



Mosaic Copy Number Variation in Human Neurons

Michael J. McConnell *et al.*

Science **342**, 632 (2013);

DOI: 10.1126/science.1243472

This copy is for your personal, non-commercial use only.

If you wish to distribute this article to others, you can order high-quality copies for your colleagues, clients, or customers by [clicking here](#).

Permission to republish or repurpose articles or portions of articles can be obtained by following the guidelines [here](#).

The following resources related to this article are available online at www.sciencemag.org (this information is current as of November 4, 2013):

Updated information and services, including high-resolution figures, can be found in the online version of this article at:

<http://www.sciencemag.org/content/342/6158/632.full.html>

Supporting Online Material can be found at:

<http://www.sciencemag.org/content/suppl/2013/10/30/342.6158.632.DC1.html>

A list of selected additional articles on the Science Web sites **related to this article** can be found at:

<http://www.sciencemag.org/content/342/6158/632.full.html#related>

This article **cites 46 articles**, 16 of which can be accessed free:

<http://www.sciencemag.org/content/342/6158/632.full.html#ref-list-1>

This article has been **cited by** 1 articles hosted by HighWire Press; see:

<http://www.sciencemag.org/content/342/6158/632.full.html#related-urls>

A balance between inductive and repressive signals determines optimal flowering time for successful reproduction in seasonal environments (28, 29). Our data demonstrate the importance of the repressors SVP, FLM- β , and FLC, which affect flowering at different but partially overlapping temperature ranges (fig. S19). However, we cannot exclude the possibility that the importance of SVP and FLM- β may not extend to *Arabidopsis* winter annual accessions. At lower temperatures, increased SVP protein interacts with FLM- β to repress flowering by direct binding to downstream targets. At higher temperatures, SVP protein levels decline, allowing expression of the downstream target genes and enabling flowering at that temperature. We propose that repressive activity of SVP requires FLM- β in the same complex and that SVP protein stability modulates the abundance of the SVP-FLM- β repressor complex to regulate ambient temperature-responsive flowering. SVP-like proteins have conserved functions across plant species (30, 31); therefore, control of SVP-FLM- β repressor complex abundance could be a general strategy for plants to adjust the balance of inductive and repressive signals under fluctuating temperature conditions.

References and Notes

1. A. Srikanth, M. Schmid, *Cell. Mol. Life Sci.* **68**, 2013–2037 (2011).
2. S. Penfield, *New Phytol.* **179**, 615–628 (2008).

3. J. H. Lee, J. S. Lee, J. H. Ahn, *J. Plant Biol.* **51**, 321–326 (2008).
4. S. Sung, R. M. Amasino, *Annu. Rev. Plant Biol.* **56**, 491–508 (2005).
5. N. A. Delk, K. A. Johnson, N. I. Chowdhury, J. Braam, *Plant Physiol.* **139**, 240–253 (2005).
6. A. H. Fitter, R. S. Fitter, *Science* **296**, 1689–1691 (2002).
7. B. I. Cook, E. M. Wolkovich, C. Parmesan, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 9000–9005 (2012).
8. P. Q. Craufurd, T. R. Wheeler, *J. Exp. Bot.* **60**, 2529–2539 (2009).
9. A. B. Nicotra *et al.*, *Trends Plant Sci.* **15**, 684–692 (2010).
10. J. T. Anderson, D. W. Inouye, A. M. McKinney, R. I. Colautti, T. Mitchell-Olds, *Proc. Biol. Sci.* **279**, 3843–3852 (2012).
11. U. Johanson *et al.*, *Science* **290**, 344–347 (2000).
12. S. D. Michaels, R. M. Amasino, *Plant Cell* **11**, 949–956 (1999).
13. C. C. Sheldon *et al.*, *Plant Cell* **11**, 445–458 (1999).
14. J. Lempe *et al.*, *PLOS Genet.* **1**, e6 (2005).
15. S. Balasubramanian, S. Sureshkumar, J. Lempe, D. Weigel, *PLOS Genet.* **2**, e106 (2006).
16. J. D. Werner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **102**, 2460–2465 (2005).
17. K. C. Scortecchi, S. D. Michaels, R. M. Amasino, *Plant J.* **26**, 229–236 (2001).
18. S. Balasubramanian, D. Weigel, *Plant Signal. Behav.* **1**, 227–228 (2006).
19. S. de Folter *et al.*, *Plant Cell* **17**, 1424–1433 (2005).
20. X. Gu *et al.*, *Nat. Commun.* **4**, 1947 (2013).
21. J. H. Lee *et al.*, *Genes Dev.* **21**, 397–402 (2007).
22. A. M. Wilczek *et al.*, *Science* **323**, 930–934 (2009).
23. D. M. Bond, E. S. Dennis, E. J. Finnegan, *Plant Cell Environ.* **34**, 1737–1748 (2011).
24. J. J. Kim *et al.*, *Plant Physiol.* **159**, 461–478 (2012).
25. H. J. Cho *et al.*, *FEBS Lett.* **586**, 2332–2337 (2012).

26. J. W. Wang, B. Czech, D. Weigel, *Cell* **138**, 738–749 (2009).
27. C. A. Helliwell, C. C. Wood, M. Robertson, W. James Peacock, E. S. Dennis, *Plant J.* **46**, 183–192 (2006).
28. I. Bäurle, C. Dean, *Cell* **125**, 655–664 (2006).
29. Y. Kobayashi, D. Weigel, *Genes Dev.* **21**, 2371–2384 (2007).
30. J. H. Lee, S. H. Park, J. H. Ahn, *Plant Sci.* **185–186**, 97–104 (2012).
31. B. Trevaskis *et al.*, *Plant Physiol.* **143**, 225–235 (2007).

Acknowledgments: We thank R. Amasino, C. Dean, I. Lee, I. Hwang, C. M. Park, J. Puterill, D. Weigel, and P. Wigge for kindly providing research materials and D. H. Kim, B. R. Kim, H. J. Cho, S. H. Kim, S. M. Hong, H. S. Jung, and S. J. Yoo for their technical assistance. J.H.L. was supported by a Korea University Grant. D.P. and M.S. were supported by the ERA-NET Plant Genomics BLOOMNET project (SCHM1560/7-1). This work was supported by a National Research Foundation of Korea grant funded by the Korea government (Ministry of Science, ICT, and Future Planning) (2008-0061988) to J.H.A. The authors declare no conflicts of interest. GenBank accession numbers are KF496888 to KF496899 (*FLM* sequences of *Arabidopsis* WT accessions). Materials described here are available from J.H.A. The genomic FLM-GFP line will be distributed for noncommercial research under a standard materials transfer agreement by the Max Planck Institute for Developmental Biology.

Supplementary Materials

www.sciencemag.org/content/342/6158/628/suppl/DC1
Materials and Methods
Figs. S1 to S19
Tables S1 to S5
References (32–56)

28 May 2013; accepted 3 September 2013

Published online 12 September 2013;

10.1126/science.1241097

Mosaic Copy Number Variation in Human Neurons

Michael J. McConnell,^{1,2,7,8,9} Michael R. Lindberg,⁷ Kristen J. Brennan,^{1*} Julia C. Piper,^{1,2,†} Thierry Voet,^{3,4} Chris Cowing-Zitron,¹ Svetlana Shumilina,⁷ Roger S. Lasken,^{5,6} Joris R. Vermeesch,³ Ira M. Hall,^{7,9,‡} Fred H. Gage^{1,‡}

We used single-cell genomic approaches to map DNA copy number variation (CNV) in neurons obtained from human induced pluripotent stem cell (hiPSC) lines and postmortem human brains. We identified aneuploid neurons, as well as numerous subchromosomal CNVs in euploid neurons. Neurotypic hiPSC-derived neurons had larger CNVs than fibroblasts, and several large deletions were found in hiPSC-derived neurons but not in matched neural progenitor cells. Single-cell sequencing of endogenous human frontal cortex neurons revealed that 13 to 41% of neurons have at least one megabase-scale de novo CNV, that deletions are twice as common as duplications, and that a subset of neurons have highly aberrant genomes marked by multiple alterations. Our results show that mosaic CNV is abundant in human neurons.

Neuronal genomes exhibit elevated levels of aneuploidy (1–3) and retrotransposition (4–6) relative to other cell types; this finding has fueled speculation that somatic genome variation may contribute to functional diversity in the human brain (7–10). The prevalence of copy number variations (CNVs) has been difficult to assess, given the limited ability of conventional genome-wide methods to detect CNVs that are rare within a population of cells, as most somatic mutations are expected to be. Recently, two methods have been developed to map large-scale CNVs in single cells: microarray analysis of multiple displacement amplification (MDA) products (11) and single-cell

sequencing (12). Here, we applied both of these approaches to single human neurons.

We examined human neurons from two neurotypic sources (fig. S1A): (i) human induced pluripotent stem cells [i.e., hiPSC-derived neurons (fig. S2)] and (ii) human postmortem frontal cortex (FCTX) neurons (fig. S3). We employed fluorescence-activated cell sorting (FACS) to obtain neurons from neuronogenic hiPSC cultures using synapsin:green fluorescent protein (GFP) expression and from postmortem tissue using NeuN immunostaining (13). After multiple displacement amplification (MDA) (14), we hybridized single hiPSC-derived neuronal genomes to Affymetrix 250K single-

nucleotide polymorphism (SNP) arrays [as in (11)]. We subjected single neurons from postmortem tissue to Illumina DNA sequencing using a custom version of the single-cell sequencing protocol developed by Navin *et al.* (12), which combines the GenomePlex whole-genome amplification method with Nextera-based library preparation (15). We developed stringent quality-control measures to ensure that only the highest-quality amplification reactions and data sets were included in downstream analyses (see methods).

To detect CNVs, we first aggregated raw copy number measurements over very large genomic intervals. We then selected interval sizes that were 1 to 2 orders of magnitude larger than the local

¹Laboratory of Genetics, Salk Institute for Biological Studies, La Jolla, CA 92037, USA. ²Crick-Jacobs Center for Theoretical and Computational Biology, Salk Institute for Biological Studies, La Jolla, CA 92037, USA. ³Center for Human Genetics, KU Leuven, 3000 Leuven, Belgium. ⁴Single-Cell Genomics Centre, Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. ⁵J. Craig Venter Institute, San Diego, CA 92121, USA. ⁶Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA 92093, USA. ⁷Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA 22908, USA. ⁸Center for Brain Immunology and Glia, University of Virginia, Charlottesville, VA 22908, USA. ⁹Center for Public Health Genomics, University of Virginia, Charlottesville, VA 22908, USA.

*Present address: Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA.

†Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA.

‡Corresponding author. E-mail: irahall@virginia.edu (I.M.H.) and gage@salk.edu (F.H.G.)

amplification biases reported for single-cell DNA amplification (16, 17). For SNP array data, we calculated the median copy number in 100-probe bins, which corresponds to a mean genomic interval of 666 kb; for sequencing data, we measured read-depth in bins composed of 500 kb of uniquely mappable sequence (mean size of 687 kb). CNVs were identified using circular binary segmentation (18) combined with strict filtering based on the number of consecutive bins identified by segmentation and the amplitude of CNV predictions relative to the noise (median absolute deviation) of each data set. These methods and filtering criteria resulted in a mean CNV size detection limit of 6.7 Mb for SNP array data and 3.4 Mb for sequencing data. A subset ($n = 7$) of the MDA-amplified hiPSC-derived neurons, analyzed by both SNP array and sequencing, showed high concordance (fig. S1B and fig. S4). Subchromosomal deletions (Fig. 1, A and C) and duplications (Fig. 1, B and D) were identified in both groups of neurons.

We examined neurons from three hiPSC lines, referred to as C, D, and E, that were generated from three different individuals as neurotypic controls for a hiPSC-based disease model (19). Analysis of bulk DNA from C and D line donor fibroblasts or hiPSC-derived neural progenitor

cells (NPCs) revealed no clonal genomic aberrations. Of 40 single neurons analyzed [for C, ($n = 21$); D, ($n = 6$); E, ($n = 13$)], 27 had copy number profiles consistent with bulk DNA, but 13 had unique genomes. In total, we identified seven whole-chromosome gains, four whole-chromosome losses, and 12 subchromosomal CNVs (range: 7 to 156 Mb) in 13 hiPSC-derived neurons (Fig. 2A, fig. S5, and table S1). Each CNV was identified in merely one neuron, which suggests that the CNVs are not early clonal events but rather are unique to single cells or distinct lineages.

The CNVs detected in C and D line hiPSC-derived neurons were distinct from those seen in either C or D line fibroblasts or NPCs (Fig. 2). Of 29 fibroblasts, 6 had single CNVs (range: 5.2 to 27.7 Mb) and one was aneuploid ($-22, -X$) (Fig. 2A). Among 19 hiPSC-derived NPCs, only 6 duplications were observed (Fig. 2A). Technical replicates of five fibroblasts and three hiPSC-derived neurons showed high concordance, and principal component analysis also showed that replicates from each individual neuron clustered distinctly from both the fibroblasts and the other two neurons (fig. S2E). Comparison of CNVs in the three cell types (Fig. 2B) showed that neurons have significantly larger CNVs than fibroblasts

(Kolmogorov-Smirnov test, $P < 0.001$). In addition, we found deletions only in hiPSC-derived neurons and not in hiPSC-derived NPCs.

We performed two additional experiments to confirm that low-level aneuploidy and CNVs occur in single fibroblasts. First, we obtained single-cell clones by limiting dilution. Each single fibroblast was expanded to ~20 sister cells over 7 days; then, we obtained individual sister fibroblasts from three different clonal expansions. In one of these clones, chromosome missegregation was observed as a gain of Chr2 in one cell and a loss of Chr2 in a sister cell (Fig. 3A). Non-clonal CNVs were also detected, so we performed a second experiment using fluorescence in situ hybridization (FISH) for a common hiPSC CNV on Chr20 (20) and for ChrX. Consistent with genomic analysis of bulk DNA, 20 metaphase spreads from this population karyotyped as euploid, but 13 out of 200 were aneuploid for ChrX (Fig. 3B) and 26 out of 200 nuclei had a Chr20 CNV (Fig. 3C). These data show that two distinct approaches (SNP array and FISH) detect large nonclonal CNVs that arise in single human cells in culture.

We next sought to determine whether mosaic CNVs were also present in FCTX neurons from postmortem human brains. For these experiments

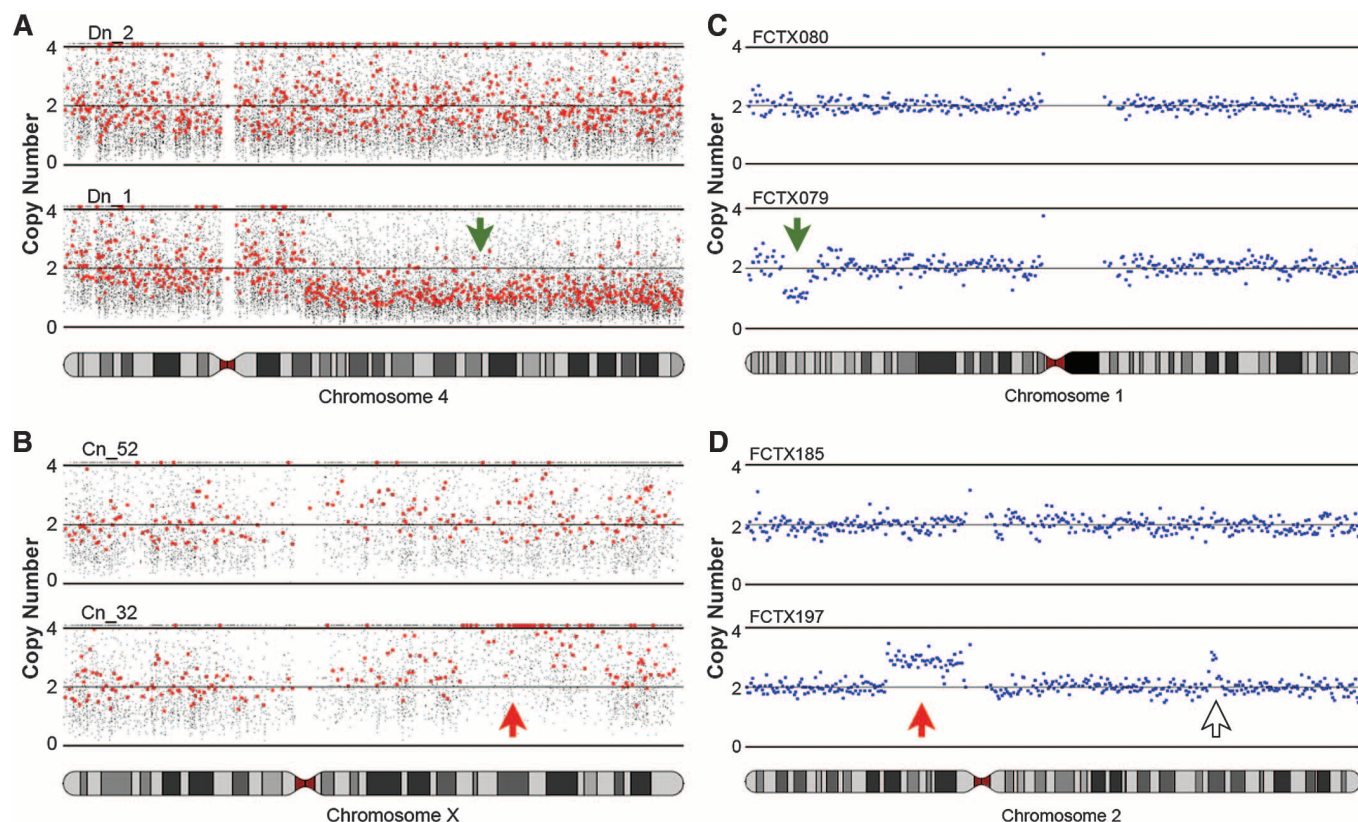


Fig. 1. Mosaic copy number variation (CNV) is detected in human neurons. (A and B) Subchromosomal deletions (green down arrow) and duplications (red up arrow) are observed in hiPSC-derived neurons. (A) Neuron Dn_1 has a deletion on chromosome (chr) 4q (bottom); neuron Dn_2 has no CNV on Chr4 (top). Small gray dots show the predicted copy number at individual SNPs; red dots show every 30th SNP. (B) Neuron Cn_32 has a duplication on ChrXq (bottom); neuron Cn_2 does not (top). (C and D) Single-cell sequencing reveals

subchromosomal deletions (green down arrow) and duplications (red up arrow) in FCTX neurons. (C) FCTX079 has a deletion on Chr1p (bottom); FCTX080 does not (top). Blue dots show raw copy number predictions obtained by read-depth analysis (mean window size ~687 kb; see methods). (D) Neuron FCTX197 has a duplication on Chr2p (bottom), whereas FCTX185 does not (top). Another likely duplication on Chr2q in FCTX197 (open arrow) comprised only four consecutive bins and therefore failed our five-bin confidence threshold.

we used the single-cell sequencing method (12), which offers superior sensitivity to microarray approaches because of the digital nature of DNA sequence data (12, 21). After benchmarking the sequencing approach with trisomic male fibroblasts in which we identified 100% trisomy 21 and monosomy X (Fig. 3D, fig. S6, and table S2), we sequenced 110 FCTX neurons from three different individuals [a 24-year-old female (NICHD Brain Bank ID no. 5125; $n = 19$), a 26-year-old male (ID no. 1583; $n = 41$), and a 20-year-old female (ID no. 1846; $n = 50$)] and used strict filtering criteria to identify high-confidence CNVs (see methods) composed of five or more consecutive bins. We identified 100% monosomy X and Y in the 41 male neurons (Fig. 4A, fig. S7, and table S3) as expected, and simulation experiments indicate that our methods detected CNVs at high sensitivity and specificity, with a predicted mean false-negative rate of 17% and a predicted mean false-discovery rate of 0.6% (fig. S8; see methods).

We identified one or more somatic CNVs in 45 of the 110 (41%) FCTX neurons analyzed (Fig. 4, fig. S7, and table S2). The vast majority of somatic CNVs were subchromosomal alterations ranging in size from 2.9 to 75 Mb, although we also identified one putative chromosome gain and two losses where CNV calls affected >50% of the chromosome (e.g., FCTX155) (Fig. 4A). Subchromosomal CNVs were distributed throughout the genome, and in only one case did two independent CNVs share the same breakpoints (a 3-Mb subtelomeric deletion on Chr16 in FCTX198 and FCTX224 (fig. S7 and table S2). However, a number of loci were affected by multiple

“small” CNVs less than 20 Mb in size ($N = 133$), and small CNVs were preferentially found at telomeres (Fig. 4B), with 23.3% extending to the chromosome end (2067-fold enrichment by Monte-Carlo simulation, see methods). Small CNVs are not enriched with features known to affect genome stability, such as transposons, segmental duplications, or fragile sites; neither are they enriched with germline CNVs or known genes (fig. S9). Subchromosomal deletions were prevalent in each of the three individuals and were twice as common as duplications, on average, which might be explained by a bias toward DNA loss in nondividing postmitotic neurons; however, the third individual (no. 1846) was unique in also showing abundant duplications (fig. S3, D to G). These results demonstrate that somatic CNVs are a common feature of neuronal genomes and suggest that the relative abundance of different CNV classes may vary among individuals.

The overall high mutational load that we report in neurons is predominantly due to a small number of cells with highly aberrant genomes. Whereas the majority of FCTX neurons exhibited 0 (59%) or 1 or 2 CNVs (25%), 17 cells (15%) accounted for 108 of the 148 CNV calls (73%), and 7 cells accounted for nearly half (49%) of all calls (Fig. 4C). Aberrant cells are marked by multiple copy number switches on distinct chromosomes, with interdigitated altered and unaltered segments that adhere well to the expectation of integer-like copy number states measured by digital DNA sequencing technology. Similar, if less dramatic, examples of this phenomenon were apparent in hiPSC neurons, where several

cells harbored multiple alterations. For example, hiPSC-derived neuron Cn_32 had five events: loss of Chr13, three duplications, and one deletion (fig. S10). Similarly, two FCTX neurons had more than 10 events. One of these, FCTX155, was aneuploid for most of Chr2 and had 18 deletions and one duplication (Fig. 4A). We did not observe similarly aberrant copy number profiles among the 16 control fibroblasts analyzed by sequencing (fig. S6) or among the 42 fibroblasts or 19 NPCs analyzed by SNP array (fig. S5). Taken together, these results suggest that a subset of neurons is especially prone to large-scale genome alterations.

Single-cell genome analysis is inherently challenging, because all existing approaches require amplification of the genome before measurement; thus, validation is impossible because one cannot know the state of a single-cell's genome before it was amplified. However, several lines of evidence argue that the vast majority of events we report are true CNVs. First, we used methods that were previously validated on clonally related cell populations, including tumors (12) and eight-cell embryos (11). Second, we report megabase-scale CNVs that are orders of magnitude larger than the amplicons generated by whole-genome amplification. Indeed, previous studies have noted that amplification artifacts tended to be small (<10 kb) and distributed relatively uniformly across the genome (16, 17); therefore, simple amplification effects cannot readily explain the large-scale deviations in copy number that we observe. It is also difficult to explain how such effects could cause both gains and losses of DNA that

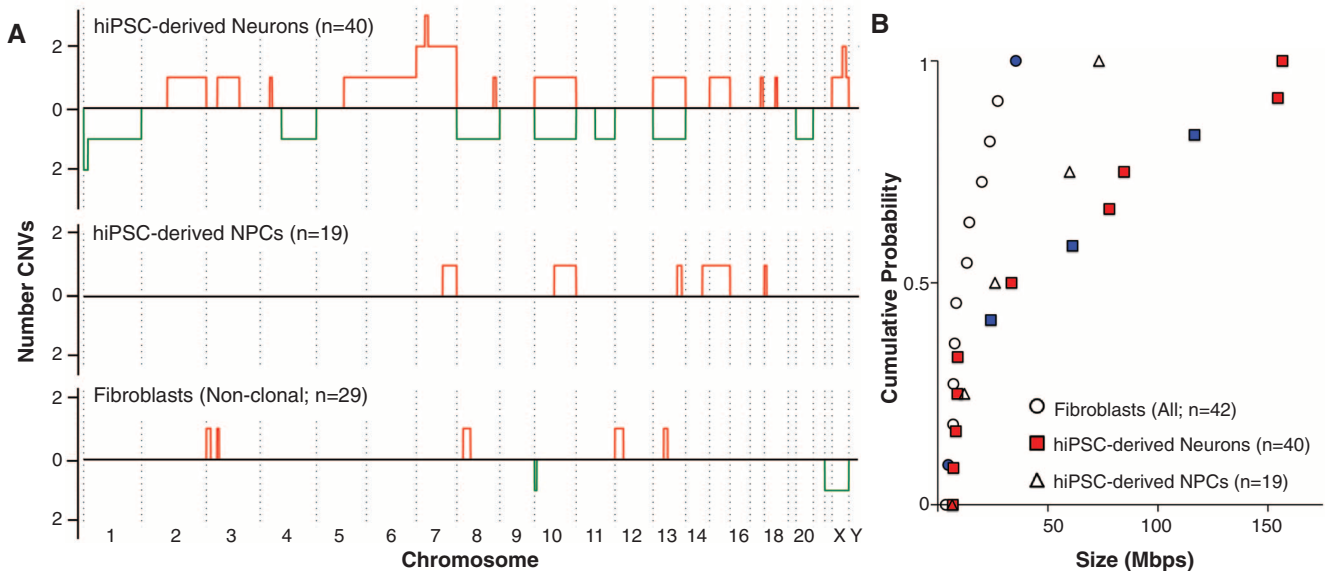


Fig. 2. Large CNVs are found in hiPSC-derived neurons. (A) Whole and subchromosomal duplications (red) and deletions (green) are summarized for 40 hiPSC-derived neurons (top). The y axis value represents the number of times each genomic interval was deleted (below in green) or duplicated (above in red). CNVs were detected in 9 out of 21 C neurons (Cn), 2 out of 6 D neurons (Dn), and 2 out of 13 E neurons (En). In donor hiPSC-derived NPC populations (middle), CNVs were detected in 1 out of 10 D NPCs (Dp) and 3 out of 9 C NPCs (Cp). In donor fibroblast populations (bottom), CNVs were detected in 7 out of

20 D fibroblasts (Df) and 0 out of 9 C fibroblasts (Cf). Note that chromosomes are not plotted to scale because data are summarized in 100-SNP bins. (B) Subchromosomal CNVs in fibroblasts were significantly smaller than in hiPSC-derived neurons (Kolmogorov-Smirnov test, $P < 0.001$). No deletions were observed in NPCs. Deletions are denoted with blue markers; all other markers indicate duplications. Aneuploidies are not included in this plot. For completeness, subchromosomal CNVs from clonal fibroblasts (Fig. 3) were included in this plot, bringing the total n to 42 fibroblasts.

produce integral copy number values by sequencing. Third, the postmortem interval is unlikely to contribute significantly to our results, because DNA degradation cannot generate duplications and because we observed large deletions in both FCTX and hiPSC-derived neurons. Fourth, Monte-Carlo simulation experiments showed that our CNV detection methods identify hemizygous gains and losses at high sensitivity and are not affected by random fluctuations in sequence coverage.

Fifth, we have employed strict quality-control measures to exclude data sets with uneven or noisy amplification or that (in the case of sequence data) do not exhibit expected integer-like copy number profiles (see methods). Finally, and perhaps most important, many of our CNV calls appear to be extremely high quality based on their size, amplitude, and integer-like properties (see Fig. 4A, fig. S6, and fig. S7), and a subset (30 to 56%) is robust to a series of increasingly strict CNV detection

parameters (fig. S11). At increased stringency, the overall number of CNVs diminishes but the core results do not change: CNVs are apparent in a significant fraction of neurons (13 to 24%), there is a predominance of deletions relative to duplications (fig. S11A), and we observe a subset of neurons with highly aberrant genomes marked by multiple copy number oscillations (fig. S11D). Therefore, although we cannot definitively exclude the possibility of as-yet-undescribed single-cell amplification

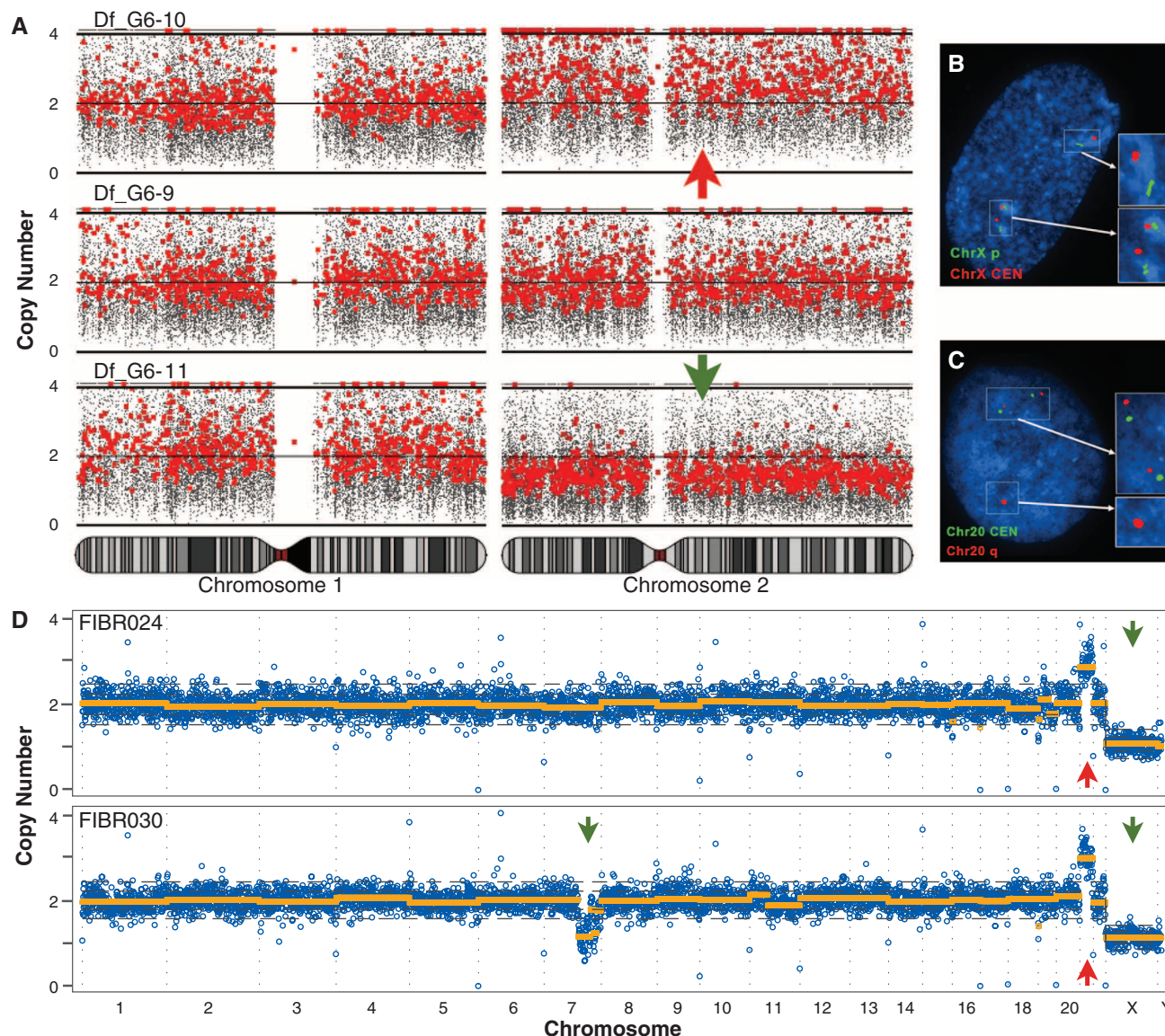


Fig. 3. Large CNVs are found in cultured fibroblasts. (A) Single fibroblasts obtained by limiting dilution were expanded to a population of ~20 clonal fibroblasts after 7 days in vitro (DIV). In one clonal population, a reciprocal chromosome missegregation event was detected. One fibroblast was trisomic for Chr2 (top) and a sister was monosomic for Chr2 (bottom). Chromosome 1 is shown along with the third euploid cell. (B and C) Two groups of Df (passages 7 and 8) were summarized in (Fig. 2A); a parallel culture of the p7 group was sent for karyotyping and FISH. Out of 20 metaphase chromosome spreads, 20 were euploid. (B) FISH was performed for a ChrX p arm telomere (green) and ChrX centromere (red). Out of 200 nuclei, 13 were aneuploid. (C) FISH was performed for the Chr20 centromere (green) and Chr20 CNV (red). Out of 200 nuclei, 26 had the CNV. (D) Single-cell sequencing of two male fibroblasts with

karyotypically defined trisomy 21. Genome-wide copy number profiles show that, in both cells, most of the genome is present at two copies, Chr21 is present at three copies, and ChrX is present at one copy. In addition, we identified a large deletion on Chr7q in FIBR030. DNA copy number (y axis) was calculated by read-depth analysis of variably sized genomic windows containing 500 kb of uniquely mappable sequence (blue), and CNVs were detected by circular binary segmentation (orange). Green (down) and red (up) arrows denote deletions and duplications, respectively, that were identified by segmentation and passed filtering criteria. Reported CNVs comprise five or more consecutive bins and exceed two median absolute deviations (MADs). Dotted gray lines show 1 and 2 MADs from the median copy number of each data set. See figs. S6 and S7 for plots of additional cells.

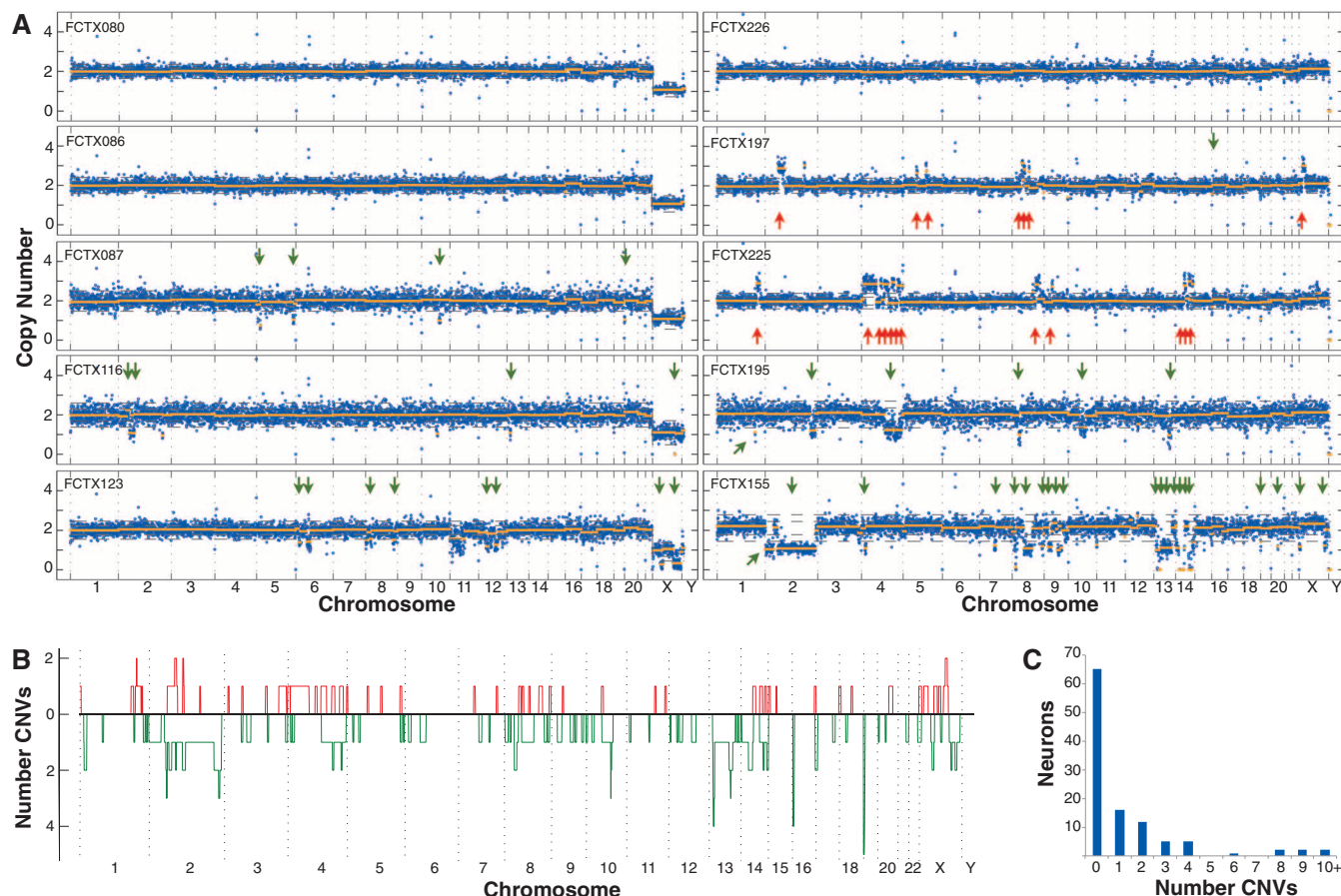


Fig. 4. Identification of CNVs in postmortem neurons using single-cell sequencing. (A) Genome-wide copy number profiles of five male (left) and five female (right) neurons from two individuals, no. 1583 and no. 1846, respectively. Data are plotted exactly as in Fig. 3D. Arrows denote deletions (green, down and at an angle in FCTX195 and 155) and duplications (red, up) that were identified by copy number segmentation and passed filtering criteria. Note that single-copy

"losses" of ChrX in cells from male individual no. 1583 are not indicated by arrows, but were identified in 100% of cells. See fig. S7 for plots of all cells. (B) Whole and subchromosomal duplications (red) and deletions (green) are summarized for the 110 FCTX neurons as in Fig. 2A. (C) The number of individual neurons (y axis) that exhibited a given number of CNVs (x axis). See fig. S11 for results at different CNV detection stringency thresholds.

artifacts, the above observations strongly argue that the central results and conclusions of our study are not attributable to technical factors.

Using three completely independent single-cell approaches (SNP array, sequencing, and FISH), we find that a subset of cultured fibroblasts has megabase-scale CNVs. Recently, small CNVs (<1 Mb) have been estimated to occur in skin fibroblasts at a frequency of perhaps 30%; however, no large CNVs were reported in this study (22). In order to study single somatic cells, Abyzov *et al.* (22) reprogrammed fibroblasts and performed deep whole-genome sequencing on the hiPSC cell lines that emerged. In contrast, we analyzed single cultured fibroblasts directly using lower-resolution methods that cannot resolve small CNVs (<1 Mb). Given that many large CNVs are expected to be deleterious and may adversely affect reprogramming or clonal expansion in culture, we believe that the two findings are not inconsistent.

Our single-cell genomic analysis of human neurons extends the observation of somatic mosaicism in the nervous system to the single-cell level. Several studies using bulk DNA from somatic tissues, including brain, have found CNVs among mono-

zygotic twins (23) and in different organs or brain regions from the same individual (24, 25). These studies were only able to detect CNVs present in >10% of the cells in the bulk sample and, thus, have only provided a coarse assessment of somatic mosaicism. We have shown that mosaic CNV is abundant in human neurons. Additional work will be required to address the full spectrum of somatic mutation in neurons and other cell lineages; however, it is possible that some neuronal lineages acquire genomic instability during development, which leads to subsequent diversification of neuronal genomes, or that individual neurons become prone to large-scale mutational events because of widespread DNA damage. A recent study has implicated electrophysiological activity as a source of double-strand DNA breaks in neurons (26), and small circular DNAs caused by excision have been reported in multiple somatic cell types, including neurons (27, 28). Additionally, retrotransposon activity is known to cause subchromosomal deletions and other rearrangements in human cells (29–32); thus, higher levels of retrotransposon activity during human neurogenesis (5, 33) may also contribute to the prevalence of CNVs in neuronal genomes.

The effect of somatic genome diversification on neuronal function remains unknown. One straightforward hypothesis is that neurons with different genomes will have distinct molecular phenotypes because of altered transcriptional or epigenetic landscapes. We expect that ongoing development of single-cell technologies will allow for this hypothesis to be tested by measuring multiple states of the same neuron (e.g., the genome and the epigenome, transcriptome, or proteome). We have shown that hiPSC-derived neurons recapitulate somatic variation, as observed in endogenous human neurons; thus, hiPSCs may offer a tractable system for applying single-cell approaches to understanding the consequences of somatic mosaicism. In the future, the ability to manipulate and measure genomic diversity in human neural circuits *in vitro* may help to reveal the consequences of somatic mosaicism in the brain.

References and Notes

1. S. K. Rehen *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13361–13366 (2001).
2. S. K. Rehen *et al.*, *J. Neurosci.* **25**, 2176–2180 (2005).
3. Y. B. Yurov *et al.*, *PLOS ONE* **2**, e558 (2007).
4. A. R. Muotri *et al.*, *Nature* **435**, 903–910 (2005).

5. J. K. Baillie *et al.*, *Nature* **479**, 534–537 (2011).
6. G. D. Evrony *et al.*, *Cell* **151**, 483–496 (2012).
7. E. M. Ostertag, H. H. Kazanian, *Nature* **435**, 890–891 (2005).
8. T. Singer, M. J. McConnell, M. C. Marchetto, N. G. Coufal, F. H. Gage, *Trends Neurosci.* **33**, 345–354 (2010).
9. S. L. Martin, *Nature* **460**, 1087–1088 (2009).
10. D. M. Bushman, J. Chun, *Semin. Cell Dev. Biol.* **24**, 357–369 (2013).
11. E. Vanneste *et al.*, *Nat. Med.* **15**, 577–583 (2009).
12. N. Navin *et al.*, *Nature* **472**, 90–94 (2011).
13. K. L. Spalding, R. D. Bhardwaj, B. A. Buchholz, H. Druid, J. Frisén, *Cell* **122**, 133–143 (2005).
14. F. B. Dean *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 5261–5266 (2002).
15. A. Adey *et al.*, *Genome Biol.* **11**, R119 (2010).
16. R. S. Lasken, *Biochem. Soc. Trans.* **37**, 450–453 (2009).
17. R. S. Lasken, T. B. Stockwell, *BMC Biotechnol.* **7**, 19 (2007).
18. A. B. Olshen, E. S. Venkatraman, R. Lucito, M. Wigler, *Biostatistics* **5**, 557–572 (2004).
19. K. J. Brennand *et al.*, *Nature* **473**, 221–225 (2011).
20. L. C. Laurent *et al.*, *Cell Stem Cell* **8**, 106–118 (2011).
21. T. Baslan *et al.*, *Nat. Protoc.* **7**, 1024–1041 (2012).
22. A. Abyzov *et al.*, *Nature* **492**, 438–442 (2012).
23. C. E. Bruder *et al.*, *Am. J. Hum. Genet.* **82**, 763–771 (2008).
24. M. O'Huallachain, K. J. Karczewski, S. M. Weissman, A. E. Urban, M. P. Snyder, *Proc. Natl. Acad. Sci. U.S.A.* **109**, 18018–18023 (2012).
25. A. Piotrowski *et al.*, *Hum. Mutat.* **29**, 1118–1124 (2008).
26. E. Suberbielle *et al.*, *Nat. Neurosci.* **16**, 613–621 (2013).
27. T. Maeda *et al.*, *Biochem. Biophys. Res. Commun.* **319**, 1117–1123 (2004).
28. Y. Shibata *et al.*, *Science* **336**, 82–86 (2012).
29. N. Gilbert, S. Lutz-Prigge, J. V. Moran, *Cell* **110**, 315–325 (2002).
30. P. A. Callinan *et al.*, *J. Mol. Biol.* **348**, 791–800 (2005).
31. N. Gilbert, S. Lutz, T. A. Morrish, J. V. Moran, *Mol. Cell. Biol.* **25**, 7780–7795 (2005).
32. D. E. Symer *et al.*, *Cell* **110**, 327–338 (2002).
33. N. G. Coufal *et al.*, *Nature* **460**, 1127–1131 (2009).

Acknowledgments: We thank D. Husband (Salk), L. Moore (Salk), S. Jackmaert (KU Leuven), R. Layer (University of Virginia) and R. Clark (University of Virginia) for technical assistance; A. Prorock and Y. Bao (UVA Sequencing Core) for DNA sequencing; and all members of the Gage laboratory for critical feedback on the project. We thank M. L. Gage for editorial comments. F.H.G. thanks the Center for Academic Research and Training in Anthropogeny (CARTA) for support and perspective. This work was supported by a Crick-Jacobs

Junior Fellowship to M.J.M.; a University of Leuven (KU Leuven) SymbioSys grant (PFV/10/016) to J.R.V. and T.V.; a Mather's Family Foundation grant, a NIH TR01 (R01 MH095741), the J.P.B. Foundation, Annette Merle-Smith, and a Helmsley Foundation grant to F.H.G.; and an NIH New Innovator Award (DP20D006493-01) and Burroughs Wellcome Fund Career Award to I.M.H. Human tissue was obtained from the National Institute for Child Health and Human Development (NIH) Brain and Tissue Bank for Developmental Disorders at the University of Maryland, Baltimore, MD, contract HHSN2752009000011C, ref. no. N01-HD-9-011. The hiPSC lines used in this study are available from the Coriell Cell Repository. Microarray data have been deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GSE51538), and DNA sequence data have been deposited in the NCBI Short-Read Archive (SRP030642).

Supplementary Materials

www.sciencemag.org/content/342/6158/632/suppl/DC1
Materials and Methods
Figs. S1 to S11
Tables S1 to S3
References (34–46)

19 July 2013; accepted 1 October 2013
10.1126/science.1243472

Resident Neural Stem Cells Restrict Tissue Damage and Neuronal Loss After Spinal Cord Injury in Mice

Hanna Sabelström,¹ Moa Stenudd,¹ Pedro Réu,^{1,2} David O. Dias,¹ Marta Elfineh,¹ Sofia Zdunek,¹ Peter Damberg,³ Christian Göritz,¹ Jonas Frisén^{1*}

Central nervous system injuries are accompanied by scar formation. It has been difficult to delineate the precise role of the scar, as it is made by several different cell types, which may limit the damage but also inhibit axonal regrowth. We show that scarring by neural stem cell–derived astrocytes is required to restrict secondary enlargement of the lesion and further axonal loss after spinal cord injury. Moreover, neural stem cell progeny exerts a neurotrophic effect required for survival of neurons adjacent to the lesion. One distinct component of the glial scar, deriving from resident neural stem cells, is required for maintaining the integrity of the injured spinal cord.

Scar formation in the injured spinal cord limits secondary damage by providing mechanical stability and restricting infiltration by inflammatory cells (1–3) but also contributes to the failure of severed axons to regrow (2, 4–6). Resident neural stem cells give rise to the majority of new astrocytes making the glial scar in the injured spinal cord (7, 8). Transplantation of stem cells, or stem cell–derived cells, to the injured spinal cord can improve functional recovery (9). The mechanisms underlying this effect are not fully understood, but trophic effects as well as remyelination of spared axons appear most important (9–13). Modulating the response of endogenous neural stem cells may offer an alternative to cell transplantation, but this requires an understanding of the function of these cells in response to spinal cord injury.

The neural stem cells in the adult mouse spinal cord constitute a small cell population, denoted ependymal cells, lining the central canal (8, 14–16). To address the role of ependymal cells in the spinal cord injury response, we selectively blocked their generation of progeny by deleting all *Ras* genes, which are required for cells to go through the G₁ phase of mitosis (17, 18). We established FoxJ1-CreER mice, which allow conditional genetic recombination specifically in ependymal cells in the adult spinal cord after administration of tamoxifen (8), homozygous for *H-Ras* and *N-Ras* null alleles and homozygous for floxed *K-Ras* alleles (17). The mice also carried a *R26R-YFP* (YFP, yellow fluorescent protein) reporter allele to allow visualization of recombination. Tamoxifen was administered to adult mice to delete *K-Ras* (we refer to these mice as FoxJ1-rasless), and matched mice with the same genotype receiving vehicle (referred to as FoxJ1 mice) were used as controls (fig. S1).

Ependymal cell proliferation was selectively reduced in the intact spinal cord (Fig. 1, A and B),

as well as after a dorsal funiculus incision in FoxJ1-rasless mice (Fig. 1, C to E, and fig. S2, A to E). Ependymal cell progeny starts migrating from the ependymal layer toward the injury site within 3 days after injury in control mice, where it almost exclusively differentiates to scar-forming astrocytes (7, 8, 14, 19). There was no migration of recombined cells from the ependymal layer in FoxJ1-rasless mice (Fig. 1F and fig. S2, F and G), establishing this as a suitable system for assessing the role of neural stem cell progeny after spinal cord injury.

We made transverse incisions by cutting the dorsal funiculus and dorsal horns at C4 in matched FoxJ1 and FoxJ1-rasless mice (*n* = 14 in each group), and we analyzed these animals 14 weeks later. All FoxJ1 control mice developed a dense glial scar at the site of the lesion (Fig. 1G). In contrast, only 3 out of 14 FoxJ1-rasless mice developed largely normal scars at the injury site. The majority of FoxJ1-rasless mice (79%) failed to form compact scar tissue and had varying degrees of tissue defects, ranging from small cavities (21%) and less compact scars with larger cavities (29%) to a single large cyst (29%) occupying the lesion area (Fig. 1, H to J).

Pericytes play a key role in spinal cord scar formation by giving rise to the fibrotic compartment of the scar (18). The fibrotic compartment was enlarged in FoxJ1-rasless mice compared with FoxJ1 control mice, suggesting that increased fibrosis may partly compensate for the absence of ependymal cell progeny (fig. S3A). Similarly, there appeared to be a compensatory increase in scarring by resident astrocytes, which were unrecombined and not deriving from ependymal cells (Fig. 1, G to J, and fig. S3B). Resident astrocytes are molecularly and, potentially, functionally distinct from astrocytes generated by ependymal cells, which are mostly negative for glial fibrillary acidic protein (8). In addition to resident astrocytes (20), oligodendrocyte lineage cells have also been suggested to generate scar-forming astrocytes in the

¹Department of Cell and Molecular Biology, Karolinska Institute, SE-171 77 Stockholm, Sweden. ²Center for Neuroscience and Cell Biology, University of Coimbra, 3004-517 Coimbra, Portugal. ³Department of Clinical Science, Intervention and Technology, Karolinska Institute, SE-171 77 Stockholm, Sweden.

*Corresponding author. E-mail: jonas.frisen@ki.se