

A and B). Conversely, increased expression of WT CLIP-170-1 over endogenous CLIP-170 led to elevated dendritic complexity, as previously shown (16), whereas expression of mutant CLIP-170-1 did not (Fig. 4, C and D). In N2A cells, WT and FEED1 mutant CLIP-170-1 were expressed at similar levels (fig. S7). Live imaging showed that they localized to MT plus ends similarly (Fig. 4, E and F; fig. S8, A to C; and movie S10) and that the mutant did not alter MT dynamics (Fig. 4G; fig. S8, D and E; and movies S10 and S11). Thus, CLIP-170 interactions with formins play an important role in coordinating MT and actin dynamics to regulate neuronal process formation.

Here we have shown that CLIP-170 interacts tightly with formins to substantially increase both the rate of actin filament elongation and the duration of elongation in the presence of CP. CLIP-170 is part of a mechanism that enables growing MT plus ends to trigger rapid assembly of actin filaments in vitro, directly linking MT and actin dynamics. This mechanism was consistent in a physiological setting, where EB1 and CLIP-170 colocalized on MT plus ends, as well as with previous reports that growing MT plus ends survey the actin-rich cortex (10) and that ~10% of mDial puncta in cells colocalize with MT plus ends (32). In neurons, CLIP-170 interactions with formins were required for proper dendritic branching. Similar mechanisms may explain the colocalization and cofunctioning of CLIP-170 and mDial in phagocytic cup formation (5) and reduced actin-based protrusive activity in neuronal growth cones after CLIP-170 silencing (18, 19, 35, 36).

REFERENCES AND NOTES

1. P. Forscher, S. J. Smith, *J. Cell Biol.* **107**, 1505–1516 (1988).
2. O. C. Rodriguez et al., *Nat. Cell Biol.* **5**, 599–609 (2003).
3. M. A. Chesarone, A. G. DuPage, B. L. Goode, *Nat. Rev. Mol. Cell Biol.* **11**, 62–74 (2010).
4. T. M. Svitkina et al., *J. Cell Biol.* **160**, 409–421 (2003).
5. E. Lewkowicz et al., *J. Cell Biol.* **183**, 1287–1298 (2008).
6. A. Sampathkumar et al., *Plant Cell* **23**, 2302–2313 (2011).
7. S. G. Martin, W. H. McDonald, J. R. Yates III, F. Chang, *Dev. Cell* **8**, 479–491 (2005).
8. S. G. Martin, S. A. Rincón, R. Basu, P. Pérez, F. Chang, *Mol. Biol. Cell* **18**, 4155–4167 (2007).
9. D. M. Suter, P. Forscher, *J. Neurobiol.* **44**, 97–113 (2000).
10. W. C. Salmon, M. C. Adams, C. M. Waterman-Storer, *J. Cell Biol.* **158**, 31–37 (2002).
11. C. H. Coles, F. Bradke, *Curr. Biol.* **25**, R677–R691 (2015).
12. M. Chesarone, C. J. Gould, J. B. Moseley, B. L. Goode, *Dev. Cell* **16**, 292–302 (2009).
13. D. R. Kovar, E. S. Harris, R. Mahaffy, H. N. Higgs, T. D. Pollard, *Cell* **124**, 423–435 (2006).
14. A. Akhmanova et al., *Cell* **104**, 923–935 (2001).
15. K. C. Slep, R. D. Vale, *Mol. Cell* **27**, 976–991 (2007).
16. D. Neukirchen, F. Bradke, *J. Neurosci.* **31**, 1528–1538 (2011).
17. R. Dixit et al., *Proc. Natl. Acad. Sci. U.S.A.* **106**, 492–497 (2009).
18. J.-H. Weng et al., *Nat. Chem. Biol.* **9**, 636–642 (2013).
19. R. Beaven et al., *Mol. Biol. Cell* **26**, 1491–1508 (2015).
20. M. Chesarone-Cataldo et al., *Dev. Cell* **21**, 217–230 (2011).
21. J. A. Eskin, A. Rankova, A. B. Johnston, S. L. Alioto, B. L. Goode, *Mol. Biol. Cell* **27**, 828–837 (2016).
22. J. R. Kuhn, T. D. Pollard, *Biophys. J.* **88**, 1387–1402 (2005).
23. C. J. Gould et al., *Curr. Biol.* **21**, 384–390 (2011).
24. D. Breitsprecher et al., *Science* **336**, 1164–1168 (2012).
25. J. B. Moseley et al., *Mol. Biol. Cell* **15**, 896–907 (2004).
26. S. H. Zigmund et al., *Curr. Biol.* **13**, 1820–1823 (2003).
27. E. S. Harris, F. Li, H. N. Higgs, *J. Biol. Chem.* **279**, 20076–20087 (2004).
28. M. A. Wear, J. A. Cooper, *Trends Biochem. Sci.* **29**, 418–428 (2004).
29. J. P. Bombardier et al., *Nat. Commun.* **6**, 8707 (2015).
30. S. Shekhar et al., *Nat. Commun.* **6**, 8730 (2015).
31. P. Pierre, J. Scheel, J. E. Rickard, T. E. Kreis, *Cell* **70**, 887–900 (1992).
32. Y. Wen et al., *Nat. Cell Biol.* **6**, 820–830 (2004).
33. J. Gaillard et al., *Mol. Biol. Cell* **22**, 4575–4587 (2011).
34. F. Bartolini et al., *J. Cell Biol.* **181**, 523–536 (2008).
35. L. Swiech et al., *J. Neurosci.* **31**, 4555–4568 (2011).
36. C. Erck et al., *Proc. Natl. Acad. Sci. U.S.A.* **102**, 7853–7858 (2005).
37. K. T. Applegate et al., *J. Struct. Biol.* **176**, 168–184 (2011).

ACKNOWLEDGMENTS

We thank S. Paradis for guidance on experiments using neurons, D. Breitsprecher for pioneering the MT-actin co-reconstitution system, J. Gelles for guidance on single-molecule analysis, S. Jansen for guidance with cell culture, H. Higgs for providing INF1 and INF2 proteins, and L. Cassimeris for providing pGFP-EB1. This research was supported by NIH grant GM083137 to B.L.G. and Brandeis NSF Materials Research Science and Engineering Center grant 142038.

J.L.H.-R. was supported in part by a fellowship from the Leukemia and Lymphoma Society and in part by NIH training grant T32NS007292. J.L.H.-R. and B.L.G. designed the experiments and wrote the manuscript, J.L.H.-R. performed the experiments and data analysis, J.A.E. performed data analysis, A.R. built reagents and performed preliminary experiments, and K.K. performed overexpression analysis in neurons and assisted in other neuronal work. We declare no conflicts of interest. The supplementary materials contain additional data.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6288/1004/suppl/DC1
Materials and Methods
Figs. S1 to S9
References (38–48)
Movies S1 to S11

29 December 2015; accepted 5 April 2016
10.1126/science.aaf1709

GENE EVOLUTION

Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals

Xun Lan^{1,3*} and Jonathan K. Pritchard^{1,2,3*}

Gene duplication is a fundamental process in genome evolution. However, most young duplicates are degraded by loss-of-function mutations, and the factors that allow some duplicate pairs to survive long-term remain controversial. One class of models to explain duplicate retention invokes sub- or neofunctionalization, whereas others focus on sharing of gene dosage. RNA-sequencing data from 46 human and 26 mouse tissues indicate that subfunctionalization of expression evolves slowly and is rare among duplicates that arose within the placental mammals, possibly because tandem duplicates are coregulated by shared genomic elements. Instead, consistent with the dosage-sharing hypothesis, most young duplicates are down-regulated to match expression levels of single-copy genes. Thus, dosage sharing of expression allows for the initial survival of mammalian duplicates, followed by slower functional adaptation enabling long-term preservation.

Gene duplications are a major source of new genes and ultimately of new biological functions (1). However, recently arisen gene duplicates tend to be functionally redundant and thus susceptible to loss-of-function mutations that degrade one of the copies into a pseudogene. The average half-life of new primate duplicates has been estimated at just 4 million years (2). This raises the question of what evolutionary forces govern the persistence of young duplicates.

Various models have been proposed to understand why some duplicate pairs do survive over long evolutionary time scales (3). Dosage-balance models focus on the importance of maintaining correct stoichiometric ratios in gene expression between different genes (4–6) and likely explain

how gene copies are maintained after whole-genome duplication (WGD), because subsequent gene losses would disrupt dosage balance (6, 7).

Alternatively, functional partitioning of duplicates can occur, either by neofunctionalization (one copy gains new functions) or subfunctionalization (the copies divide the ancestral functions between them). The duplication-degeneration-complementation (DDC) model proposes that complementary degeneration of regulatory elements causes the two copies to be expressed in different tissues, such that both copies are required to provide the overall expression of the ancestral gene (8). Similarly, neofunctionalization of expression could lead to one gene copy gaining function in a tissue where the parent gene was not expressed. Functional divergence may also occur at the protein level (9), but this is thought to be a slow process, with initial divergence more often occurring through changes in gene regulation (10).

It is currently unclear which factors are most important for long-term survival of gene duplications in mammals, where most duplications arise

through segmental duplications or retrotranspositions that increase copy numbers of just one or a few genes. These small-scale duplications most likely disrupt overall dosage balance and should thus favor gene loss rather than preservation.

We therefore set out to investigate whether gene expression data across tissues in human and mouse support either model of duplicate preservation. We analyzed RNA-sequencing (RNA-seq) data from 10 individuals for each of 46 diverse human tissues collected by the Genotype-Tissue Expression (GTEx) project (11) and replicated our main conclusions using RNA-seq from 26 diverse mouse tissues (12).

We developed a computational pipeline to identify duplicate gene pairs in the human genome (13). After excluding annotated pseudogenes, we identified 1444 high-confidence reciprocal best-hit duplicate gene pairs with >80% alignable coding sequence and >50% average sequence identity. We used synonymous divergence, d_s , as a proxy for divergence time, while noting that divergence of some gene pairs may be affected by nonallelic homologous gene conversion in young duplicates. Additional analyses using the phylogenetic distribution of duplicates to refine date estimates were highly concordant with results based on d_s alone (figs. S5 to S7). We estimate that d_s for duplicates that arose at the time of the human-mouse split averages ~0.45 and that most pairs with $d_s > \sim 0.7$ predate the origin of the placental mammals (figs. S3 and S4). Thus, most of our analysis focuses on duplicates that likely arose within the mammalian lineage and postdate the early vertebrate whole-genome duplications.

Accurate measurement of expression in gene duplicates can be challenging if RNA-seq reads map well to both gene copies. Mapping may also be biased if the two copies have differential homology with other genomic locations. To overcome these challenges, we estimated expression ratios using only paralogous positions for which reads from both copies would map uniquely to the correct gene (13). This approach is related to a method for measuring allele-specific expression (14). These strict criteria mean that some very young genes are excluded from our expression analyses as unmappable, but, for the remaining genes, simulations show that our pipeline yields highly accurate, unbiased estimates of expression ratios (fig. S1).

This read-mapping pipeline allowed us to classify duplicates into categories on the basis of their coexpression patterns (13). First, within each pair, we classified the gene with higher overall expression as the “major” gene and its partner as the “minor” gene. We then defined a gene pair as potentially sub- or neofunctionalized if both the major and minor copy are significantly more highly expressed than the other in at least one tissue each (at least a twofold difference and $P < 0.001$ with paired t test) (Fig. 1A). We refer to pairs with consistent asymmetry as asymmetrically expressed duplicates (AEDs) if the major gene is significantly more highly expressed in at least 1/3 of tissues where either gene is expressed and not expressed at a significantly lower level than its partner in any tissue (Fig. 1B). The remaining duplicates were classified as having no difference, although many of these pairs show weaker levels of asymmetry.

Few duplicate pairs show evidence of sub- or neofunctionalization of expression (Fig. 2, A to C). Moreover, most gene pairs with such patterns are very old, dating to before the emergence of the placental mammals: For duplicates with $d_s < 0.7$, just 15.2% of duplicates are classified as potentially sub- or neofunctionalized in expression. Given that even modest variation in expression profiles across tissues would meet our criteria for subfunctionalization, the fraction of truly subfunctionalized duplicates may be even lower.

We also found similar levels of potential subfunctionalization in a mouse data set (12) that, unlike GTEx, includes fetal tissues (fig. S14). We examined whether subfunctionalization might instead be occurring through differential splicing of exons; however, we found little evidence for this (fig. S20). Last, we hypothesized that subfunctionalization might be more prevalent in gene pairs with higher tissue specificity (because they likely have more tissue-specific enhancers), but this is not the case (fig. S13).

Although relatively scarce, the genes identified as potentially subfunctionalized exhibit systematic differences from other duplicates. First, subfunctionalized gene pairs are expected to be under stronger selective constraint than genes without diverged expression, because the two copies are not functionally redundant. Consistent with this, we find that putatively subfunctionalized genes tend to have a higher fraction of rare variants in human polymorphism data (15) ($P = 2 \times 10^{-5}$ for missense mutations; Kolmogorov-Smirnov test) (Fig. 2D). Second, we hypothesized that if subfunctionalized genes have distinct

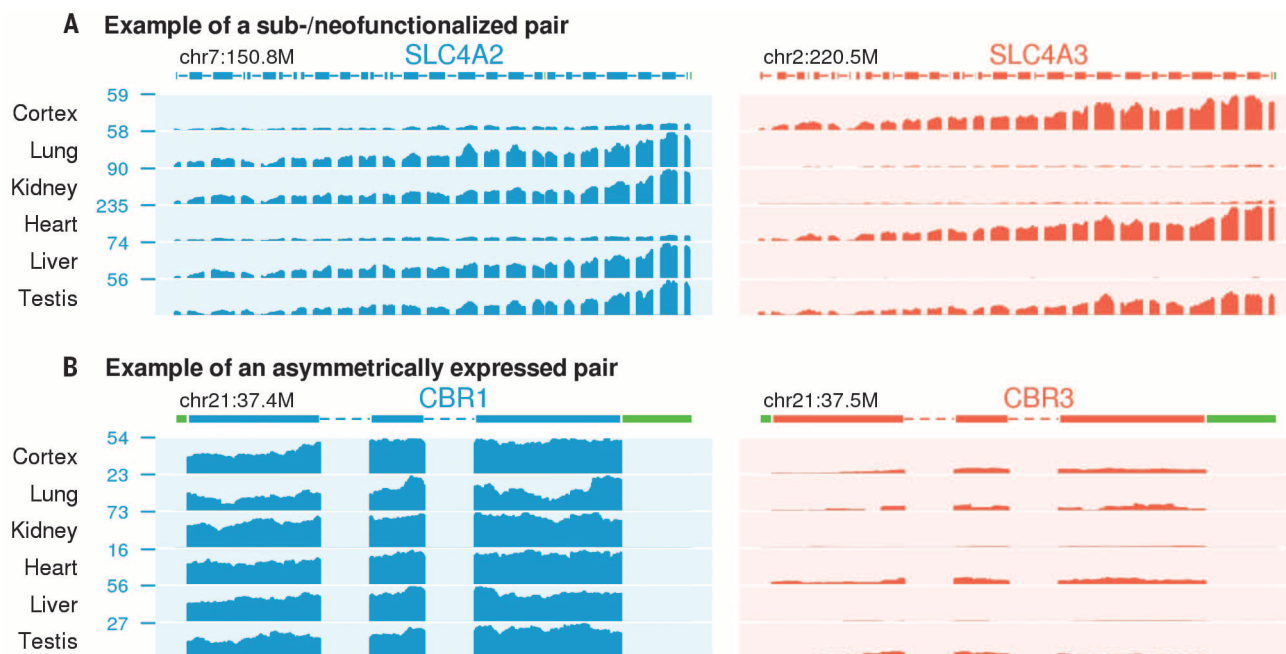


Fig. 1. Expression profiles of duplicate genes. (A) A gene pair whose expression profile is consistent with sub- or neofunctionalization: i.e., each gene is significantly more highly expressed than the other in at least one tissue. (B) An asymmetrically expressed gene pair. Expression of CBR1 exceeds expression of CBR3 in all tissues. Introns shortened for display purposes. The y axis shows read depth per billion mapped reads. Green regions in the gene models are unmappable.

functions, then they may be associated with distinct genetic diseases. Examining a database of gene associations with disease (16), we found a correlation between the degree of expression subfunctionalization and the number of diseases reported for only one member of the gene pair ($P = 5 \times 10^{-12}$, controlling for relevant covariates; Wald test) (Fig. 2E and table S3).

In sharp contrast to the expectations of subfunctionalization, many duplicate pairs exhibit systematically biased expression, as seen in some species after whole-genome duplication (17). Across all duplicate pairs, the mean expression of the less-expressed gene is 40% that of its duplicate (Fig. 2, B and C) ($P \sim 0$, relative to a model with no true asymmetry). Among duplicates that likely arose within the placental mammals ($d_S < 0.7$), 52.6% of duplicate pairs are AEDs, com-

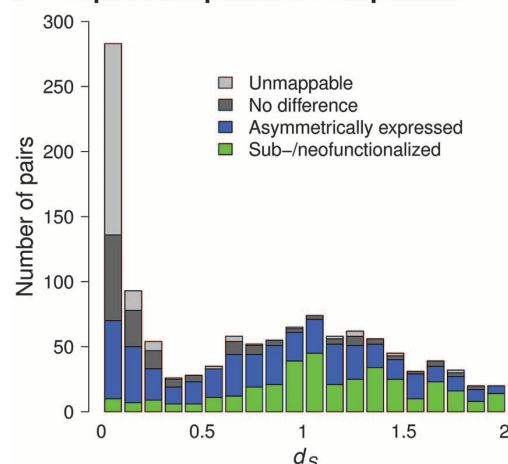
pared with just 15.2% that are potentially subfunctionalized. As might be expected, the minor genes at AEDs show evidence of reduced selective constraint relative to their duplicate partners, both within the human population (fig. S23) and between species (fig. S21). Furthermore, in gene pairs with asymmetric expression, the minor genes tend to be associated with significantly fewer diseases ($P = 8 \times 10^{-7}$; Wald test) (Fig. 2E). Nonetheless, despite their reduced importance, minor genes are not dispensable: 97% of minor genes have $d_N/d_S < 1$, a hallmark of protein-coding constraint (fig. S21).

Together, these results show that subfunctionalization of expression evolves slowly. However, we noticed much higher rates of sub- or neofunctionalization for duplicates located on different chromosomes, compared with duplicates in

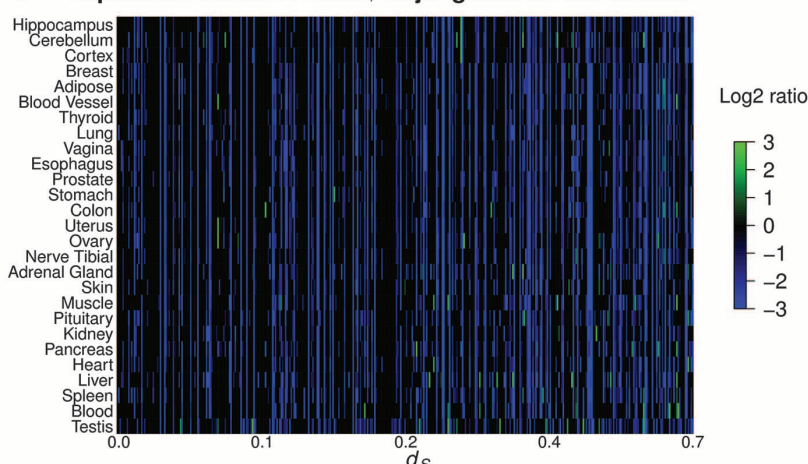
tandem ($P = 5 \times 10^{-23}$; Fisher's exact test) (fig. S24). We thus wanted to understand whether separation of duplicates enables subfunctionalization or whether the higher rate simply reflects the greater age of separated duplicates. Most duplicates arise as segmental duplications (18) and are close together in the genome: 87% of young gene pairs ($d_S < 0.1$) are on the same chromosome (Fig. 3A). Duplicates may subsequently become separated as the result of chromosomal rearrangements; however, this is a slow process. It is not until $d_S = 0.6$ that half of gene duplicates are found on different chromosomes.

Even controlling for duplicate age, however, there is a strong signal that genomic separation is a key factor enabling expression divergence (Fig. 3B). Separated duplicates have roughly 50% lower correlation of expression across tissues:

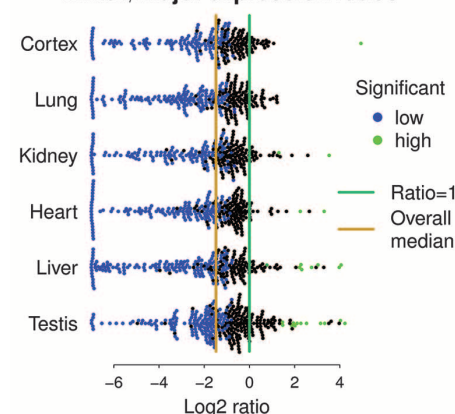
A Expression patterns of duplicates



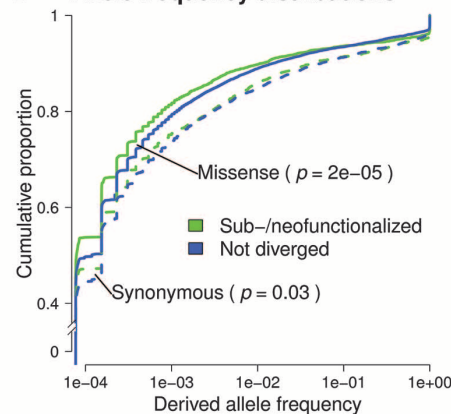
B Expression ratios of minor/major genes across tissues



C Minor/major expression ratios



D Allele frequency distributions



E Associated diseases

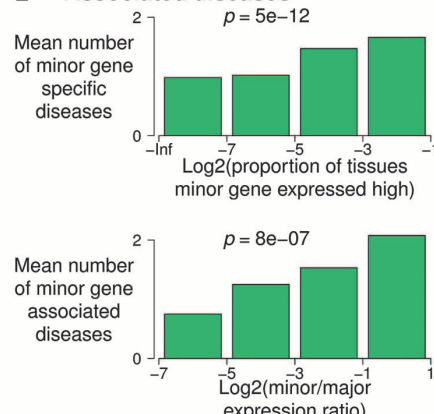


Fig. 2. Properties of subfunctionalized genes. (A) Classification of gene pairs by expression patterns. For context, note that duplicates arising at the human-mouse split would have $d_S \sim 0.45$. (B) Heat map of expression ratios for duplicate pairs. For each duplicate pair (plotted in columns), the ratios show the tissue-specific expression level of the minor gene relative to its duplicate. Green indicates evidence for subfunctionalization; consistently blue columns indicate AEDs. Black indicates tissue ratios not significantly different from 1 ($P > 0.001$). (C) Distributions of expression ratios in different tissues (minor genes/major genes). Ratios significantly >1 marked in green.

(D) Frequency spectra of human polymorphism data (15) for synonymous and nonsynonymous variants in subfunctionalized duplicates (green) and duplicates without significant expression differences (black). The plots show cumulative derived allele frequencies at segregating sites. The lines that climb more steeply (subfunctionalized genes) have a higher fraction of rare variants, indicating stronger selective constraint. (E) Disease burden of minor genes is highly correlated with degree of subfunctionalization (top) and overall expression relative to major genes (bottom). Data in (B), (C), and (D) are for $d_S < 0.7$.

$P = 3 \times 10^{-30}$, controlling for age by d_S in a multiple regression model (table S4 and fig. S26); $P = 6 \times 10^{-18}$, controlling for age by phylogenetic distribution (fig. S7). Further, we see the same effect in a paired test of duplicates that are separated in human but not mouse, or vice versa (Fig. 3C). Notably, duplicate age itself is a much weaker predictor ($P = 2 \times 10^{-6}$ for d_S) than is genomic separation ($P = 3 \times 10^{-30}$) (table S4). [In contrast to correlation across tissues, the asymmetry of mean expression is uncorrelated with whether the duplicates are on the same chromosome or not ($P = 0.9$, controlling for d_S ; Wald test).]

These results echo previous observations that, in general, genes that are close in the genome tend to be coregulated, with correlated expression (19) and often shared expression quantitative trait loci (eQTLs) (20). This effect is yet stronger for duplicates: Gene expression is more correlated for tandem duplicates than for singleton neighbors ($P = 10^{-19}$; t test) (Fig. 3D), and duplicates share eQTLs at higher rates than matched singletons ($P = 6 \times 10^{-4}$ and 5×10^{-4} in two data sets; Fisher's exact test) (13, 20, 27). Further, duplicates show higher connectivity by whole-genome chromosome conformation capture (Hi-C) (22), including higher numbers of promoter-promoter links than neighboring singletons (Fig. 3E) (mean ef-

fect size = 1.7-fold, $P = 3 \times 10^{-6}$; Wald test) (13). Promoter-promoter links may reflect a tendency of coregulated genes to be transcribed simultaneously within transcription factories (23). In contrast, duplicates on different chromosomes show no evidence of Hi-C linkage. In summary, we hypothesize that tandem duplicates tend to be highly coregulated and that genomic separation is a key factor enabling independent evolution.

Thus far, our results argue that expression subfunctionalization evolves slowly, in large part because tandem duplicates tend to be coregulated. An alternative explanation for the initial survival of duplicates is that they are both necessary to produce the required expression dosage (6). However, in contrast to whole-genome duplications, the small-scale duplications that are typical in mammals would initially disrupt dosage of the duplicated genes relative to all other genes. Thus, if dosage sharing is important in mammals, this would suggest that after tandem duplication, the duplicates should rapidly evolve reduced expression. Subsequent loss of either gene would then cause a deficit of expression and be deleterious.

To evaluate this, we analyzed the expression of human duplicates that arose since the human-macaque split, using RNA-seq data from six tissues in human and macaque (Fig. 4A) (13, 24). Indeed, there is a very clear signal that both hu-

man copies tend to evolve reduced expression, such that the median summed expression of the human duplicates is close to the expression of the singleton orthologs in macaque (median expression ratio 1.11; this is significantly less than the 2:1 expression ratio expected on the basis of copy number, $P = 3 \times 10^{-7}$; t test). Interestingly, polymorphic duplicates also show partial down-regulation, whereas the youngest fixed duplicates are about as down-regulated as older pairs, suggesting that reduced expression occurs rapidly (fig. S19). In contrast, we find no evidence for coding adaptation in these relatively young duplicate pairs (fig. S16). Thus, dosage sharing may be a frequent first step in the preservation of tandem duplicates. However, although dosage sharing evolves quickly, it is notable that duplicate genes remain less conserved than singleton genes over long evolutionary time scales ($d_S \leq 0.7$, or roughly the age of placental mammals) (Fig. 4B and fig. S22).

We propose that down-regulation is a key first step enabling the initial survival of duplicates, followed by dosage sharing, as suggested for WGDs (Fig. 4C) (6). In this view, the early survival of young duplicates is a race between down-regulation to achieve dosage balance versus mutational degradation of one copy. If dosage balance is achieved, then the relative expression levels of

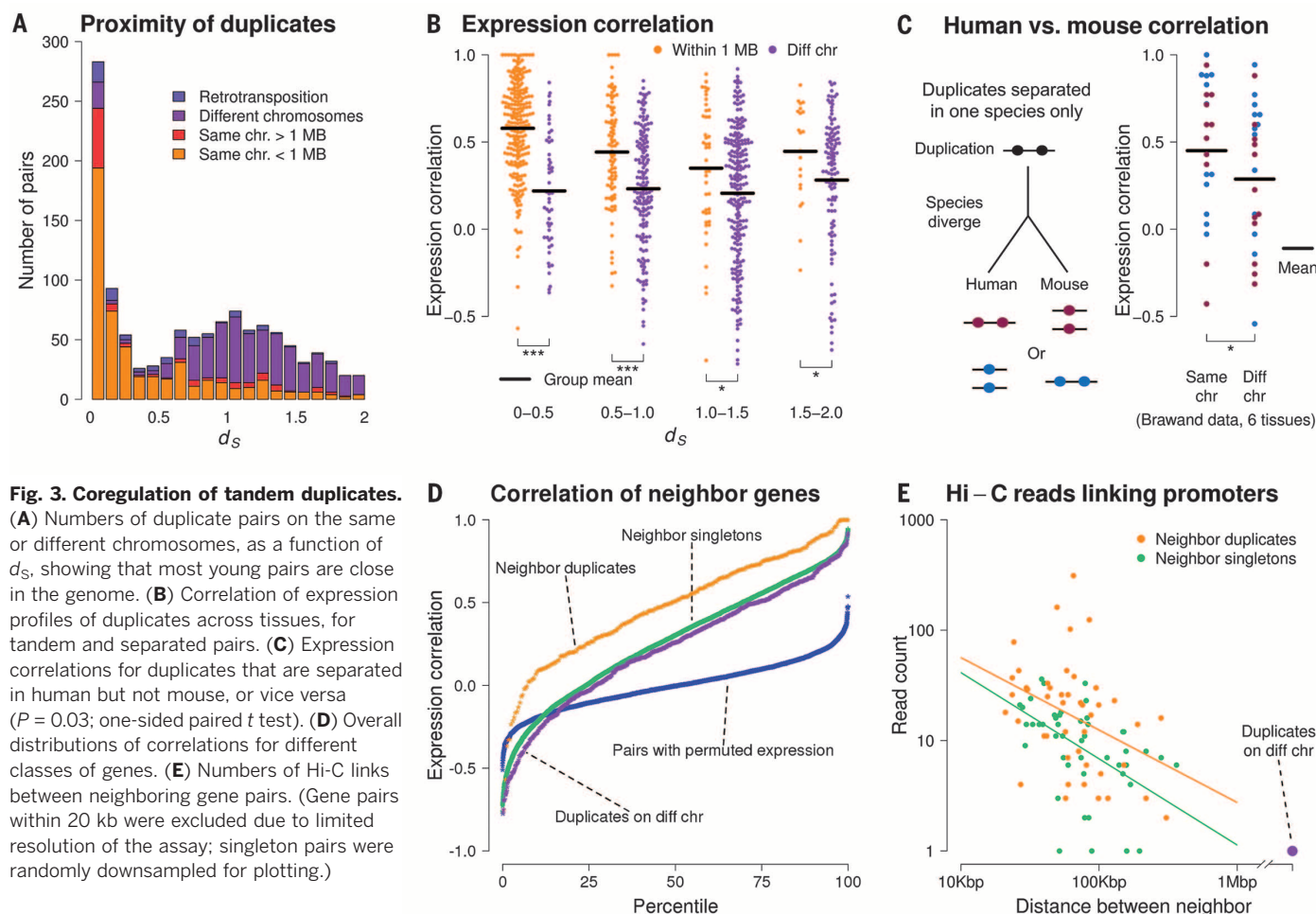


Fig. 3. Coregulation of tandem duplicates. (A) Numbers of duplicate pairs on the same or different chromosomes, as a function of d_S , showing that most young pairs are close in the genome. (B) Correlation of expression profiles of duplicates across tissues, for tandem and separated pairs. (C) Expression correlations for duplicates that are separated in human but not mouse, or vice versa ($P = 0.03$; one-sided paired t test). (D) Overall distributions of correlations for different classes of genes. (E) Numbers of Hi-C links between neighboring gene pairs. (Gene pairs within 20 kb were excluded due to limited resolution of the assay; singleton pairs were randomly downsampled for plotting.)

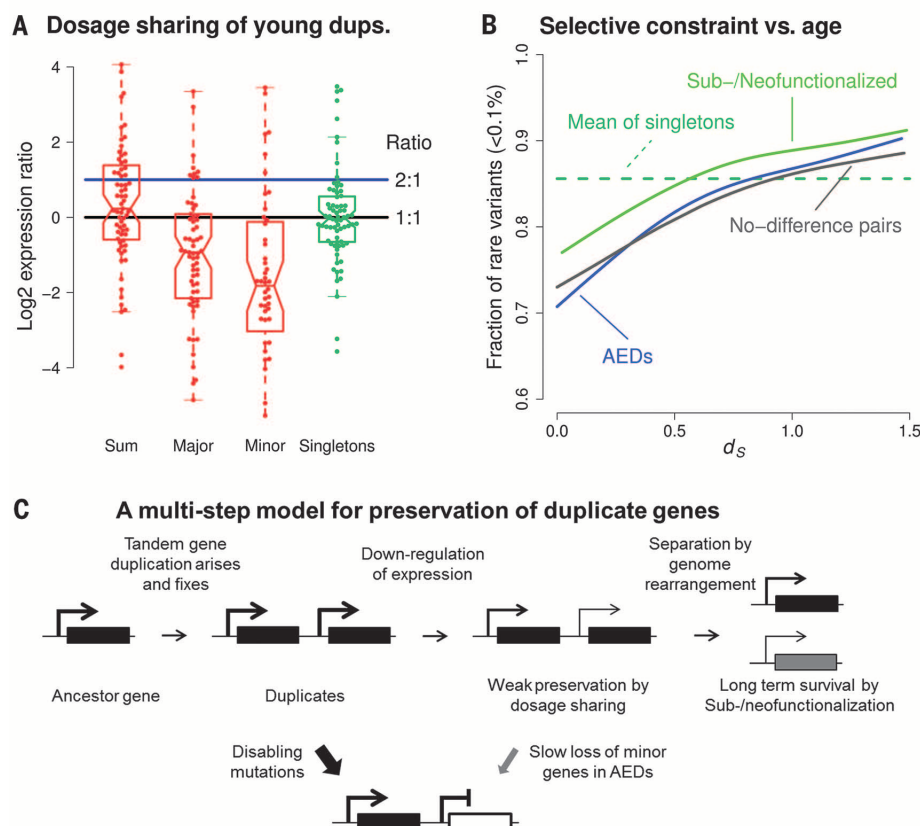


Fig. 4. Long-term survival of duplicate genes. (A) Expression levels of young duplicates compared to their macaque orthologs in six tissues (24), for human duplicates that are single-copy genes in macaque. Sum shows the summed expression of both duplicates, relative to expression of the macaque orthologs in the same tissues. “Major” and “Minor” show corresponding ratios for major and minor genes separately, classified using GTEx data. The green data show a random set of singleton orthologs. Each tissue-gene expression ratio is plotted separately. (B) The strength of purifying selection in humans increases with duplicate age. The fraction of rare missense variants in a large human data set (15) is used as a proxy for the strength of purifying selection. (C) Conceptual model of duplicate gene evolution. Other transitions not explicitly shown would occur at lower but nonzero rates.

the two genes evolve slowly as a random walk due to constraint on their combined expression (7, 25). Both copies tend to evolve under reduced constraint, especially for minor genes of AEDs. Genomic separation frees expression of the dup-

licates to evolve independently and may also encourage protein adaptation, potentially leading to true functional differentiation and long-term survival. In summary, we find that subfunctionalization of expression evolves slowly

in mammals due to coregulation of tandem duplicates and that rapid evolution of dosage sharing may be the most frequent first step to duplicate preservation.

REFERENCES AND NOTES

- G. C. Conant, K. H. Wolfe, *Nat. Rev. Genet.* **9**, 938–950 (2008).
- M. Lynch, J. S. Conery, *J. Struct. Funct. Genomics* **3**, 35–44 (2003).
- H. Innan, F. Kondrashov, *Nat. Rev. Genet.* **11**, 97–108 (2010).
- A. Stoltzfus, *J. Mol. Evol.* **49**, 169–181 (1999).
- W. Qian, B. Y. Liao, A. Y. Chang, J. Zhang, *Trends Genet.* **26**, 425–430 (2010).
- G. C. Conant, J. A. Birchler, J. C. Pires, *Curr. Opin. Plant Biol.* **19**, 91–98 (2014).
- J. F. Gout, M. Lynch, *Mol. Biol. Evol.* **32**, 2141–2148 (2015).
- A. Force *et al.*, *Genetics* **151**, 1531–1545 (1999).
- C. R. Baker, V. Hanson-Smith, A. D. Johnson, *Science* **342**, 104–108 (2013).
- I. Wapinski, A. Pfeffer, N. Friedman, A. Regev, *Nature* **449**, 54–61 (2007).
- GTEx Consortium, *Science* **348**, 648–660 (2015).
- T. Babak *et al.*, *Nat. Genet.* **47**, 544–549 (2015).
- See supplementary materials and methods on Science Online.
- B. van de Geijn, G. McVicker, Y. Gilad, J. K. Pritchard, *Nat. Methods* **12**, 1061–1063 (2015).
- W. Fu *et al.*, *Nature* **493**, 216–220 (2013).
- K. Peng *et al.*, *Nucleic Acids Res.* **41**, D553–D560 (2013).
- E. W. Ganko, B. C. Meyers, T. J. Vision, *Mol. Biol. Evol.* **24**, 2298–2309 (2007).
- J. A. Bailey *et al.*, *Science* **297**, 1003–1007 (2002).
- A. T. Ghanbarian, L. D. Hurst, *Mol. Biol. Evol.* **32**, 1748–1766 (2015).
- A. Battle *et al.*, *Genome Res.* **24**, 14–24 (2014).
- T. Lappalainen *et al.*, *Nature* **501**, 506–511 (2013).
- S. S. Rao *et al.*, *Cell* **159**, 1665–1680 (2014).
- A. Feuerborn, P. R. Cook, *Trends Genet.* **31**, 483–490 (2015).
- D. Brawand *et al.*, *Nature* **478**, 343–348 (2011).
- K. Y. Popadin *et al.*, *Am. J. Hum. Genet.* **95**, 660–674 (2014).

ACKNOWLEDGMENTS

This work was funded by NIH grants ES025009 and MH01825 and by the Howard Hughes Medical Institute. We thank H. Fraser for prepublication access to data (12) and H. Fraser, A. Fu, A. Harpak, Y. I. Li, D. Petrov, P. C. Phillips, M. Przeworski, A. Stoltzfus, and the anonymous reviewers for comments or discussion. J.K.P. is on advisory boards for 23andMe and DNAnexus, with stock options in both.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6288/1009/suppl/DC1
Materials and Methods
Supplemental Text
Figs. S1 to S27
Tables S1 to S5
Supplementary Files S1 to S3
References (26–79)

9 November 2015; accepted 2 April 2016
10.1126/science.aad8411

Coregulation of tandem duplicate genes slows evolution of subfunctionalization in mammals

Xun Lan and Jonathan K. Pritchard

Science **352** (6288), 1009-1013.
DOI: 10.1126/science.aad8411

Evolutionary maintenance of gene duplications

Understanding genetic redundancy—the maintenance of multiple copies of a gene after duplication—and its relevance to genetic evolution have long been debated. Lan and Pritchard examined gene duplicates within human and other mammalian genomes. The expression of genes appears to be controlled by dosage balance and tight coregulation of tandem duplicates. They found little evidence for gene copies evincing significantly different expression patterns. However, such changes can evolve later, after gene copies become physically separated within the genome and thus are no longer jointly regulated.

Science, this issue p. 1009

ARTICLE TOOLS

<http://science.sciencemag.org/content/352/6288/1009>

SUPPLEMENTARY MATERIALS

<http://science.sciencemag.org/content/suppl/2016/05/18/352.6288.1009.DC1>

REFERENCES

This article cites 75 articles, 16 of which you can access for free
<http://science.sciencemag.org/content/352/6288/1009#BIBL>

PERMISSIONS

<http://www.sciencemag.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of Service](#)