012	RP24-149H16	C-terminus	68, 72
13	DNCL1	mouse	ENSG00000088986	ENSMUSG00000009013	RP23-87P14	C-terminus	10
14	CDC16	mouse	ENSG00000130177	ENSMUSG00000038416	RP24-298H15	C-terminus	14, 22, 55, 60, 66, 71, 72
15	MIS12	mouse	ENSG00000167842	ENSMUSG00000040599	RP23-372F19	C-terminus	18, 24
16	MIS12	human	ENSG00000167842		CTD-2300O9	C-terminus	18, 24
17	LAP1B	mouse	ENSG00000143337	ENSMUSG00000026466	RP24-321N12	C-terminus	43, 47, 66, 76
18	INCENP	mouse	ENSG00000149503	ENSMUSG00000024660	RP23-141I14	C-terminus	105
19	AurB	mouse	ENSG00000178999	ENSMUSG00000020897	RP23-23C3	C-terminus	39
20	Rab5C	mouse	ENSG00000108774	ENSMUSG00000019173	RP23-258L5	N-terminus	23
21	BUB1	mouse	ENSG00000169679	ENSMUSG00000027379	RP23-160E11	C-terminus	122
22	BUB1B	mouse	ENSG00000156970	ENSMUSG00000040084	RP24-153E22	C-terminus	119
23	CDC20	mouse	ENSG00000117399	ENSMUSG00000006398	RP23-118J14	C-terminus	55
24	ACTR1A	mouse	ENSG00000138107	ENSMUSG00000025228	RP23-196G7	C-terminus	42
25	LIP8	mouse	ENSG00000170037	ENSMUSG00000032782	RP23-334K19	C-terminus	13, 52, 101, 104
26	WDR8	mouse	ENSG00000116213	ENSMUSG00000029029	RP23-156P1	C-terminus	25, 44, 49, 51
27	PPP2R1A	mouse	ENSG00000105568	ENSMUSG00000007564	RP24-239A7	C-terminus	65
28	TBL1X	mouse	ENSG00000101849	ENSMUSP00000085549	RP23-235B20	C-terminus	57, 62
29	HP1-beta	mouse	ENSG00000108468	ENSMUSG00000018666	RP24-288H19	C-terminus	21
30	PTTG1	human	ENSG00000164611		RP11-729E7	C-terminus	22
31	WAPAL	human	ENSG00000062650		RP11-383M14	C-terminus	45, 133, 137, 140, 142
32	WAPAL	mouse	ENSG00000062650	ENSMUSG00000041408	RP23-478G5	C-terminus	45, 133, 137, 140, 142
33	CDC27	human	ENSG00000004897		RP11-671L22	C-terminus	92
34	PLK1	human	ENSG00000166851		CTD-2201P20	C-terminus	27, 68
35	WDR67	mouse	ENSG00000156787	ENSMUSG00000022364	RP24-273K19	C-terminus	17, 75, 116, 124
36	ATXN3	mouse	ENSG00000066427	ENSMUSG00000021189	RP24-305B15	C-terminus	35, 41, 43
37	PDHX	mouse	ENSG00000110435	ENSMUSG00000010914	RP23-16F14	C-terminus	54
38	PSD6	mouse	ENSG00000163636	ENSMUSG00000021737	RP23-178L19	C-terminus	45
39	RAD51	mouse	ENSG00000051180	ENSMUSG00000027323	RP23-360K10	C-terminus	26, 37
40	MYST1	mouse	ENSG00000103510	ENSMUSG00000030801	RP24-306M15	C-terminus	52, 53
41	HSD17B7	mouse	ENSG00000132196	ENSMUSG00000026675	RP23-447P5	C-terminus	9, 12, 34, 37, 38
42	BUB1	human	ENSG00000169679		RP11-625K21	C-terminus	122
43	TTK	human	ENSG00000112742		RP11-689M7	C-terminus	95, 97
44	KNTC2	human	ENSG00000080986		RP11-265P7	C-terminus	74
45	ZWINT	human	ENSG00000122952		RP11-1144M12	C-terminus	14, 18, 25, 31
46	TMEM29	human	ENSG00000179304		CTD-2009K1	C-terminus	24
47	SA2	mouse	ENSG00000101972	ENSMUSG00000025862	RP23-221C5	C-terminus	141, 146
48	SGOL1	mouse	ENSG00000129810	ENSMUSG00000023940	RP24-185F17	C-terminus	25, 33, 64
49	TUBG1	mouse	ENSG00000131462	ENSMUSG00000035198	CT7-399F17	C-terminus	51
50	RPL35	mouse	ENSG00000136942	ENSMUSG00000039928	RP23-428G9	C-terminus	11, 14
51	PCNA	mouse	ENSG00000132646	ENSMUSG00000027342	RP23-428F16	C-terminus	29
52	RANBP2	mouse	ENSG00000153201	ENSMUSG00000003226	RP24-112B8	C-terminus	358
53	Tpr	mouse	ENSG00000047410	ENSMUSG00000006005	RP23-83J12	C-terminus	16, 266, 267
54	Rab5C	mouse	ENSG00000108774	ENSMUSG00000019173	RP23-258L5	C-terminus	23
55	CDC16	human	ENSG00000130177		RP11-450H16	C-terminus	14, 22, 55, 60, 66, 71, 72
56	NEK9	mouse	ENSG00000119638	ENSMUSG00000034290	RP23-331P12	C-terminus	107
57	NUP133	mouse	ENSG00000069248	ENSMUSG00000039509	RP24-295P8	C-terminus	24, 118, 120, 129
58	RPL23A	mouse	ENSG00000198242	ENSMUSG00000071532	RP23-299K11	C-terminus	18
59	CapH	mouse	ENSG00000121152	ENSMUSG00000034906	RP23-454M7	C-terminus	82
60	RAB8A	mouse	ENSG00000167461	ENSMUSG00000003037	RP23-280B22	C-terminus	24
61	CDCA5	mouse	ENSG00000146670	ENSMUSG00000024791	RP23-151I3	C-terminus	27
62	PAPSS1	mouse	ENSG00000138801	ENSMUSG00000028032	RP24-394D10	C-terminus	71
63	RASFF1(A)	mouse	ENSG00000068028	ENSMUSG00000010067	RP24-134D22	C-terminus	10, 16, 22, 31, 39
64	Apc1	mouse	ENSG00000153107	ENSMUSG00000014355	RP23-464B2	C-terminus	216
65	Apc5	mouse	ENSG00000089053	ENSMUSG00000029472	RP23-302F1	C-terminus	25, 85
66	CDC23	mouse	ENSG00000094880	ENSMUSG00000024370	RP23-259E2	C-terminus	68
67	AurB	human	ENSG00000178999		CTD-2604F21	C-terminus	39
68	TPX2	mouse	ENSG00000088325	ENSMUSG00000027469	RP24-370E11	C-terminus	86, 89

Supplementary Table 2: Localization and Mass-spec result in comparison with the literature data

	Transgene information						Localization				Protein interaaction			
	symbol	gene name	gene size ATG-Stop [kbp]	predicted protein size [kDa]	species	Tag	positive/total	positive (%)	Localization	Supporting reference	Mw kDa	Mascot	Seq Cov	Selected co-purifying proteins
1	CDC2	cell division cycle 2	16278	34	mouse	C	133 / 170	78,2	(centrosome) - NV	[1]	34,1	1728	72,1%	Cyclins A2, B1, B2; Cks1+2, p27kip1, Separase
2	DYNC1H1	dynein heavy chain 1	65474	532	mouse	C	106 / 133	79,7	Cytoplasm, (spindle) - NV	[2-4]	532,0	7518	29,8%	Dynein complex (6 proteins)
3	DYNC1I2	dynein intermediate chain 2	51318	70	mouse	C	81 / 201	40,3	Cytoplasm, spindle	[2, 3, 5]	68,4	1521	50,5%	Dynein complex (6 proteins)
4	DYNLL1	dynein light chain 1	3809	10	mouse	C	189 / 194	97,4	Cytoplasm, spindle	[2, 3, 6]	10,4	477	65,2%	Dynein complex (6 proteins)
5	CDC23	Cell division cycle protein 23 homolog	20054	68	mouse	C	192 / 215	89,3	Cytoplasm, (spindle) - NV	[5]	56,5	1386	52,7%	APC (11 proteins), BubR1, Cdc20, Bub3, Mad2
6	MIS12	MIND kinetochore complex component (mouse)	7524	24	mouse	C	177 / 222	79,7	Kinetochore	[6]	24,1	798	69,9%	13 kinetochore proteins
7	MIS12	MIND kinetochore complex component (human)	4418	24	human	C	226 / 240	94,2	Kinetochore	[6]	24,1	509	43,9%	13 kinetochore proteins
8	TOR1AIP1	torsin-1A-interacting protein 1 (LAP1B)	29512	76	mouse	C	188 / 209	90,0	Nuclear envelop	[7, 8]	66,8	1313	52,4%	bait only
9	INCENP	inner centromere protein	27122	105	mouse	C	168 / 215	78,1	Inner centromere, spindle midzone, midbody	[9]	101,2	537	10,6%	Aurora-B, Borealin
10	AURKB	aurora kinase B	6018	40	mouse	C	132 / 196	67,3	Inner centromere, spindle midzone, midbody	[10]	39,3	770	48,1%	INCENP, Borealin, Survivin
11	RAB5C	Ras-related protein Rab-5C	23178	23	mouse	N	89 / 133	66,9	Cytoplasmic dots	[11]	23,4	1080	85,6%	GDP dissociation inhibitor 2
12	STAG2	cohesin subunit SA-2, Stromal antigen 2 (SA2)	125100	141	mouse	C	127 / 201	63,2	Chromatin-associated, cytoplasm - NV	[12]	141,3	2268	40,0%	Cohesin proteins: Smc1, Smc3, Scc1, Pds5A, Pds5B, WAPL
13	SGOL1	shugoshin-like 1	13253	60	mouse	C	112 / 146	76,7	Kinetochore	[13]	59,0	430	18,8%	PP2A subunits: PR65-alpha, SET(I-2PP2A), B56-delta
14	TUBG1	gamma-1 tubulin	6288	51	mouse	C	180 / 189	95,2	Centrosome	[14]	51,1	1011	38,4%	Gamma-tubulin complex (5 proteins)
15	PCNA	proliferating cell nuclear antigen	3894	29	mouse	C	176 / 220	80,0	Nuclear dots	[15]	29,1	557	45,2%	p15PAF, NME1

References:

1. Pockwinse, S., et al., *Cell cycle independent interaction of CDC2 with the centrosome, which is associated with the nuclear matrix-intermediate filament scaffold*. Proc Natl Acad Sci U S A, 1997. **94**(7): p. 3022-7.
2. Pfarr, C., et al., *Cytoplasmic dynein is localized to kinetochores during mitosis*. Nature, 1990. **345**(6272): p. 263-5.
3. Dick, T., et al., *Cytoplasmic dynein (ddlc1) mutations cause morphogenetic defects and apoptotic cell death in Drosophila melanogaster*. Mol Cell Biol, 1996. **16**(5): p. 1966-77.
4. Pfister, K.K., et al., *Genetic analysis of the cytoplasmic dynein subunit families*. PLoS Genet, 2006. **2**(1): p. e1.
5. Ollendorff, V. and D.J. Donoghue, *The serine/threonine phosphatase PP5 interacts with CDC16 and CDC27, two tetratricopeptide repeat-containing subunits of the anaphase-promoting complex*. J Biol Chem, 1997. **272**(51): p. 32011-8.
6. Goshima, G., et al., *The role of Ppe1/PP6 phosphatase for equal chromosome segregation in fission yeast kinetochore*. Embo J, 2003. **22**(11): p. 2752-63.
7. Yang, L., T. Guan, and L. Gerace, *Integral membrane proteins of the nuclear envelope are dispersed throughout the endoplasmic reticulum during mitosis*. J Cell Biol, 1997. **137**(6): p. 1199-210.
8. Kondo, Y., et al., *Molecular cloning of one isotype of human lamina-associated polypeptide 1s and a topological analysis using its deletion mutants*. Biochem Biophys Res Commun, 2002. **294**(4): p. 770-8.
9. Vagnarelli, P. and W. Earnshaw, *Chromosomal passengers: the four-dimensional regulation of mitotic events*. Chromosoma, 2004. **113**(5): p. 211-22.
10. Carmena, M. and W.C. Earnshaw, *The cellular geography of aurora kinases*. Nat Rev Mol Cell Biol, 2003. **4**(11): p. 842-54.
11. Barbieri, M., et al., *Epidermal growth factor and membrane trafficking. EGF receptor activation of endocytosis requires Rab5a*. J Cell Biol, 2000. **151**(3): p. 539-50.
12. Sumara, I., et al., *Characterization of vertebrate cohesin complexes and their regulation in prophase*. J Cell Biol, 2000. **151**(4): p. 749-62.
13. Kitajima, T., et al., *Human Bub1 defines the persistent cohesion site along the mitotic chromosome by affecting Shugoshin localization*. Curr Biol, 2005. **15**(4): p. 353-9.
14. Wiese, C. and Y. Zheng, *Microtubule nucleation: gamma-tubulin and beyond*. J Cell Sci, 2006. **119**(Pt 20): p. 4143-53.
15. Leonhardt, H., et al., *Dynamics of DNA replication factories in living cells*. J Cell Biol, 2000. **149**(2): p. 271-80.

SUPPLEMENTARY METHODS

1 Generation of BAC transgenic constructs

1.1.1 Materials and equipment

- Plate shaker.
- 96 well electroporator (Harvard Apparatus, Holliston, MA, USA)
- Multichannel pipettes.
- 2ml square deep well 96 well plates.
- Plate sealing film (ABgene).
- Breathable plate sealing film (Corning).
- YNEB medium (Bacto yeast extract 0.75%, Bacto Nutrient Broth 0.8%).
- Sterile 10% glycerol (filter sterilize and keep in the cold room).
- SOC medium (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4 and 20 mM glucose).
- LB medium.
- Tetracycline (10 mg/ml in ethanol)
- Chloramphenicol (30 mg/ml in ethanol).
- Kanamycin (50 mg/ml in H_2O).
- Sterile 50% glycerol.
- L-Arabinose 10% solution (freeze 5 ml aliquots)
- BAC isolation kit - NucleoBond PC 100, Macherey-Nagel, 740 573.100

1.2 Bioinformatics

Mouse and human BAC end pairs and their genomic coordinates (on genome assemblies mm8 and hg18 respectively) were downloaded from the UCSC genome browser web site (http://hgdownload.cse.ucsc.edu/downloads.html). Sequence information for genes of interest was retrieved from the Ensembl databases[1]. The BAC selection criteria were (i) the BAC must contain the gene of interest; (ii) the gene should be at least 10 kb from each end of the BAC to allow for the inclusion of regulatory regions; (iii) the BAC should be as small as possible

since smaller BACs are easier to handle. PCR primers were designed using the Primer3 program[2].

1.3 High throughput generation of BAC transgenes by 96 well recombineering

1.3.1 Generation of tagging cassettes by PCR

1. Set up 50 µl PCR reactions in a 96 well plate format. Use of proofreading polymerase is recommended. Use high quality oligonucleotides, at least HPLC purified. The optimal annealing temperature is 57-63°C. Do not use more cycles than necessary (20-25) to reduce the risk of PCR introduced mutations. Increase the amount of template if necessary to reduce the number of cycles.
2. Purify the products using Eppendorf Perfectprep PCR Cleanup 96 according to manufacture protocol.
3. Elute in 500 µl of distilled water (to final concentration of about 5 ng/µl).

1.3.2 Rearray of the BAC clones in a 96well plate.

4. Plate from glycerol stock onto L agar plate containing 12,5 µg/ml chloramphenicol.
5. Grow overnight at 37°C.
6. Pick single colonies and inoculate into 96 well plate with 1ml of LB plus chloramphenicol 12,5 µg/ml. Keep track of the order. Save the coordinates of each clone and label the plates accordingly.
7. Grow overnight at 30°C.
8. The next day transfer 200 µl of the overnight culture to each of 2 plates with 800 µl of LB plus 25% glycerol and 12,5 µg/ml chloramphenicol.
9. Seal the plates, label them and store in 2 separate freezers.

1.3.3 Transformation of the pSC101gbaA(tetR) plasmid

Start with fresh overnight culture of the BAC clone. If starting from frozen stock, grow a fresh overnight culture first.

10. Pipette 1 ml of YENB plus 12,5 µg/ml chloramphenicol to each 96 well.
11. Inoculate with 40 µl overnight culture of the BAC clones.
12. Seal the plate with breathable plate seal and grow for 2h at **37°C**, 900 rpm.
13. Spin down for 10 min at 5000 g.
14. Toss away the supernatant and press the plate onto a paper tower (made of 5-6 layers of clean paper tissue).
15. Add 1 ml of ice-cold sterile 10% glycerol to each plate.
16. Seal the plate with plastic seal and shake at 1400 rpm for 1 min.
17. Spin down for 10 min at 5000 g.
18. Toss away the supernatant and press the plate onto a paper tower (made of 5-6 layers of clean paper tissue).
19. Add 100 µl of pSC101gbaA(tetR) in H_2O (0,1 ng/µl).
20. Pipette up and down 5 times to suspend the pellet and transfer to the electroporation cuvette.
21. Electroporate at 2500V.
22. Transfer the cells to a new plate with 1ml of SOC medium per well.
23. Seal with the breathable film and let the cells recover for 1h at 30°C, 900 rpm.
24. Transfer 100 µl to a new plate with 1ml of YENB plus 12,5 µg/ml chloramphenicol and 5 µg/ml tetracycline.
25. Seal with breathable seal film and incubate overnight at **30°C**.
26. The next day transfer 200 µl of the overnight culture to each of 2 plates with 800 µl of LB plus 25% glycerol, 12,5 µg/ml chloramphenicol and 5 µg/ml tetracycline.
27. Seal the plates, label them and store in 2 separate freezers.

1.3.4 Tagging by Red/ET recombination

Start with fresh overnight culture of the BAC clone plus pSC101gbaA(tetR). If starting from frozen stock, grow a fresh overnight culture first.

28. Pipette 1 ml of YENB plus 12,5 μg/ml chloramphenicol and 5 μg/ml tetracycline to each 96 well.
29. Inoculate with 40 μl overnight culture of the BAC plus pSC101gbaA(tetR).
30. Seal the plate with breathable plate seal and grow for 2h at **30°C**, 900 rpm.
31. Add 20 μl of 10% L-Arabinose to each well, seal again with breathable seal and grow for further 1h at **37°C**, 900 rpm.
32. Spin down for 10 min at 5000 g.
33. Toss away the supernatant and press the plate onto a paper tower (made of 5-6 layers of clean paper tissue).
34. Add 1,2 ml of cold sterile 10% glycerol to each plate.
35. Seal the plate with plastic seal and shake at 1400 rpm for 1 min.
36. Spin down for 10 min at 5000 g.
37. Toss away the supernatant and press the plate onto a paper tower (made of 5-6 layers of clean paper tissue).
38. Add 100 μl of PCR product in H_2O (5 ng/μl).
39. Pipette up and down 5 times to suspend the pellet and transfer to the electroporation cuvette.
40. Electroporate at 2500V.
41. Transfer the cells to new plate with 1ml of SOC medium per well.
42. Seal with the breathable film and let the cells recover for 1h at **37°C**, 900 rpm.
43. Transfer 100 μl to a new plate with 1ml of YENB plus 12,5 μg/ml chloramphenicol and 25 μg/ml kanamycin.
44. Seal with the breathable film and incubate overnight at **37°C**.
45. The next day transfer 200 μl of the overnight culture to each of 2 plates with 800 μl of LB plus 25% glycerol, 12,5 μg/ml chloramphenicol and 25 μg/ml kanamycin.
46. Seal the plates, label them and store in 2 separate freezers.

1.3.5 Check PCR

47. To check the tagged clone of interest streak the glycerol stock on a selective L-agar (12,5 µg/ml chloramphenicol and 25 µg/ml kanamycin).
48. Using a pipette tip touch a single colony and pipette up and down in the PCR mix (on ice). Do not take too much bacteria, as that will inhibit the PCR reaction.

PCR mix:

H2O	9,95	µl
F	500	nM
R	500	nM
10X buffer	3	µl
dNTP	250	µM
Taq	0,3	µl
total	30	µl

49. Run a PCR reaction as follows:

	95 - 5'	
35x	94 - 20"	
	55 - 20"	
	68 - 2' 30"	
	68 - 5'	

(You may need to optimize the conditions for different primer pairs)

50. Check 3 µl of the reaction on 0,8% agarose gel.

1.4 Purification of BAC transgenes for transfection

51. Inoculate a fresh colony of modified BAC into 50 ml of LB media supplemented with chloramphenicol and kanamycin and grow overnight at 37°C.
52. Harvest bacteria from LB culture (OD600 ~1.8-2.0) by centrifugation at 4,500g for 15 min at 4°C and discard supernatant.
53. Proceed with BAC isolation by following the protocol "Low-copy plasmid purification: Maxi/BAC" of the Nucleobond AX 100 kit. Isolated BAC can be stored at 4°C for up to 4 weeks.
54. Run 10 µl of the isolated BAC on a 0.8% agarose gel (70V, 1h) to verify good quality of the BAC isolation. For best transfection results, isolated BACs need to be of high quality. A large fraction of supercoiled BAC is especially important.

2 BAC transfection and generation of transgenic cells, mice and MEFs

2.1 Materials

- HeLa and MCF-7 cells culture media: HeLa cells were cultured in DMEM/Glutamax (4.5 g glucose/500 ml, Invitrogen, Germany) supplemented with 10% FCS (PAN, Germany), 100 units/ml Penicillin, and 100 µg/ml Streptomycin (Gibco, Germany). The medium for MCF-7 cells contained in addition bovine insulin (0.01 mg/ml, Sigma).
- Mouse ES cells culture media: DMEM/Glutamax (4.5g glucose/500ml, Invitrogen, Germany) supplemented with 20% FCS, 50µM beta-Mercaptoethanol (Gibco, Germany), 1x non-essential amino acids (Invitrogen, Germany), 100 units/ml Penicillin, and 100µg/ml Streptomycin (Gibco, Germany), and 2.5ml/100ml cell culture supernatant taken from LIF (Leukemia inhibitory factor) - overexpressing CHO cells.
- G418 (Geneticin) – Invitrogen, 10131-019
- Effectene – Qiagen, 301427 (4x1ml)
- Lipofectamine 2000 – Invitrogen, 11668-019
- OptiMEM I – Invitrogen, 31985-062

2.2 Stable transfection

2.2.1 1 week before stable transfection

1. Thaw a vial of HeLa or mouse ES cells and culture the cells about 1 week prior stable transfection.

2.2.2 Day 1: day before stable transfection

2. Plate 200,000 cells into tissue-culture dishes (60 mm or 6-well plate, see Table 1a and 1b for details). Use one dish for each modified BAC to be transfected. Additionally, consider two more plates to transfect an unmodified BAC (negative control) as well as a verified BAC serving as a positive transfection control.

2.2.3 Day 2: day of stable transfection

3. Place isolated BAC and transfection reagents (Effectene or Lipofectamine 200) on ice.
4. Prepare the transfection mix for each BAC to be transfected using a separate 1.5 ml cup. Transfection is performed according to the manufactures protocol supplied with each transfection reagent. For details on amounts refer to Table 1.
5. Add the entire transfection mix drop-wise to the cells and mix by gently rotating the whole dish horizontal.

2.2.4 Day 3: 1 day after stable transfection

6. Change complete cell culture media the next day.
7. Day 4: 2 days after stable transfection
8. Change media and culture the cells in complete media supplemented with G418 (see Table 2).

2.2.5 Day 7 (5 days after stable transfection) to day 21

9. Increases G418 concentration (see Table 2) and change media every 2-3 days to remove dead cells and cell debris. 2 weeks after transfection, cells transfected with unmodified BAC (negative control) are dead or critically decimated. Distinct stable colonies are visible in the cell dishes transfected with modified BACs. Cells transfected with unmodified BAC need to be completely dead before further processing of the resistant clones transfected with the gene of interest.

2.2.6 3 weeks after transfection

10. Stable colonies are pooled after trypsinization of the cells. Subsequently, cells are amplified to increase the cell number for freeze-downs and further analysis (western blot, immunofluorescence, pull-down assays).

Table 1:

Cell line	Cell number	Isolated BAC (µg)	EC Buffer (µl)	Enhancer (µl)	Effectene (µl)
Hela-Kyoto Or MCF7	200,000/ 28cm2	50-250	Up to 150	8	10
Mouse ES cells	200,000/ 10cm2	150-350	10	Up to 500	Mouse ES cells

Table 2:

	G418 concentration (µg/ml)		
Cell line	Day 4 – day 6 of transfection	Day 7 till end of selection	Maintenance
Hela-Kyoto	50	400	500
MCF7	50	400	400
Mouse ES cells	250	250	250

2.3 Generation of transgenic mouse embryos

Fully ES cell derived embryos laser-assisted injection technology was used to inject stable mPCNA-LAP ES cells into the eight-cell morula stage of C56BL/6 mice using

laser-assisted injection technology[3]. At 13.5 dpc, 10 embryos were analyzed for positive GFP expression.

2.4 Isolation of mouse embryonic fibroblasts (MEFs)

Trypsinize the tissue of a GFP-positive embryo (13.5 dpc) and carefully resuspend the cells in DMEM/Glutamax (4.5g glucose/500ml, Invitrogen, Germany) supplemented with 10% FCS, 100 units/ml Penicillin, and 100 µg/ml Streptomycin (Gibco, Germany).

3 Antibodies:

Western blot - mouse monoclonal anti-GFP (Roche); IP and ChIP - Goat anti-GFP (raised against His-tagged full-length eGFP and affinity-purified with GST-tagged full-length eGFP), goat anti-FoxA1 (ab5089) from Abcam, goat anti-XBP1 (sc-7160X), and normal goat IgG (sc-2028) from Santa Cruz Biotechnologies (Santa Cruz).

4 Protein localization

4.1 Immunofluorescencne

Immunofluorescence stainings were performed on cells grown on glass coverslips. After fixing cells in 3% PFA supplemented with 5 mM EGTA, 1mM $MgCl_2$ and 2% sucrose, they were washed, then permeabilized using PBS/0.1% Triton X-100 and eventually blocked with PBS/0.2% Fish Skin Gelatin (Sigma, Germany). Subsequently, cells were incubated with primary antibodies: goat anti-GFP (antibody facility, MPI-CBG, Germany), mouse anti-tubulin (DM1a, Sigma, Germany); and the following secondary antibodies: donkey anti-goat IgG-Alexa488 and donkey anti-mouse Alexa594 (Invitrogen, Germany). DNA was counterstained using Hoechst 33342 (100 µg/ml). Images were taken on an Olympus IX70 equipped with the imaging system DeltaVision RT and deconvolved using softWoRx software (Applied Precision, USA).

4.2 Time-lapse movies of HeLa and mouse ES cells

16h prior to imaging, 15000 HeLa cells and 8000 ES cells, respectively, were seeded into 8-chamber-LabTek Chambered Coverglass (Nalge Nunc, Rochester, NY, USA), resulting in a confluence of 50-60%. For confocal imaging the LabTeks were placed in a heated chamber and the culture medium was supplemented with additional 20mM Hepes (Sigma, Germany). Images were taken every 3 min for 18 h on an Olympus IX70 equipped with the imaging system DeltaVision RT and deconvolved using softWoRx software (Applied Precision, USA).

5 Fluorescence in situ hybridisation

HeLa cell lines transgenic with mouse BACs containing the AURKB or KIAA1387 gene, respectively, were subjected to fluorescence in situ hybridization (FISH). As probes we used the respective mouse BACs (RP23-417I6 or RP23-393O10) and a human BAC for the homologous regions (RP11-1322B12 or RP11-554J4). BACs were labeled by nicktranslation with digoxigenin-dUTP or biotin-dUTP and correct chromosomal localization was confirmed by FISH on normal, diploid human and mouse metaphase chromosome spreads together with the respective chromosomal library (data not shown). Metaphase spreads from transgenic HeLa cells were prepared using standard methods, subjected to FISH and recorded with a CCD-Camera attached to a Zeiss Axiophot epifluorescence microscope. For 3D-FISH, cells were fixed with 4% freshly made buffered paraformaldehyde, permeabilized and subjected to FISH as previously described[4], with ToPro-3 (Invitrogen Molecular Probes, Karlsruhe, Germany) as a DNA counterstain. 3D-microscopy was performed on a Leica SP1 confocal laser scanning microscope (Leica-Microsystems, Wetzlar, Germany) with laser lines for 488 (for FITC and Alexa488), 561 (for Cy3) and 633 nm (for ToPro-3) with a 100x NA 1.4 objective and a voxel size of 0.08 x 0.08 x 0.24 μm^3. Projections of confocal sections were made in ImageJ 1.38j (http://rsb.info.nih.gov/ij/) and signal intensity was adjusted for RGB-overlays by linear scaling with the 'levels' command in Adobe Photoshop.

6 FACS analysis

For FACS analysis of the GFP fluorescence intensity distribution normalized to DNA content the cells were stained with 83 µM DRAQ5 (Squarix biotechnology, Germany). The GFP and DRAQ5 fluorescence was measured on a FACS Calibur (Becton Dickinson). The ratio of GFP/DRAQ5 was calculated using FlowJo software (Tree Star, Inc., USA).

7 Affinity purification and mass spectrometry

7.1 Materials

• HeLa Culture Medium: Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% foetal bovine serum, 0.2 mM L-glutamine, 100 U/ml penicillin and 100 µg/ml streptomycin. Include G-418 (Gibco/Invitrogen) at a concentration appropriate for maintenance of the tagged transgene (typically 500 µg/ml).

• LAP Buffer: 50 mM HEPES-KOH, pH 7.5, 5 mM EDTA, 150 mM KCl, 10% (v/v) glycerol.

• Protease Inhibitor Mix ('PIM', 1000×): 10 mg/ml each of leupeptin, pepstatin and chymostatin, in DMSO.

• Leupeptin-Pepstatin Mix ('LP Mix', 1000×): 1 mM each of leupeptin and pepstatin, in DMSO.

• Lysis Buffer: LAP Buffer plus 1% (v/v) Triton X-100, 20 mM beta-glycerophosphate, 10 mM NaF, 10 mM Na-pyrophosphate, 1× PIM, 0.1 mM PMSF, 1 mM Na3VO4, 1 mM DTT.

• Wash Buffer: LAP Buffer plus 0.5% (v/v) NP-40, 20 mM beta-glycerophosphate, 10 mM NaF, 10 mM Na-pyrophosphate, 1× PIM, 0.1 mM PMSF, 1 mM Na3VO4, 1 mM DTT.

• Cleavage Buffer: LAP Buffer plus 0.5% (v/v) NP-40, 20 mM beta-glycerophosphate, 10 mM NaF, 10 mM Na-pyrophosphate, 1× LP Mix, 0.1 mM PMSF, 1 mM Na3VO4, 1 mM DTT.

• Elution Buffer: 0.1 M glycine, pH 2.0.

7.2 Procedure

1. HeLa cell culture: grow HeLa-BAC cells on monolayers in HeLa Culture Medium. To obtain mitotically-arrested cells, culture cells on five 25×25-cm

tissue culture trays until approx. 75% confluence has been obtained, then add nocodozole to 0.1 µg/ml in the medium, and incubate for 18 h.

2. Cell harvesting: scrape the cells into their medium on the trays, transfer into a centrifuge vessel, and pellet the cells. Wash the pellet in ice-cold PBS, spin again and aspirate the supernatant. Snap-freeze the cell pellet (typically 2-3 ml, approx 3.2×108 cells) and store at –80˚C.
3. Steps 3-10 should be performed at 4˚C or on ice.
4. Making cell suspension: thaw the cell pellet, add sufficient Lysis Buffer to bring the total volume to 5 ml, and resuspend the pellet thoroughly.
5. Cell lysis: using a syringe, force the cell suspension through a 20 G needle ten times. Keep the suspension on ice for 5 min, then repeat the syringing.
6. Clarifying the extract: Centrifuge at maximum speed (e.g. 20,000 rpm in a Sorvall SS-34 rotor) for 15 min. Transfer the supernatant into a fresh tube.
7. Anti-GFP immunoprecipitation: add clarified extract to 50 µl AffiPrep-Protein A beads (Bio-Rad), crosslinked with anti-GFP antibody (1 µg antibody per µl beads, pre-equilibrated with Lysis Buffer), and incubate end-over-end for 1 h. Spin the beads briefly, remove the supernatant.
8. Washing anti-GFP beads: add 2 ml Wash Buffer, incubate end-over-end for 1 min, spin and aspirate supernatant; repeat five times. Wash the beads twice more with 2 ml Cleavage Buffer; spin and aspirate the supernatant.
9. PreScission cleavage: add 300 µl Cleavage Buffer to the beads plus 3 µl PreScission protease (2 U/µl, GE Healthcare). Rotate end-over-end for 30 min, spin briefly and transfer supernatant into a fresh tube. Add 100 µl Cleavage Buffer to the beads, mix briefly, centrifuge as before and transfer the supernatant to the previous tube.
10. S-protein pull-down: add the combined supernatants from the previous step to 50 µl S-protein agarose beads (Novagen, pre-equilibrated with

Wash Buffer), and incubate end-over-end for 1 h. Spin the beads briefly, remove the supernatant.

11. Washing S-protein beads: add 1 ml Wash Buffer, incubate end-over-end for 1 min, spin and aspirate supernatant; repeat this twice more. Then wash the beads twice with 1 ml of LAP Buffer, then twice with 1 ml of 150 mM KCl.
12. Elution of protein complexes: add 50 µl Elution Buffer to the beads, mix and incubate for 1 min at room temperature. Spin the beads, transfer the supernatant to a fresh tube. Repeat the elution and pool the second eluate with the first. Neutralise the eluates by adding Tris-HCl, pH 9.2, and checking with pH-indicator paper until the pH = approx. 8.0.
13. SDS-PAGE – silver staining: take a 20-µl sample of the neutralised eluate and boil with SDS-sample buffer. Electrophorese the sample on a gradient SDS-PAGE gel. Silver stain the gel to visualise the purified proteins.
14. In-solution digestion and mass spectrometry: the remaining eluate should be reduced, alkylated, trypsinised and analysed by LC-MS/MS as described by Gregan et al. (Nat Protocols, 2007, Vol2, no5, p.1145). Note that where the tagged bait is a mouse protein, the protein database used for MS data searching should contain the mouse bait as well as the human proteome sequences.

8 Chromatin immunopurification

8.1 LAP tag based ChIP

Cells at 80% confluency (~5x10^6 cells per ChIP) were cross-linked with 1% formaldehyde for 10 minutes at 37°C, and quenched with 125 mM glycine. The fixed cells were washed twice with cold PBS and scraped, and transferred into 1 ml PBS containing protease inhibitors (Roche). After centrifugation at 700 g for 4 minues at 4°C, the cell pellets were resuspended in 100 µl ChIP lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.1] with protease inhibitors) and sonicated at 4°C with Bioruptor (Diagenode) (30 seconds ON and 30 seconds OFF at highest power for 12 minutes). The sheared chromatin with an average fragment length of ~500 bp) was centrifuged at 10,000 g for 10 minutes at 4°C). 100 µl of the supernatant was used for ChIP or as input. A 1:10 dilution of the solubilized chromatin in ChIP buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 167 mM NaCl. 16.7 mM Tris-HCl [pH 8.1]) was incubated at 4°C overnight with 4 µg/ml of specific antibodies. We used the following antibodies for ChIP: goat anti-GFP (raised against His-tagged full-length eGFP and affinity-purified with GST-tagged full-length eGFP), goat anti-FoxA1 (ab5089) from Abcam, goat anti-XBP1 (sc-7160X), and normal goat IgG (sc-2028) from Santa Cruz Biotechnologies (Santa Cruz). Immunoprecipitations were carried out by incubating with 40 µl pre-cleared Protein G Sepharose beads (Amersham Bioscience) for 1 hour at 4°C, followed by five washes for 10 minutes with 1ml of

the following buffers: Buffer I: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 150 mM NaCl; Buffer II: 0.1% SDS, 1% Triton X-100, 2 mM EDTA, 20 mM Tris-HCl [pH 8.1], 500 mM NaCl; Buffer III: 0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1 mM EDTA, 10 mM Tris-HCl [pH 8.1]; twice with TE buffer [pH 8.0]. Elution from the beads was performed twice with 100 μl ChIP elution buffer (1% SDS, 0.1 M $NaHCO_3$) at room temperature (RT) for 15 minutes. Protein-DNA complexes were de-crosslinked by heating at 65°C in 192 mM NaCl for 16 hours. DNA fragments were purified using QiaQuick PCR Purification kit (Qiagen) and eluted into 30 μl H_2O according to the manufacturer's protocol after treatment with RNase A and Proteinase K. For ChIP-PCR assays, 3 μl of the ChIPed DNA were amplified using the forward primer 5'-TTTTTGAGGGACGTGTTTCA-3' and the reverse primer 5'-AGTCTTTTCCCAGGAGTATTGC-3' in the *PIT1* promoter with the following PCR conditions: initial denaturation at 94 °C for 5 minutes, 32 cycles of 94°C for 40 seconds, 58 °C for 30 seconds, and 72 °C for 30 seconds, followed by a final elongation step of 72°C for 7 minutes. For ChIP-on-Chip, both ChIP DNA and input DNA were subjected to the linker-mediated PCR amplification, i.e., treated with Polynucleotide Kinase (NEB), blunt-ended using Klenow Polymerase (NEB), ligated to linkers (oligonucleotides 5'-AGAAGCTTGAATTCGAGCAGTCAG-3' annealed to 5'-CTGCTCGAATTCAAGCTTCT-3') using T4 ligase (NEB) and amplified by PCR with the following conditions: initial denaturation at 94 °C for 5 minutes, 32 cycles of 94°C for 40 seconds, 55°C for 30 seconds, and 72 °C for 75

seconds, followed by a final elongation step of 72°C for 10 minutes. A dNTP mixture containing dUTP (at ratio dTTP:dUTP = 4:1) was used in the above linker-mediated PCR amplification. ChIP and Input DNA samples were further fragmented and end-labeled with biotin with GeneChip® WT Double-Stranded DNA Terminal Labeling Kit (Affymetrix). The resulting labeled samples were hybridized to the 4R array (including chromosomes 4, 15, 18, and 20) from Affymetrix GeneChip® Human Tiling 2.0R Array Set following the Affymetrix® Chromatin Immunoprecipitation Assay Protocol. Independent biological triplicates were performed for each transcription factor, as well as the control (input).

8.2 Analysis of ChIP-on-Chip tiling array data

ChIP-on-Chip tiling array data were normalized and analyzed as previously described [5-7]. Briefly, raw tiling array data were quantile-normalized within ChIP and control groups [8] and scaled to a median intensity of 500. For each genomic position of a 25-mer probe on the array, a local dataset composed of intensities for all adjacent probes within a window of ±250 bp was generated. A one-tailed Wilcoxon rank sum test was then applied to test the null hypothesis that intensities of ChIP and control experiments within the local dataset are from the same distribution, against the alternative that the distribution of the ChIP data is positively shifted when compared to the distribution of the control data. This test procedure was performed in a sliding window across all tiled genomic regions. Significantly enriched probes, defined by applying a P-value cut-off of 1e-4, were locally extended by merging adjacent enriched probes within 100 bp and these

merged regions (with a minimum length cutoff = 100 bp) were defined as transcription factor bound regions. When comparing significant binding regions identified by ChIP with antibodies against the endogenous transcription factor (FOXA1 or XBP1) and the LAP tag, the common binding regions are considered only if the centers of two regions are within 1 kb.

References:

1. Hubbard, T.J. et al. Ensembl 2007. *Nucleic acids research* **35**, D610-617 (2007).
2. Rozen, S. & Skaletsky, H. Primer3 on the WWW for general users and for biologist programmers. *Methods in molecular biology (Clifton, N.J* **132**, 365-386 (2000).
3. Poueymirou, W. et al. F0 generation mice fully derived from gene-targeted embryonic stem cells allowing immediate phenotypic analyses. *Nat Biotechnol* **25**, 91-99 (2007).
4. Hepperger, C., Otten, S., von Hase, J. & Dietzel, S. Preservation of large-scale chromatin structure in FISH experiments. *Chromosoma* **116**, 117-133 (2007).
5. Cawley, S. et al. Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. *Cell* **116**, 499-509 (2004).
6. Bernstein, B.E. et al. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169-181 (2005).
7. Bernstein, B.E. et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).
8. Bolstad, B.M., Irizarry, R.A., Astrand, M. & Speed, T.P. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* **19**, 185-193 (2003).