

# Smoking is associated with mosaic loss of chromosome Y

Jan P. Dumanski,<sup>1,2\*</sup> Chiara Rasi,<sup>1,2</sup> Mikael Lönn,<sup>3</sup> Hanna Davies,<sup>1,2</sup> Martin Ingelsson,<sup>4</sup> Vilmantas Giedraitis,<sup>4</sup> Lars Lannfelt,<sup>4</sup> Patrik K. E. Magnusson,<sup>5</sup> Cecilia M. Lindgren,<sup>6,7</sup> Andrew P. Morris,<sup>6,8</sup> David Cesarini,<sup>9</sup> Magnus Johannesson,<sup>10</sup> Eva Tiensuu Janson,<sup>11</sup> Lars Lind,<sup>11</sup> Nancy L. Pedersen,<sup>5</sup> Erik Ingelsson,<sup>2,11</sup> Lars A. Forsberg<sup>1,2\*</sup>

Tobacco smoking is a risk factor for numerous disorders, including cancers affecting organs outside the respiratory tract. Epidemiological data suggest that smoking is a greater risk factor for these cancers in males compared with females. This observation, together with the fact that males have a higher incidence of and mortality from most non-sex-specific cancers, remains unexplained. Loss of chromosome Y (LOY) in blood cells is associated with increased risk of nonhematological tumors. We demonstrate here that smoking is associated with LOY in blood cells in three independent cohorts [TwinGene: odds ratio (OR) = 4.3, 95% confidence interval (CI) = 2.8 to 6.7; Uppsala Longitudinal Study of Adult Men: OR = 2.4, 95% CI = 1.6 to 3.6; and Prospective Investigation of the Vasculature in Uppsala Seniors: OR = 3.5, 95% CI = 1.4 to 8.4] encompassing a total of 6014 men. The data also suggest that smoking has a transient and dose-dependent mutagenic effect on LOY status. The finding that smoking induces LOY thus links a preventable risk factor with the most common acquired human mutation.

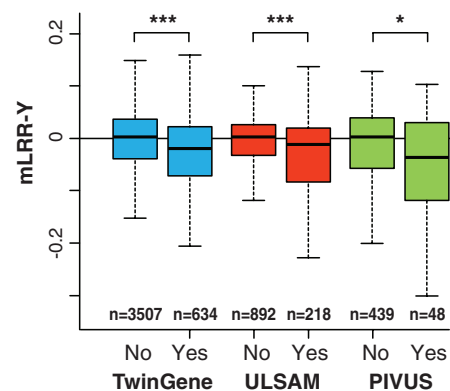
**T**obacco smoking killed ~100 million people during the 20th century and is projected to kill ~1 billion people during the current century, assuming that the current frequency of smoking is retained (1, 2). Lung cancer is the prime cause of cancer-associated death in relation to smoking. However, smoking is also a risk factor for tumors outside the respiratory tract, and these are more common in males than females [hazard ratio in males: 2.2, 95% confidence interval (CI) = 1.7 to 2.8; in females: 1.7, 95% CI = 1.4–2.1] (2). Moreover, males have a higher incidence and mortality from most non-

sex-specific cancers, disregarding smoking status, and this fact is largely unexplained by known risk factors (3, 4). A recent analysis of noncancerous blood cells revealed that a male-specific chromosomal aberration, acquired mosaic loss of chromosome Y (LOY), is associated with an increased risk of nonhematological tumors among aging males (5).

Here, we analyzed possible causes of LOY by studying 6014 men from three independent prospective cohorts—TwinGene,  $n = 4373$  (6, 7); Uppsala Longitudinal Study of Adult Men (ULSAM),  $n = 1153$  (8); and Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS),  $n = 488$  (9)—from which comprehensive epidemiological records are available (tables S2 to S4). We included the following environmental, lifestyle, and clinical factors in the analyses: smoking, age, hypertension, exercise habits, diabetes, body mass index, low-density lipoprotein cholesterol, high-density lipoprotein cholesterol, education level, and alcohol intake. We also included genotyping quality as a confounder in the regression analyses, to adjust for possible influence of experimental noise. Similar definitions of factors were used in all cohorts, as outlined in tables S2 to S5 and described in detail in the materials and methods section of the supplementary materials. Estimation of LOY was based on single-

nucleotide polymorphism (SNP)-array data using the 2.5MHumanOmni and HumanOmniExpress beadchips in the ULSAM and PIVUS/TwinGene studies, respectively (fig. S1). The estimation of the degree of mosaicism and scoring of LOY was undertaken using the continuous median logR ratio (mLRR-Y) estimate, calculated from SNP-array data as the median of the logR ratio of all SNP probes within the male-specific part of chromosome Y (MSY), as described previously (5). An mLRR-Y estimate close to zero indicates a normal chromosome Y state, whereas more negative mLRR-Y values denote an increasing level of blood cells with LOY. To facilitate comparisons between the three cohorts, we corrected the mLRR-Y values for all participants, using cohort-specific correction constants, as explained in the supplementary materials (figs. S1 and S2).

LOY was by far the most common postzygotic mutation found in the three cohorts. The age range at sampling in ULSAM and PIVUS was 70.7 to 83.6 years and 69.8 to 70.7 years, respectively, and we found LOY in 12.6% of ULSAM participants and 15.6% of PIVUS participants (figs. S3 and S4). The age range at sampling in TwinGene was 48 to 93 years, and the frequency of LOY in the entire cohort was 7.5% (fig. S5).



**Fig. 1. The association between smoking status and the level of LOY (i.e., mLRR-Y) in three independent cohorts.** In all cohorts, these unadjusted analyses indicate that the current smokers (Yes) (table S5) had a significantly higher degree of mosaic LOY in blood, compared with noncurrent smokers (No), composed of never-smokers and previous smokers. \*\*\* $P < 0.001$ ; \* $P < 0.05$  (Kolmogorov-Smirnov tests: TwinGene,  $D = 0.15$ ,  $P = 1.131 \times 10^{-11}$ ; ULSAM,  $D = 0.15$ ,  $P = 0.0006$ ; PIVUS,  $D = 0.23$ ,  $P = 0.0203$ ). The definitions used for LOY scoring and the entire ranges of mLRR-Y data observed in each cohort are shown in figs. S3 to S5.

<sup>1</sup>Department of Immunology, Genetics and Pathology, Uppsala University, Uppsala, Sweden. <sup>2</sup>Science for Life Laboratory, Uppsala University, Uppsala, Sweden. <sup>3</sup>Södertörn University, School of Life Sciences, Biology, Huddinge, Sweden. <sup>4</sup>Department of Public Health and Caring Sciences, Uppsala University, Uppsala, Sweden. <sup>5</sup>Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden. <sup>6</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. <sup>7</sup>Broad Institute of Massachusetts Institute of Technology and Harvard University, Cambridge, Massachusetts, USA. <sup>8</sup>Department of Biostatistics, University of Liverpool, Liverpool, UK. <sup>9</sup>Center for Experimental Social Science, New York University, New York, NY 10012, USA. <sup>10</sup>Department of Economics, Stockholm School of Economics, Stockholm, Sweden. <sup>11</sup>Department of Medical Sciences, Uppsala University, Uppsala, Sweden.

\*Corresponding Author. E-mail: jan.dumanski@igp.uu.se (J.P.D.); lars.forsberg@igp.uu.se (L.A.F.).

However, in TwinGene participants aged 70 years or older, 15.4% had LOY, which is similar to the LOY frequency in the other cohorts in the same age range. In TwinGene participants younger than 70 years, only 4.1% were affected by LOY. Thresholds for LOY scoring, at the lower 99% confidence limit of the distributions of experimental mLRR-Y variation (TwinGene =  $-0.1324$ , ULSAM =  $-0.1024$ , and PIVUS =  $-0.1182$ ), were used for frequency calculations, as explained in the supplementary materials and figs. S3 to S5. At this degree of mosaicism, ~10% of the analyzed nucleated blood cells from each sample are affected by LOY. Mosaic LOY in blood detected with SNP arrays was validated using whole-genome next-generation sequencing (NGS) in 100 random participants in the ULSAM cohort. There was 100% concordance in LOY scoring between results from SNP array and NGS data (5).

We found a strong association between smoking and LOY status in the three independent cohorts. Current smokers had a significantly higher degree of LOY mosaicism, compared with nonsmokers and past smokers, in unadjusted models (Fig. 1) and in the multivariable models adjusting for the above-mentioned potential confounders (table S1, analysis of covariance: TwinGene  $F_{(1,1666)} = 45.4$ ,  $P = 2.225 \times 10^{-11}$ ; ULSAM  $F_{(1,968)} = 17.6$ ,  $P = 2.984 \times 10^{-5}$ ; PIVUS  $F_{(1,385)} = 9.1$ ,  $P = 0.0028$ ). Apart from smoking, the only other factor significantly associated with LOY was age, which was observed in TwinGene, with higher degree of LOY in older participants. The age range in ULSAM and PIVUS was narrower (see above), which may explain why age had no effect on LOY in these regression models (tables S1 and S5). To assess the risk of LOY in blood cells of smokers, we used logistic regression adjusting for the same confounders as in table S1. Participants were LOY-scored as 1 or 0 before these analyses based on the continuous mLRR-Y estimate, using the same threshold as for the estimations of LOY frequencies—i.e., the lower 99% CIs of the experimentally induced mLRR-Y variation (figs. S3 to S5). The adjusted odds ratio (OR) estimates from logistic regressions were highest in TwinGene (OR = 4.3, 95% CI = 2.8 to 6.7), followed by PIVUS (OR = 3.5, 95% CI = 1.4 to 8.4) and ULSAM (OR = 2.4, 95% CI = 1.6 to 3.6). The corresponding unadjusted ORs are given in table S6. Based on these calculations, we estimate that current smokers in the studied cohorts had a 2.4 to 4.3 times greater risk of displaying LOY compared with nonsmokers. Furthermore, among the current smokers in the large TwinGene cohort, we found a strong dose-response effect with more LOY in heavy smokers—i.e., smokers with LOY had been smoking significantly more pack-years compared to smokers without LOY (Fig. 2C).

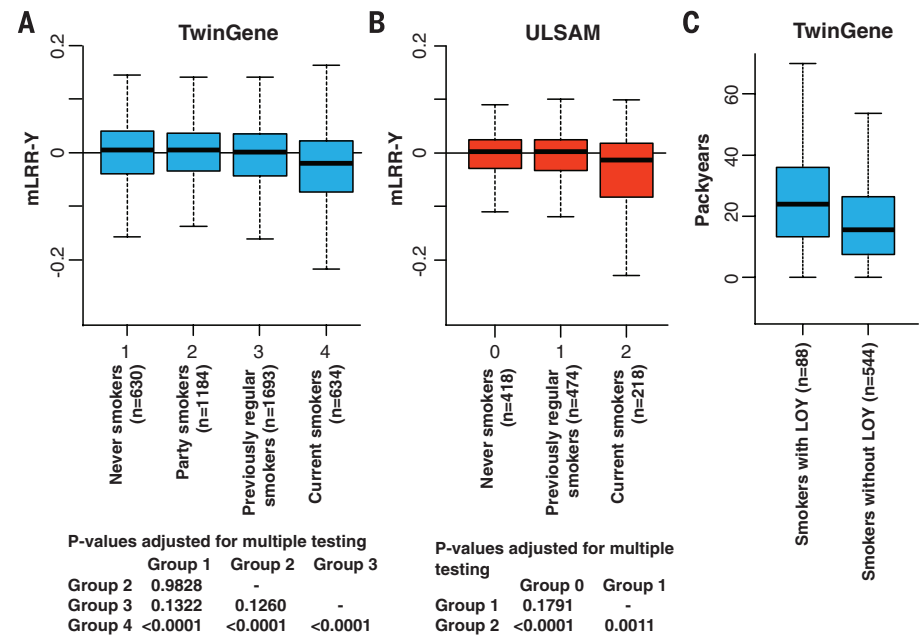
Our results suggest that the association between smoking and LOY is valid for current smokers only (Fig. 2, A and B). Previous epidemiological studies showed that smoking cessation at any age is associated with dramatically reduced death rates. For smokers who quit at 25 to 34 years of age, survival was nearly identical with those who had never smoked (2, 10). We analyzed the level of LOY after smoking cessa-

tion in the ULSAM and the TwinGene cohorts by using LOY data in past regular smokers (table S5). No difference in LOY frequency between never-smokers and previously regular smokers was found (Fig. 2, A and B). One possible explanation for these results is that the previous smokers with LOY died off faster than the rest of the cohort. Another, and perhaps more likely, explanation is that LOY is induced and sustained by smoking and that LOY is a dynamic and reversible process.

Whether the LOY induced by smoking plays a direct role in cancer is unclear. One hypothesis is that smoking is clastogenic—that is, it induces many chromosomal abnormalities, including an incidental loss of the Y chromosome. In this scenario, LOY would be a neutral passenger mutation and a reporter of a general tendency of chromosome missegregation in mitosis, which is enhanced by smoking and associated with risk for cancer and mortality. A second hypothesis is that LOY in blood cells is a causative factor in cancer development, possibly through effects on tumor immunosurveillance (11). We conducted a preliminary test of this hypothesis by investigating possible functional consequences of LOY in sorted blood cells from three ULSAM survivors scored with LOY and still alive at the age of 91 years. We sorted cells from three compartments (granulocytes, CD4+ T-lymphocytes, and CD19+ B-lymphocytes) in subjects that displayed LOY in earlier serial analyses of whole blood, performed

at four time points during two decades (fig. S6). The three subjects were free from cancer diagnoses at the time of blood collection at 91 years. The data from these experiments are tantalizing because (i) the percentage of cells with LOY differed between different compartments of the hematopoietic system, and (ii) ULSAM-1412 suggests that LOY might be an oligo-clonal process, because cells derived from myeloid and some (but not all) lymphoid progenitors display LOY. These preliminary results support the second hypothesis. If LOY were a phenotypically neutral passenger mutation, one would expect that LOY cells would be randomly distributed within all components of the hematological system. It was recently shown that LOY status of blood cells is associated with a higher risk for all-cause mortality as well as a higher risk for nonhematological cancers and that it can be considered as a biomarker of male carcinogenesis (5). We hypothesized that a disrupted tumor-immunosurveillance in LOY-affected cells could help explain the connection between LOY status of noncancerous blood cells and risk for tumors in other tissues (5, 11). These results also support the second hypothesis and the increasingly recognized view that chromosome Y carries many vital functions in biological processes beyond sex determination and sperm production (12–16).

Our results are consistent with a previously described dynamic nature of expanding-contracting noncancerous cell clones in blood affected with



**Fig. 2. Differences in degree of LOY between different smoking categories within TwinGene (A) and ULSAM (B), as defined in table S5.** In TwinGene (A), there was a significant difference between four smoking categories [analysis of variance (ANOVA):  $F_{(3,4137)} = 22.2$ ,  $P = 3.028 \times 10^{-14}$ ], and results from post hoc analysis adjusting  $P$  values for multiple testing using a Tukey post hoc test are displayed. In ULSAM (B), there was also a significant difference between three smoking categories (ANOVA:  $F_{(2,1107)} = 12.2$ ,  $P = 5.812 \times 10^{-6}$ ), and post hoc analysis is shown. In both cohorts, the current smokers had a significantly higher degree of LOY compared with all other categories. The average degree of LOY in the previously regular smokers was not significantly different from the average degree of LOY in the never-smokers in both cohorts. (C) A dose-response effect within current smokers in TwinGene, with men smoking the most pack-years also being associated with higher degree of LOY, as defined in fig. S5 (Kolmogorov-Smirnov test:  $D = 0.2244$ ,  $P = 0.0010$ ).

mosaic genetic aberrations—i.e., it appears that the relative frequency of cells from a cell clone can first increase and then decrease later in life (5, 17, 18). In the present analyses, LOY was detected in  $\geq 10\%$  of blood cells from about 15% of elderly males in three cohorts (figs. S3 to S5). The cell clones with LOY were likely detectable in our analyses because they are enriched due to an increased proliferative potential as a consequence of LOY, which is in agreement with chromosome Y containing tumor suppressor genes. Recent analysis of >8200 tumor-normal pairs suggest that two genes (*ZFY* and *UTY*, from the male-specific part of Y) have properties of tumor suppressors (19). Interestingly, both genes have homologs on chromosome X and escape X inactivation (19, 20). Moreover, other analyses of various tumor collections show that chromosome Y is lost from numerous types of tumors in frequencies ranging from 15 to 80% of cases (21–24). Thus, counting both LOY in noncancerous blood clones and in transformed tumor cells, nullisomy Y is among the most common, if not the most common, human mutation. The results presented here suggest that this aneuploidy, affecting 1.6% of the genome, is likely induced by smoking.

In conclusion, we show that LOY is more common in current smokers compared with non-current smokers in three cohorts (Fig. 1 and table S1), that the effect from smoking on LOY is dose dependent, and that this effect appears to be transient, as it disappears after smoking cessation (Fig. 2). Epidemiological observations suggest that smoking could be a greater risk factor for cancer outside the respiratory tract in males compared with females (2, 4, 10). Moreover, males have a higher incidence and mortality from most sex-unspecific cancers (3, 4). The molecular mechanisms behind these observations are not well understood, but LOY, being a male-specific, smoking-induced risk factor, could provide a missing link and help explain these sex differences.

#### REFERENCES AND NOTES

- P. Jha, *Natl. Rev.* **9**, 655–664 (2009).
- P. Jha et al., *N. Engl. J. Med.* **368**, 341–350 (2013).
- M. B. Cook, K. A. McGlynn, S. S. Devesa, N. D. Freedman, W. F. Anderson, *Cancer Epidemiol. Biomarkers Prev.* **20**, 1629–1637 (2011).
- G. Edgren, L. Liang, H. O. Adami, E. T. Chang, *Eur. J. Epidemiol.* **27**, 187–196 (2012).
- L. A. Forsberg et al., *Nat. Genet.* **46**, 624–628 (2014).
- P. Lichtenstein et al., *Twin Res. Hum. Genet.* **9**, 875–882 (2006).
- P. K. Magnusson et al., *Twin Res. Hum. Genet.* **16**, 317–329 (2013).
- H. Hedstrand, *Ups. J. Med. Sci. Suppl.* **19**, 1–61 (1975).
- L. Lind, N. Fors, J. Hall, K. Marttala, A. Stenborg, *Arterioscler. Thromb. Vasc. Biol.* **25**, 2368–2375 (2005).
- M. J. Thun et al., *N. Engl. J. Med.* **368**, 351–364 (2013).
- G. P. Dunn, A. T. Bruce, H. Ikeda, L. J. Old, R. D. Schreiber, *Nat. Immunol.* **3**, 991–998 (2002).
- B. Lemos, L. O. Araripe, D. L. Hartl, *Science* **319**, 91–93 (2008).
- B. Lemos, A. T. Branco, D. L. Hartl, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 15826–15831 (2010).
- A. G. Clark, *Nature* **508**, 463–465 (2014).
- D. W. Bellott et al., *Nature* **508**, 494–499 (2014).
- D. Cortez et al., *Nature* **508**, 488–493 (2014).
- L. A. Forsberg et al., *Am. J. Hum. Genet.* **90**, 217–228 (2012).
- L. A. Forsberg, D. Absher, J. P. Dumanski, *J. Med. Genet.* **50**, 1–10 (2013).
- T. Davoli et al., *Cell* **155**, 948–962 (2013).
- A. Schneider-Gädick, P. Beer-Romero, L. G. Brown, R. Nussbaum, D. C. Page, *Cell* **57**, 1247–1258 (1989).

- L. J. Zhang, E. S. Shin, Z. X. Yu, S. B. Li, *Chin. Med. J. (Engl.)* **120**, 2002–2005 (2007).
- N. O. Bianchi, *Mutat. Res.* **682**, 21–27 (2009).
- L. C. Silva Veiga, N. A. Bérnago, P. P. Reis, L. P. Kowalski, S. R. Rogatto, *Braz. J. Med. Biol. Res.* **45**, 172–178 (2012).
- P. H. Duijff, N. Schultz, R. Benezra, *Int. J. Cancer* **132**, 2316–2326 (2013).

#### ACKNOWLEDGMENTS

We thank K. Lindblad-Toh, R. M. Myers, C.-H. Heldin, D. H. Ledbetter, G. J. B. van Ommen, and U. Landegren for critical evaluation of the manuscript. This study was sponsored by the Swedish Cancer Society, the Swedish Research Council, the Swedish Heart-Lung Foundation and Science for Life Laboratory Uppsala to J.P.D. and the Olle Engqvist Byggnästars Foundation to L.A.F. Genotyping and next-generation sequencing were performed by the SNP&SEQ Technology Platform in Uppsala, Sweden, and supported by Wellcome Trust Grants WT098017, WT064890, and WT090532, Uppsala University, Uppsala University Hospital, the Swedish Research Council, and the Swedish Heart-Lung Foundation. The SNP&SEQ Technology Platform is part of Science for Life Laboratory at Uppsala University and supported as a national infrastructure by the Swedish Research Council. C.M.L. is a Wellcome Trust Research Career Development Fellow (086596/Z/08/Z). A.P.M. is a Wellcome Trust Senior Research

Fellow in Basic Biomedical Science. A.P.M. acknowledges funding from the Wellcome Trust under awards WT064890, WT090532, and WT098017. TwinGene was supported by the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254 and QL2-CT-2002-01254), NIH DK U01-066134, the Swedish Foundation for Strategic Research (SSF), and the Heart and Lung foundation no. 20070481. J.P.D. and L.A.F. are cofounders and shareholders in Cray Innovation AB, as well as co-inventors on Patent Application No. PCT/EP2014/071448, protecting the commercial applications of LOY for the assessment of cancer risk. Genetic variants detected in this study are available at the Database of Genomic Structural Variation (dbVar) under accession code nstd92 for ULSAM and PIVUS cohorts and accession code nstd104 for the TwinGene cohort.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/347/6217/81/suppl/DC1  
Materials and Methods  
Figs. S1 to S6  
Tables S1 to S6  
References (25–27)

4 July 2014; accepted 24 November 2014  
Published online 4 December 2014;  
10.1126/science.1262092

#### AGING

## Lysosomal signaling molecules regulate longevity in *Caenorhabditis elegans*

Andrew Folick,<sup>1</sup> Holly D. Oakley,<sup>2,3</sup> Yong Yu,<sup>2,3</sup> Eric H. Armstrong,<sup>4</sup> Manju Kumari,<sup>5\*</sup> Lucas Sanor,<sup>2</sup> David D. Moore,<sup>1,6</sup> Eric A. Ortlund,<sup>4</sup> Rudolf Zechner,<sup>5</sup> Meng C. Wang<sup>1,2,3,†</sup>

Lysosomes are crucial cellular organelles for human health that function in digestion and recycling of extracellular and intracellular macromolecules. We describe a signaling role for lysosomes that affects aging. In the worm *Caenorhabditis elegans*, the lysosomal acid lipase LIPL-4 triggered nuclear translocation of a lysosomal lipid chaperone LBP-8, which promoted longevity by activating the nuclear hormone receptors NHR-49 and NHR-80. We used high-throughput metabolomic analysis to identify several lipids in which abundance was increased in worms constitutively overexpressing LIPL-4. Among them, oleoylethanolamide directly bound to LBP-8 and NHR-80 proteins, activated transcription of target genes of NHR-49 and NHR-80, and promoted longevity in *C. elegans*. These findings reveal a lysosome-to-nucleus signaling pathway that promotes longevity and suggest a function of lysosomes as signaling organelles in metazoans.

Lysosomes contain acid hydrolytic enzymes, digesting macromolecules taken up by endocytosis and recycling dysfunctional cellular components during autophagy (1). Lysosomal deficiency is associated with human diseases. For example, loss of human lysosomal acid lipase, *LIPA*, results in severe systemic metabolic malfunction known as infantile Wolman disease (2). Here, we explored how lysosomes might generate signaling molecules that regulate aging by influencing nuclear transcription.

We analyzed a *Caenorhabditis elegans* longevity-promoting lipase, LIPL-4, which has sequence and functional similarities with human *LIPA* (fig. S1). Lipid hydrolase activity was decreased in *lipl-4(tm4417)* loss-of-function mutants at pH 4.5 but not at pH 7.4 (Fig. 1A). FLAG-tagged LIPL-4 protein was localized to intestinal lysosomes (Fig. 1, B to D, and fig. S2). Increased *lipl-4* expression is associated with longevity (3). A transgenic strain (*lipl-4 Tg*) that constitutively expressed *lipl-4* in the intestine had 55% mean life-span increase compared with wild-type (WT) animals (Fig. 1E and table S1) and delayed age-related decline of physical activity (fig. S3A). Constitutive expression of LIPL-4 without the signal peptide (*lipl-4 Tg no SP*), which was not targeted to the lysosome, caused little extension of life span (fig. S4 and table S1), which suggests that the lysosomal activity of LIPL-4 is essential for its longevity effect.

To elucidate whether lipid signals are affected by the LIPL-4 lipase, we examined fatty acid-binding proteins (FABPs) that are intracellular lipid chaperones shuttling lipid molecules between cellular

<sup>1</sup>Program in Developmental Biology, Baylor College of Medicine, Houston, TX 77030, USA. <sup>2</sup>Huffington Center on Aging, Baylor College of Medicine, Houston, TX 77030, USA.

<sup>3</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX 77030, USA. <sup>4</sup>Department of Biochemistry, Discovery and Developmental Therapeutics, Winship Cancer Institute, Emory University School of Medicine, Atlanta, GA 30322, USA. <sup>5</sup>Institute of Molecular Biosciences, University of Graz, Graz, A-8010, Austria.

<sup>6</sup>Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX 77030, USA.

\*Present address: Division of Endocrinology, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA 02215, USA.

†Corresponding author. E-mail: wmeng@bcm.edu