Codon adaptation–based control of protein expression in *C. elegans*

Stefanie Redemann[1], Siegfried Schloissnig[2], Susanne Ernst[1], Andrey Pozniakowsky[1], Swathi Ayloo[1], Antony A Hyman[1] & Henrik Bringmann[3]

We present a method to control protein levels under native genetic regulation in *Caenorhabditis elegans* by using synthetic genes with adapted codons. We found that the force acting on the spindle in *C. elegans* embryos was related to the amount of the G-protein regulator GPR-1/2. Codon-adapted versions of any *C. elegans* gene can be designed using our web tool, *C. elegans* codon adapter.

Controlling protein amounts is important for the correct function of a gene. It is therefore of interest to study the effects of increased or decreased amounts of specific proteins. Frequently, transgenes are expressed at levels too weak to be detected and thus an increase in expression is desirable, or toxic effects from overexpression are to be avoided, necessitating a decrease in expression.

We aimed to derive a method by which the amount of a protein can be adjusted between endogenous amounts and severe overexpression. The basis for our strategy was the finding that highly expressed genes preferentially use a specific set of codons[1], for which there is a greater pool of tRNAs compared to codons used in less-highly expressed genes[2,3]. The codon adaptation index (CAI) of a gene is defined as the geometric mean of the codon weights of all codons of that gene. The weight of a codon is determined by a reference set of genes that are known to be highly expressed[1]. If a gene has a CAI value of 1.0, this gene consists only of codons optimal for fast translation. If the extent of codon adaptation determines expression levels, it should be possible to fine-tune expression empirically by generating synthetic genes with different codon usage.

To test this idea, we applied it to genes involved in regulation of cell division. The movement of the spindle in *C. elegans* embryos is mediated by G-protein regulators 1 and 2 (GPR-1 and GPR-2), almost identical, partially redundant proteins. Reduction of function studies using RNAi have suggested that GPR-1/2 is required for applying force on the spindle[4–7]. We do not know, however, whether the force generation is directly related to GPR-1/2 amounts. We designed three synthetic *gpr-1* constructs with CAI values of 1.0, 0.6 or 0.3. We fused each of these synthetic *gpr-1* constructs to the same *yfp* gene and integrated each fusion into the genome of a *C. elegans* N2 starter strain, generating a separate transgenic strain for each CAI version of *yfp::gpr-1* (the three genes were named *yfp::gpr-1(synthetic, CAI 1.0)*, *yfp::gpr-1(synthetic, CAI 0.6)* and *yfp::gpr-1(synthetic, CAI 0.3)*). As controls, we used a strain into which we integrated endogenous *gpr-1* fused to *yfp* (*yfp::gpr-1 (endogenous, CAI 0.3)*) into the genome and wild-type worms not expressing any *yfp* transgene. We quantified protein expression by YFP fluorescence of one-cell stage embryos and by western blotting (**Fig. 1** and **Supplementary Fig. 1**). We designed all genes to be resistant to RNAi-mediated knockdown of the endogenous gene and to contain three artificial introns. YFP–GPR-1 expressed from the synthetic constructs localized to microtubule asters and to the cortex as described previously[4,8,9] (**Fig. 1a**). Worms expressing *yfp::gpr-1(synthetic, CAI 1.0)* were resistant to RNAi knockdown of wild-type *gpr-1* (**Supplementary Fig. 2**).

We first tested whether the synthetic versions of *gpr-1* were expressed in amounts corresponding to their CAI (**Fig. 1b**).

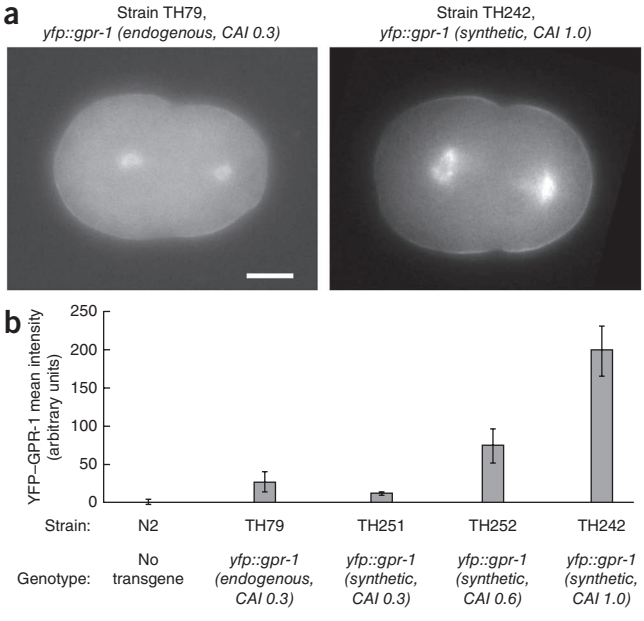

Figure 1 | Codon-adaptation of the *gpr-1* gene determines the amount of YFP–GPR-1 protein. (**a**) Fluorescence images showing localization of YFP–GPR-1 expressed from the indicated constructs in embryos at early anaphase. Images were acquired using different imaging conditions, and different contrast was used to display localization patterns. Scale bar, 10 µm. (**b**) Mean YFP fluorescence intensity over the surface of the embryo expressing the indicated constructs. *yfp::gpr-1(endogenous, CAI 0.3)*, 28 ± 13 arbitrary units (mean ± s.d.; $n = 6$ embryos); *yfp::gpr-1(synthetic, CAI 0.3)*, 12 ± 3 arbitrary units ($n = 6$ embryos); *yfp::gpr-1(synthetic, CAI 0.6)*, 75 ± 22 arbitrary units ($n = 6$ embryos); and *yfp::gpr-1(synthetic, CAI 1.0)*, 202 ± 34 arbitrary units ($n = 6$ embryos). N2, wild type expressing no transgene, 0 ± 5 arbitrary units ($n = 6$ embryos).

[1]Max Planck Institute of Molecular Cell Biology and Genetics, Dresden Germany. [2]European Molecular Biology Laboratory Heidelberg, Heidelberg, Germany. [3]Max Planck Institute for Biophysical Chemistry, Göttingen, Germany. Correspondence should be addressed to H.B. (henrik.bringmann@mpibpc.mpg.de).

RECEIVED 28 SEPTEMBER 2010; ACCEPTED 17 DECEMBER 2010; PUBLISHED ONLINE 30 JANUARY 2011; DOI:10.1038/NMETH.1565

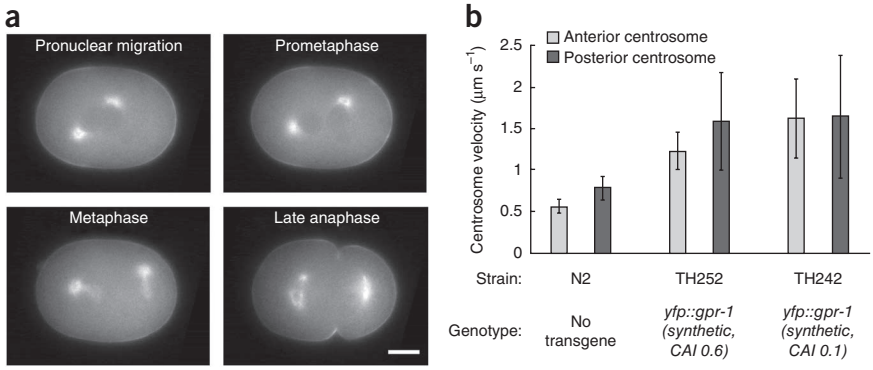

Figure 2 | Increasing GPR-1 amounts causes an increase in force acting on the mitotic spindle. (**a**) Time-lapse images of an embryo expressing *yfp::gpr-1(synthetic, CAI 1.0)* throughout the first cell division. Scale bar, 10 μm. (**b**) Histogram of centrosomal velocity after spindle cut by a UV light laser and spindle break (n = 5 embryos for all three strains; error bars, s.d.).

We measured the autofluorescence of wild-type N2 worms not expressing any *yfp* transgene and subtracted this value in all experiments. The expression of YFP–GPR-1 from *yfp::gpr-1(endogenous, CAI 0.3)* was roughly double the expression from *yfp::gpr-1(synthetic, CAI 0.3)* (P = 0.05) even though both constructs had the same CAI. The constructs contained endogenous and synthetic introns, respectively, which may have caused the difference in expression. The expression of YFP–GPR-1 from *yfp::gpr-1 (synthetic, CAI 0.6)* was significantly greater than the expression from *yfp::gpr-1(endogenous, CAI 0.3)* (P = 0.001) or *yfp::gpr-1 (synthetic, CAI 0.3)* (P < 0.0005). The expression of YFP–GPR-1 from *yfp::gpr-1(synthetic, CAI 1.0)* was significantly greater than the expression from *yfp::gpr-1(endogenous, CAI 0.3)* (P < 0.0005), *yfp::gpr-1(synthetic, CAI 0.3)* (P < 0.0005) and *yfp::gpr-1(synthetic, CAI 0.6)* (P < 0.0005) (**Fig. 1b**). We tested significance by a two-sided Welch's t-test with a confidence interval of 95%.

Western blotting using antibodies to GFP confirmed the expression levels determined by YFP fluorescence (**Supplementary Fig. 2**). The experiments showed that functional, RNAi-resistant synthetic transgenes can be made with variable CAI values. We observed consistent expression for different independently generated transgenic lines expressing the same construct, suggesting that the insertion site or copy number of insertion contributed much less to expression levels than the CAI value (**Supplementary Fig. 3**).

We next determined whether the force generated on the spindle corresponded to the amount of GPR-1 in the cell. We took time-lapse images of embryos expressing *yfp::gpr-1(synthetic, CAI 1.0)* throughout the first cell division. Unlike in wild-type embryos, the pronuclei-centrosome complex often separated before nuclear envelope breakdown during prometaphase. As a result, two half-spindles formed during metaphase. During anaphase, we often observed a snapping of the spindle and flattening of both centrosomes (**Fig. 2a**). These observations indicated greater forces acting on the centrosomes. To directly measure force, we cut the spindle with a laser. The velocity of movement of the two spindle poles after cutting was related to the force generated.

Compared centrosome velocity in wild-type N2 worms that do not have a transgene (anterior, 0.56 ± 0.08 μm s^{-1}; and posterior, 0.8 ± 0.1 μm s^{-1} (mean ± s.d.)), centrosome velocity was greater in *yfp::gpr-1(synthetic, CAI 0.6)* (anterior, 1.2 ± 0.2 μm s^{-1}, $P_{anterior}$ = 0.001; and posterior, 1.6 ± 0.6 μm s^{-1}, $P_{posterior}$ = 0.005) and in *yfp::gpr-1(synthetic, CAI 1.0)* (anterior, 1.6 ± 0.5 μm s^{-1}, $P_{anterior}$ = 0.02; and posterior 1.7 ± 0.7 μm s^{-1}, $P_{posterior}$ = 0.025). We tested significance with a two-sided Welch's t-test with a confidence interval of 95% (**Fig. 2b**).

Thus, we observed an equal increase in velocity of spindle poles in embryos that were overexpressing YFP–GPR-1 for both *yfp::gpr-1(synthetic, CAI 0.6)* and *yfp::gpr-1 (synthetic, CAI 1.0)*. However, the forces acting on the spindle appeared more symmetrical in *yfp::gpr-1(synthetic, CAI 1.0)* embryos than in *yfp::gpr-1(synthetic, CAI 0.6)* embryos.

Analysis of spindle-pole oscillations supported this finding (**Supplementary Fig. 4**). Knockout of either *gpr-1* or *gpr-2* resulted in reduced centrosome oscillation amplitude (**Supplementary Fig. 5**), consistent with previous reports[4–7]. These results indicate that the amount of GPR determines the force acting on the spindle.

We generated a program called *C. elegans* codon adapter, which is freely available at http://worm-srv3.mpi-cbg.de/codons/. The user can paste a sequence and specify the desired CAI as well as the number of synthetic introns and resistance to RNAi of the endogenous gene. Occasionally, sequences resembling splice sites are introduced during the codon-adaptation process. The *C. elegans* codon adapter has an option to avoid the introduction of splice sites. We tested the tool with two additional genes, *air-1* and *sas-4*, and obtained consistent results that the CAI determines the expression level of the protein (**Supplementary Figs. 6** and **7**).

Expression levels have been shown to depend on the stability of the mRNA at the ribosome-binding site in *Escherichia coli*[10]. We have not observed an effect of mRNA structure on expression in our experiments. However, we cannot exclude the possibility that mRNA structure affects expression, at least for some genes. If the goal is to produce synthetic genes with different CAI values that express according to their CAI, we recommend that users control the mRNA structure at the ribosome-binding site, which typically is easily possible by using the same 5′-end sequence for all the different CAI versions of the synthetic genes. As an example, the same 5′-end tag, which has not been codon-adapted, can be used for all different CAI versions. Alternatively, *C. elegans* codon adapter can be used to design the sequence of the mRNA ribosome-binding site to have both, a weak structure and the specified CAI.

Codon adaption–based control of protein levels has two major advantages over existing techniques. First, it allows preservation of 3′ and 5′ untranslated regions and thus endogenous genetic control. Second, it can potentially be applied simultaneously to many genes at a time. Theoretically, the expression level of every single gene of an organism could be modulated independently using codon adaptation, a prerequisite for genome engineering. As the endogenous CAI is generally low, this method should be applicable to any *C. elegans* gene and possibly also other model organisms such as *Drosophila melanogaster* and *Saccharomyces cerevisiae* (**Supplementary Table 1** and **Supplementary Fig. 8**).

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturemethods/.

Note: Supplementary information is available on the Nature Methods website.

ACKNOWLEDGMENTS

We thank E. Busch, M. Decker, N. Goehring and Z. Maliga for helpful discussions.

AUTHOR CONTRIBUTIONS

S.R. characterized all strains. S.S. wrote the web tool algorithm and analyzed genome-wide CAI. S.E. bombarded all constructs. A.P. cloned constructs. S.A. filmed embryos. A.A.H. mentored and financed the project. S.R., A.A.H. and H.B. wrote the paper. H.B. conceived the general synthetic gene design, cloned *gpr-1* constructs and preliminarily characterized *gpr-1* strains.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturemethods/.
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/.

1. Sharp, P.M. & Li, W.H. *Nucleic Acids Res.* **15**, 1281–1295 (1987).
2. Dong, H., Nilsson, L. & Kurland, C.G. *J. Mol. Biol.* **260**, 649–663 (1996).
3. Duret, L. *Trends Genet.* **16**, 287–289 (2000).
4. Colombo, K. *et al. Science* **300**, 1957–1961 (2003).
5. Pecreaux, J. *et al. Curr. Biol.* **16**, 2111–2122 (2006).
6. Srinivasan, D.G., Fisk, R.M., Xu, H. & van den Heuvel, S. *Genes Dev.* **17**, 1225–1239 (2003).
7. Gotta, M., Dong, Y., Peterson, Y.K., Lanier, S.M. & Ahringer, J. *Curr. Biol.* **13**, 1029–1037 (2003).
8. Barstead, R.J., Kleiman, L. & Waterston, R.H. *Cell Motil. Cytoskeleton* **20**, 69–78 (1991).
9. Bringmann, H., Cowan, C.R., Kong, J. & Hyman, A.A. *Curr. Biol.* **17**, 185–191 (2007).
10. Kudla, G., Murray, A.W., Tollervey, D. & Plotkin, J.B. *Science* **324**, 255–258 (2009).

ONLINE METHODS

Synthetic gene design and *C. elegans* transgenesis. We codon-adapted *gpr-1* using 'synthetic gene designer'[11]. The CAI was empirically adjusted by changing the 'optimality factor'. We used codon weights determined from a set of highly expressed genes[12]. Using DEQOR[13], we identified 21 nucleotide sequences in the synthetic genes that are present in *gpr-1* or *gpr-2*. We manually edited the sequences to remove such identical 21-nucleotide sequences. We also edited the sequence to remove stretches of 21-nucleotide sequence with one mismatch compared with *gpr-1* and *gpr-2*. Three introns were inserted into each synthetic gene. We used the following three introns[14]: 5′-gtaagtttaaacatatatatactaactaaccctgattatttaaattttcag-3′, 5′-gtaagtttaaacagttcggtactaactaaccatacatatttaaattttcag-3′ and 5′-gtaagtttaaacatgattttactaactaatctgatttaaattttcag-3′. The introns were placed between the third and fourth nucleotide of one of the following sequences: 5′-aagg-3′, 5′-aaga-3′, 5′-cagg-3′ or 5′-caga-3′. We distributed the introns (roughly equally spaced) across the gene. Three codons encoding glycine were inserted at the 5′ end of the synthetic gene, and TAA was used as a stop codon in all constructs. A 5′-end BamHI and a 3′-end XmaI sites were added.

Codon-adapted *air-1* and *sas-4* constructs were designed using our website, *C. elegans* codon adaptor. In the case of *air-1,* artificial introns were used as described above; for *sas-4*, endogenous genomic introns were used.

The constructs were synthesized by Geneart and were cloned into plasmid TH307 using BamHI and XmaI. Plasmid TH307 drives expression in the germline by a *pie-1* promoter and 3′ untranslated region. It also contains an *unc-119* rescue fragment. Strains expressing the synthetic constructs were generated by biolistic transformation into *unc-119(ed3)* worms[15]. Because previously described strains expressing *gpr-1(endogenous)* encoding a protein fused N-terminally to YFP had silenced their transgene expression, we generated a new strain by transforming the same plasmid as described previously[9].

Algorithm of *C. elegans* codon adaptor. *C. elegans* codon adapter uses previously published codon weights[12]. The sequence subject to optimization is represented internally as a series of codons. The core optimization routine then replaces codons continuously until the desired CAI is achieved or no codons are left that could be changed to reach the target CAI. Optimized codons are spread as evenly as possible throughout the sequence.

The binding site of the ribosome (nucleotides 1–37)[10] is optimized separately to achieve a weak RNA structure and to come as close as possible to the target CAI. For this, the algorithm randomly samples from the set of 37-mers coding for the same amino acids as the original sequence and calculates their CAI and free energy. From the result set, a sequence with weak RNA structure and favorable CAI is chosen. The Vienna RNA package[16] is used to calculate the energetic properties of the RNAs during the random sampling. The sequence at positions −4 to −1 is always 5′-aaaa-3′, as this is the consensus sequence[17].

If resistance to RNAi knockdown of the endogenous gene is desired, the algorithm proceeds by scanning the optimized sequence for segments of seven unaltered codons (21 nucleotides and 21 nucleotides with one mismatch). If such is found, the codon that causes the smallest possible change to the CAI is altered, and the scanning for unaltered segments continues from the modified position.

The final phase of the algorithm includes the removal of splice sites introduced during the previous steps. Therein, potential splice sites are first predicted using Netgene2 (refs. 18,19), then perturbed to remove consensus elements, and finally the whole sequence is optimized to compensate for potential changes to the target CAI. These three steps are repeated until all high-confidence splice sites have been removed or the upper limit for the number of iterations is reached.

***C. elegans* culture.** We grew *C. elegans* in culture as described previously[20]. We used the following *C. elegans* strains and alleles: N2, (wild type) VC1679, *gpr-1*(*ok2126*) I. RB1150, *gpr-2*(*ok1179*) III. TH26, *unc-119(ed3) III; ddEx10[gfp::sas-4; unc-119(+)].* TH41, *unc-119(ed3) III; ddIs5[gfp::air-1(K07C11.2); unc-119(+)].* TH79, *unc-119(ed3) III; ddIs21[yfp::gpr-1; unc-119(+)].* TH242, *unc-119(ed3) III; ddIs32[yfp::gpr-1(synthetic, CAI 1.0, artificial introns); unc-119(+)].* TH251, *unc-119(ed3) III; ddIs33[yfp::gpr-1(synthetic, CAI 0.3, artificial introns); unc-119(+)].* TH252, *unc-119(ed3) III; ddIs34[yfp::gpr-1(synthetic, CAI 0.6, artificial introns); unc-119(+)].* TH253, *unc-119(ed3) III; ddIs35[gpr-1 (synthetic, CAI 1.0, artificial introns); unc-119(+)].* TH315, *unc-119(ed3); ddEx23[sas-4::gfp(synthetic, CAI 0.3, endogenous introns); unc-119(+)].* TH319, *unc-119(ed3); ddEx24[sas-4::gfp(synthetic, CAI 0.6, endogenous introns) unc-119(+)].* TH329, *unc-119(ed3); ddIs62[gfp::air-1(K07C11.2, synthetic, CAI 1.0 artificial introns); unc-119(+)].* TH334, *unc-119(ed3) III; ddIs63[gfp::air-1(K07C11.2, synthetic, CAI 0.7 artificial introns); unc-119(+)].*

YFP–GPR-1 imaging. For phenotypic characterization (**Figs. 1a** and **2a**), we imaged embryos using an inverted widefield epifluorescence microscope (Axiovert, Zeiss) using a standard GFP filter set (Chroma ET) and an Orca ER camera (Hamamatsu).

For intensity measurements (**Fig. 2b**) we imaged embryos at ~20 °C with an Olympus IX71 spinning disc setup using a 100×, 1.35 NA Oil Iris UPlan Apochromat objective and Yokogawa scan head CSU10. We used a 488-nm laser for illumination. We acquired images with an Andor Ixon electron-multiplying charge-coupled device (EMCCD) 512 × 512 pixel camera at 100-ms exposure. We dissected hermaphrodites in 0.1 M NaCl buffer to retrieve embryos and mounted them onto 2% agarose pads. We processed images with ImageJ (macBiophotonics). We drew an ellipsoid shape around the embryos and measured the mean signal intensity. We set the mean signal intensity of embryos expressing no YFP as zero.

Laser ablation. We performed laser ablation experiments using a highly focused ultraviolet laser beam as described previously[21]. We observed embryos using differential interference contrast (DIC), fired five to ten shots (57 pulses, 450 Hz) at the spindle midzone just after anaphase onset (identified by the disappearance of the metaphase plate) and observed centrosome movements[22]. We manually tracked centrosomes using ImageJ.

RNA interference. We performed RNAi by feeding as described previously[9]. We used a *gpr-1* ORFeome feeding clone (11008-b4) for *gpr-1/2* RNAi[23]. For *yfp* RNAi, we PCR-amplified the *yfp* gene. We TA cloned the PCR product into L4440 and transformed the resulting plasmid into *Escherichia coli* HT115 as described previously[24,25].

Image acquisition for SAS-4 and intensity measurements for AIR-1. We imaged embryos using an upright microscope (Axiovision imager 2e, Zeiss) in GFP optics. We acquired images with a Hamamatsu Orca ER camera. We analyzed images with ImageJ.

Western blotting. We collected whole worms in M9 buffer and froze them in liquid nitrogen. We then diluted the M9 buffer containing *C. elegans* 1:2 with SDS loading buffer and incubated the mixtures for 5 min at 95 °C. For the GPR-1/2 and AIR-1 blot, we loaded the samples on 4–12% Bis-Tris gels (NuPAGE; Invitrogen) and ran them at 100 V for 90 min using MOPS buffer (NuPAGE). For the SAS-4 blot, we loaded the samples on 3–8% tris-acetate gels (NuPAGE) and ran them at 100 V for 90 min using tris-acetate SDS running buffer (Novex; Invitrogen). We blotted the gels on a nitrocellulose membrane using NuPAGE transfer buffer at 50 V for 1.5 h.

We then blocked the membrane with 5% milk powder in PBS plus Tween 20 (PBST; 10× PBST was prepared from 80 g NaCl, 2 g KCl, 14.4 g Na_2HPO_4, 2.4 g KHPO, 1 l H_2O; pH adjusted to 7.4 with HCl, 0.1% Tween 20) for 1 h at room temperature (25 °C). We incubated primary antibodies with the membrane at 4 °C overnight and secondary antibodies at room temperature for 1 h. We used ECL reagent (GE Healthcare) and Amersham Hyperfilm ECL (GE Healthcare) for detection.

We used the following blotting conditions. For tubulin detection, we used as primary antibody anti-alpha-tubulin DM1A (Sigma-Aldrich, 1:7,000) diluted in 5% milk powder in PBST, and as secondary antibody goat-anti-mouse IgG coupled to horse radish peroxidase (Bio-Rad, 1:5,000) diluted in 5% milk powder.

For GPR-1/2 detection, we used as primary antibody mouse anti-GFP (Roche; 1:5,000) diluted in 5% milk powder in PBST, and as secondary antibody anti-mouse IgG coupled to horse radish peroxidase (Bio-Rad, 1:5,000) diluted in 5% milk powder.

For GFP–SAS-4 detection, we used as primary antibody mouse anti-GFP (Roche) 1:5,000 diluted in 5% milk powder in PBST, and as secondary antibody anti-mouse IgG coupled to horse radish peroxidase (Bio-Rad, 1:5,000) diluted in 5% milk powder.

For AIR-1–GFP/GFP–AIR detection, we used as primary antibody mouse anti-GFP (Roche) 1:5,000 diluted in 5% milk powder in PBST, and as secondary antibody anti-mouse IgG coupled to horse radish peroxidase (Bio-Rad, 1:5,000) diluted in 5% milk powder.

11. Wu, G., Bashir-Bello, N. & Freeland, S.J. *Protein Expr. Purif.* **47**, 441–445 (2006).
12. Carbone, A., Zinovyev, A. & Kepes, F. *Bioinformatics* **19**, 2005–2015 (2003).
13. Henschel, A., Buchholz, F. & Habermann, B. *Nucleic Acids Res.* **32**, W113–W120 (2004).
14. Fire, A.A.J., Kelly, W., Harfe, B., Kostas, S., Hsieh, J., Hsu, M. & Xu, S. in *GFP: Freen Fluorescent Protein Strategies and Applications* (ed., Chalfie, M.) 153–168 (John Wiley and Sons New York, 1998).
15. Praitis, V., Casey, E., Collar, D. & Austin, J. *Genetics* **157**, 1217–1226 (2001).
16. Hofacker, I.L. *et al. Monatsh. Chem.* **125**, 167–188 (1994).
17. Riddle, D., Blumenthal, T., Meyer, B. & Priess, J. *C. elegans II* (Cold Spring Harbor Laboratory Press, 1997).
18. Hebsgaard, S.M. *et al. Nucleic Acids Res.* **24**, 3439–3452 (1996).
19. Brunak, S., Engelbrecht, J. & Knudsen, S. *J. Mol. Biol.* **220**, 49–65 (1991).
20. Brenner, S. *Genetics* **77**, 71–94 (1974).
21. Grill, S.W., Howard, J., Schaffer, E., Stelzer, E.H. & Hyman, A.A. *Science* **301**, 518–521 (2003).
22. Grill, S.W., Gonczy, P., Stelzer, E.H. & Hyman, A.A. *Nature* **409**, 630–633 (2001).
23. Rual, J.F. *et al. Genome Res.* **14**, 2162–2168 (2004).
24. Fraser, A.G. *et al. Nature* **408**, 325–330 (2000).
25. Kamath, R.S. *et al. Nature* **421**, 231–237 (2003).

doi:10.1038/nmeth.1565