Science Advances

AAAS

advances.sciencemag.org/cgi/content/full/2/12/e1601941/DC1

Supplementary Materials for

A single splice site mutation in human-specific *ARHGAP11B* causes basal progenitor amplification

Marta Florio, Takashi Namba, Svante Pääbo, Michael Hiller, Wieland B. Huttner

Published 7 December 2016, *Sci. Adv.* **2**, e1601941 (2016)
DOI: 10.1126/sciadv.1601941

This PDF file includes:

Supplementary Figures

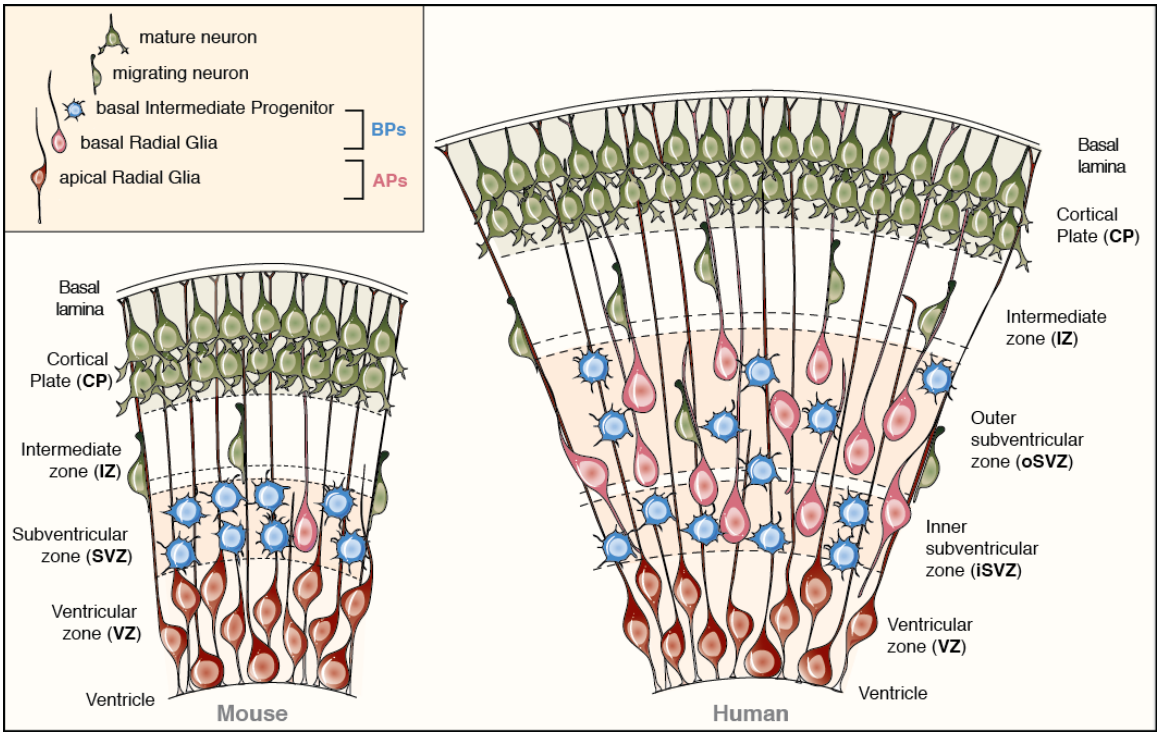

fig. S1. Cartoon illustrating APs and BPs in embryonic mouse and fetal human neocortex. Mouse is left, human is right. Note the increase in BPs and the expansion of the SVZ in fetal human neocortex. For more details, see (*13*).

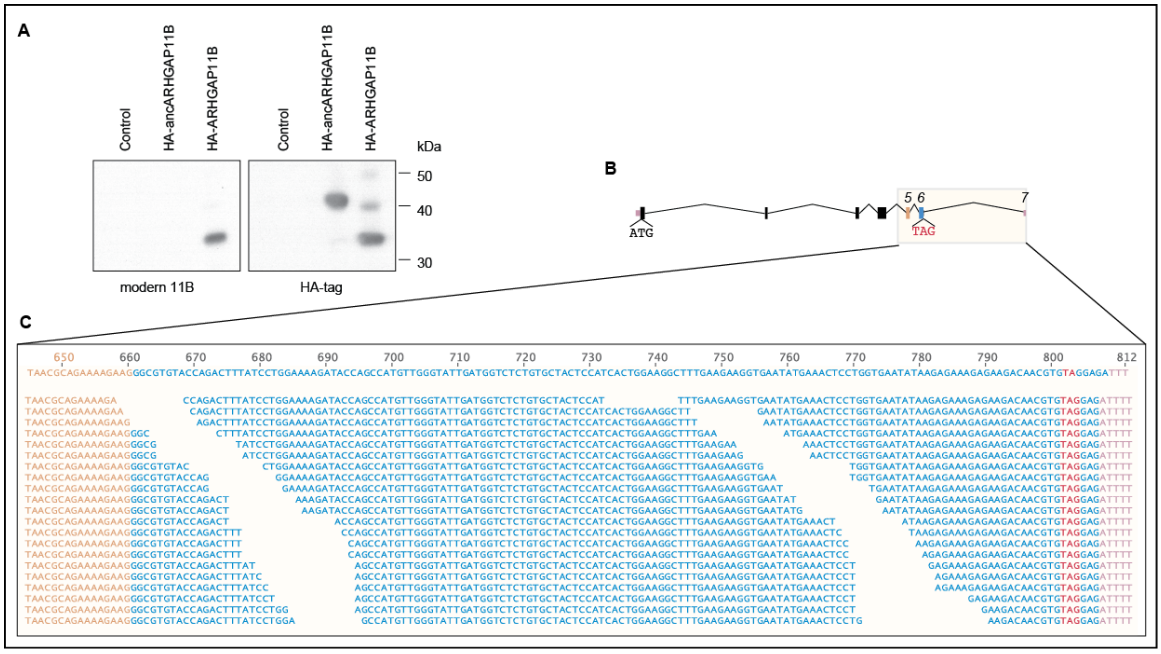

fig. S2. Presence of the novel, human-specific C-terminal sequence in modern but not ancestral ARHGAP11B. (**A**) COS-7 cells were transfected with pCAGGS (Control, empty vector), pCAGGS-*HA-ancARHGAP11B* encoding HA-tagged (4 kDa) ancARHGAP11B (37+4 kDa) or pCAGGS-*HA-ARHGAP11B* encoding HA-tagged (4 kDa) modern ARHGAP11B (30+4 kDa). Total cell lysates prepared 24 h after transfection were subjected to immunoblotting using either an anti-HA antibody (right) or a monoclonal antibody raised against the C-terminal 21 amino acid residues of modern ARHGAP11B (see Methods) (left). Note that both modern and ancestral HA-tagged ARHGAP11B are detected by the anti-HA antibody, whereas only modern but not ancestral ARHGAP11B is detected by the monoclonal antibody against the C-terminal sequence of modern ARHGAP11B. The positions of molecular weight markers are indicated. (**B**) Schematic of the unspliced *ARHGAP11B* pre-mRNA (see also Fig. 1). Yellow box highlights exon 5 (where the new splice donor site is located), exon 6 (where the translational stop codon is located) and exon 7 (where the 3'UTR is located). (**C**) RNA-seq reads overlapping to *ARHGAP11B* mRNA obtained from a RNA-seq dataset previously produced from basal radial glia purified from fetal human neocortex (*23*). A selection of reads covering *ARHGAP11B* from 5' to the new splice donor site to 3' to the translational stop codon (TAG, red) is shown, to demonstrate the presence in the relevant human cortical progenitor cells of the nucleotide sequence in the *ARHGAP11B* mRNA that encodes for the novel C-terminal sequence. The *ARHGAP11B* reference cDNA sequence and respective nucleotide numbers are shown in the top row. Nucleotides from exon 5, exon 6 and exon 7 are shown in ocher, blue and violet, respectively.

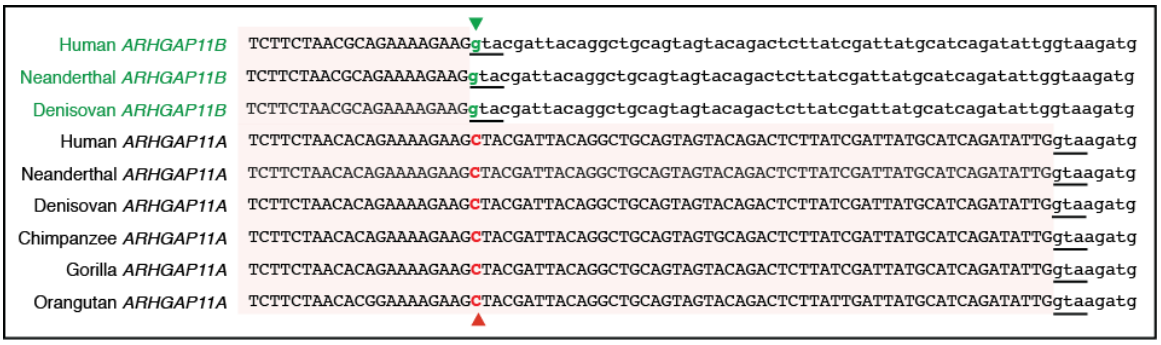

fig. S3. Reconstruction of the ancestral splice donor site. Alignment of modern human, Neanderthal and Denisovan *ARHGAP11B* and *ARHGAP11A* sequences with chimpanzee, gorilla and orangutan *ARHGAP11A* ortholog sequences. The boundaries between exons 5 (pink rectangles) and introns 5 are shown. Arrowheads indicate the C-to-G nucleotide substitution (red-to-green) in exons 5. Notice that the 'G' is only found in archaic and modern human *ARHGAP11B*. Exonic and intronic nucleotide sequences are displayed in upper and lower cases, respectively. Horizontal lines indicate splice donor sites.

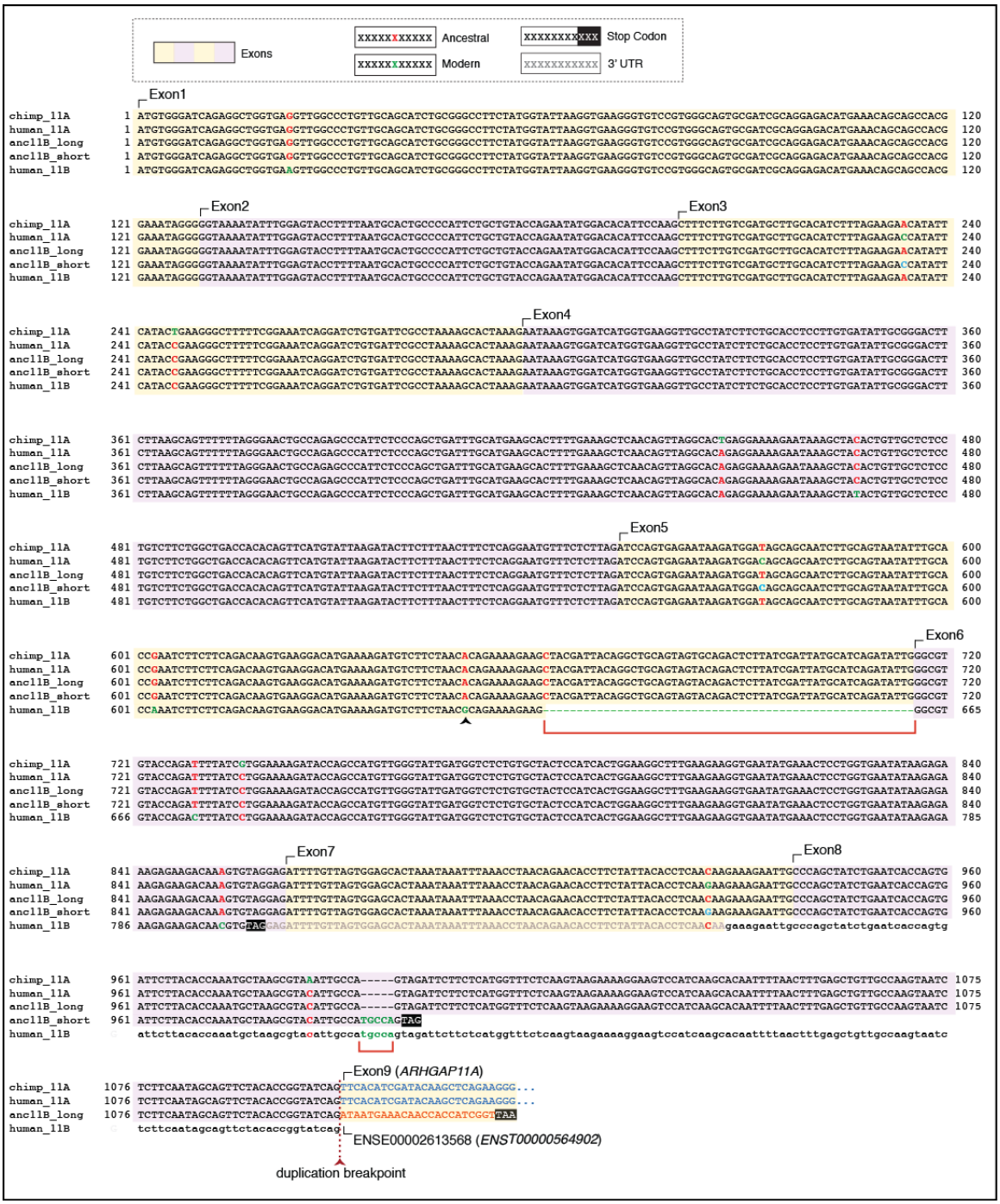

fig. S4. Ancestral *ARHGAP11B* nucleotide sequence. Alignments of the protein-encoding sequences of chimpanzee *ARHGAP11A,* modern human *ARHGAP11A*, two ancestral human *ARHGAP11B* (*anc11B_long* and *anc11B_short*; see *description of data* below) and modern human *ARHGAP11B* mRNAs. All RNA sequences are depicted with T instead of U. Red and green letters indicate ancestral and modern nucleotides, respectively, as predicted from multiple alignments with various non-human primates. Turquoise letters in *anc11B_short* reflect the use of the *ARHGAP11A* template to clone

anc11B_short. Exons are numbered and indicated by alternating colors. Arrowhead in exon 5 (at position 649 of the coding sequence) points to the A→G nucleotide substitution in the modern *ARHGAP11B* gene that leads to a threonine-to-alanine change at residue 217 (see fig. S5). Red bracket underneath the ancestral exon 5 sequences highlights the 55-nucleotide deletion in the *ARHGAP11B* mRNA that leads to a frameshift and novel human-specific C-terminal 47 amino acid sequence [see also Fig. 1 and fig. S5, and (*23*)]. Stop codons are highlighted by black background. The 3'-UTR sequence of the modern *ARHGAP11B* mRNA is indicated in grey, with its 3' end predicted from the polyA stretches present in various cDNA reads [obtained from RNA-Seq datasetset in (*23*)]; the subsequent, usually untranscribed sequence of the duplicated *ARHGAP11B* gene, which is homologous to the 3'-end of *ARHGAP11A* exons 7 and exon 8, is indicated by lowercase letters. Red bracket underneath the modern human *ARHGAP11B* exon-8 sequence indicates a 5-bp microduplication (green TGCCA; note the TGCCATGCCA repeat). This insertion is found in modern human *ARHGAP11B* and was used in the design of *anc11B_short* (see *description of data* below). Note that this insertion results in a frameshift leading to a stop codon one nucleotide 3' to the insertion. Red vertical dotted line and arrowhead indicate the breakpoint of the *ARHGAP11B* gene duplication. Note the divergence in the *ARHGAP11A* and *anc11B_long* sequences downstream of the duplication breakpoint: blue nucleotides indicate the 5' end of *ARHGAP11A* exons 9. Orange nucleotides indicate the 5' end of exon *ENSE00002613568*, found in the *ARHGAP11B* locus 3' to the breakpoint. Note the predicted in-frame stop codon found in this exon.

Description of data: A putative ancestral *ARHGAP11B* cDNA (*anc11B_short*) for experimental use was reconstructed as follows. First, we assumed that there would be no termination of transcription in exon 7, as is the case in modern *ARHGAP11B* (Fig. 1A and 1B). Next, we reverted the C→G substitution at position 661 that in the modern *ARHGAP11B* mRNA creates a new splice donor site (Fig. 1B), resulting in loss of 55 nucleotides (Fig. 1B), reading-frame-shift (Fig. 1C) and translational stop codon at the 3' end of exon 6. To this end, we cloned the first 997 bps of the *ARHGAP11A* cDNA (i.e. exons 1-7 plus part of exon 8), which contains a C at position 661. We then introduced a 5-bp insertion in exon 8 (green sequence), which is fixed at this position in the modern *ARHGAP11B* gene (*31*). This ancestral *ARHGAP11B* cDNA, which is devoid of a 5'-UTR and 3'-UTR, is characterized by the ancestral exon 5 splice donor site configuration, contains the 55 nucleotides that are lacking in the modern *ARHGAP11B* mRNA, and hence exhibits an ancestral *ARHGAP11* reading frame, with translation proceeding through exon 7 and terminating in exon 8 due to the in-frame stop codon caused by the 5-bp insertion. To account for the possibility that this 5-bp insertion occurred after the *ARHGAP11* gene duplication event, we depict a second version of a putative ancestral *ARHGAP11B* cDNA (*anc11B_long*) without this insertion. In this ancestral *ARHGAP11B* mRNA, exon 8 is predicted to be spliced to exon *ENSE00002613568*, and hence translation would proceed through exon 8 and terminate in exon *ENSE00002613568* due to an in-frame stop codon. The latter exon is transcribed (though not translated) in some splice variants of modern *ARHGAP11B* (ENST00000563110, ENST00000564902, ENST00000622744), extending its 3'-UTR.

fig. S5. Ancestral ARHGAP11B protein sequence. Alignments of amino acid sequences of chimpanzee ARHGAP11A, modern human ARHGAP11A, two ancestral ARHGAP11B (anc11B_long and anc11B_short), and modern human ARHGAP11B, as predicted from the respective mRNAs (see fig. S4). Red and green letters indicate ancestral and modern residues, respectively. Turquoise letter in anc11B_short reflects the use of the *ARHGAP11A* template to clone *anc11B_short* (see fig. S4). Encoding exons are numbered and indicated by alternating colors; overlapping colors at junctions between exons indicate residues whose respective codons overlap splice sites. Black box indicates GAP domain (ARHGAP11A, ancestral ARHGAP11B: residues 46-246; ARHGAP11B:

residues 46-220). Arrowhead at residue 217 points to a threonine-to-alanine change in modern ARHGAP11B due to the A→G nucleotide substitution in exon 5 (see fig. S4). See fig. S8 for the functional tests of the A217T modern-to-ancestral reversion. Arrowhead at residue 221 in modern ARHGAP11B points to the first amino acid of the novel human-specific C-terminal 47-residue sequence, which is due to the frameshift-inducing 55-nucleotide deletion in modern *ARHGAP11B* mRNA [see Fig. 1, fig. S4 and (*23*)]. The amino acid sequence translated from *anc11B_long* by a fusion transcript between exon 8 and *ENSE00002613568* (see fig. S4) is highlighted in pink.

Description of data: The ancestral ARHGAP11B examined experimentally in the present study, encoded by *anc11B_short*, is a 333-amino-acid protein, with 331 residues being identical to the N-terminal sequence of ARHGAP11A. The ancestral ARHGAP11B predicted by the *anc11B_long* sequence would be a 376-amino-acid protein, with 368 residues being identical to the N-terminal sequence of ARHGAP11A. Like ARHGAP11A, the ancestral ARHGAP11B encoded by *anc11B_short* contains a full GAP domain, and so would the ancestral ARHGAP11B predicted by *anc11B_long*.

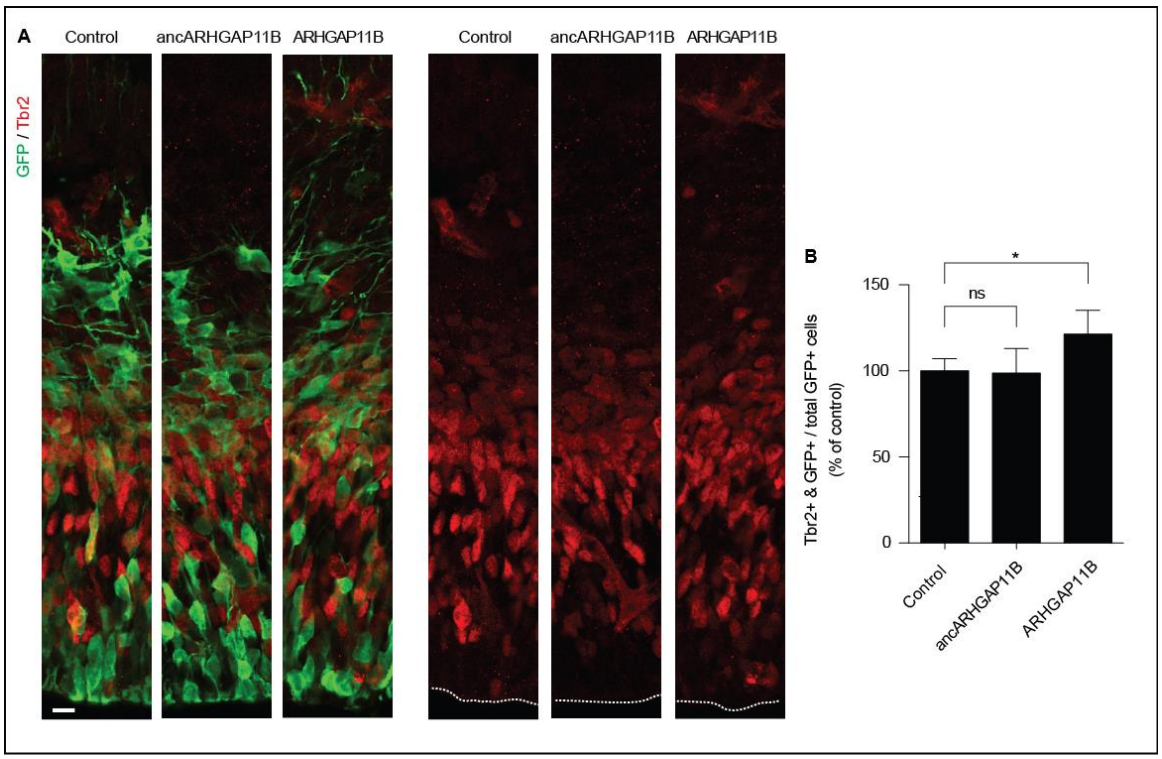

fig. S6. Modern but not ancestral ARHGAP11B increases Tbr2$^+$ progenitors in mouse developing neocortex. *In utero* electroporation of pCAGGS-GFP together with empty vector (Control), ancestral *ARHGAP11B* (*ancARHGAP11B*) and modern *ARHGAP11B* expression plasmids in E13.5 mouse neocortex followed by analysis at E14.5. (**A**) GFP (green) and Tbr2 (red) immunofluorescence. Scale bar, 10 µm. (**B**) Quantification of the proportion of GFP+ cells that are Tbr2+. Data are expressed as percentage of Control and are the mean of 3 (Control), 4 (ancARHGAP11B) and 5 (ARHGAP11B) embryos from three litters; error bars, SD; *p<0.05; ns, not significant (ANOVA with Bonferroni post-hoc test for multiple comparisons).

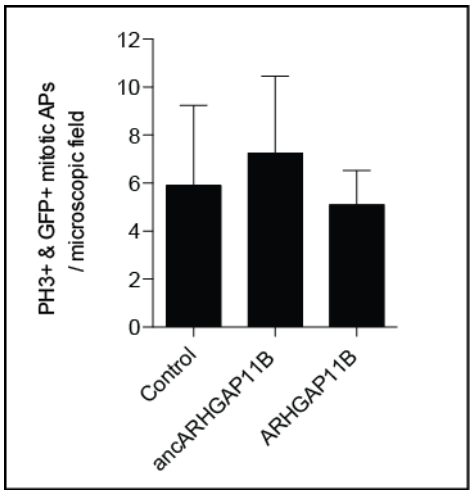

fig. S7. Neither modern nor ancestral ARHGAP11B affects the abundance of mitotic APs in mouse developing neocortex. *In utero* electroporation of pCAGGS-GFP together with empty vector (Control), ancestral *ARHGAP11B* (*ancARHGAP11B*) and modern *ARHGAP11B* expression plasmids in E13.5 mouse neocortex, followed at E14.5 by quantification of targeted mitotic APs, identified by the presence of both GFP & PH3 immunofluorescence at an apical location (see Fig. 2). Data are the mean of 4 (Control), 6 (ancARHGAP11B) and 5 (ARHGAP11B) embryos from three litters, expressed per 100-µm-wide field of cortical wall; error bars, SD; ns, not significant (ANOVA with Bonferroni post-hoc test for multiple comparisons).

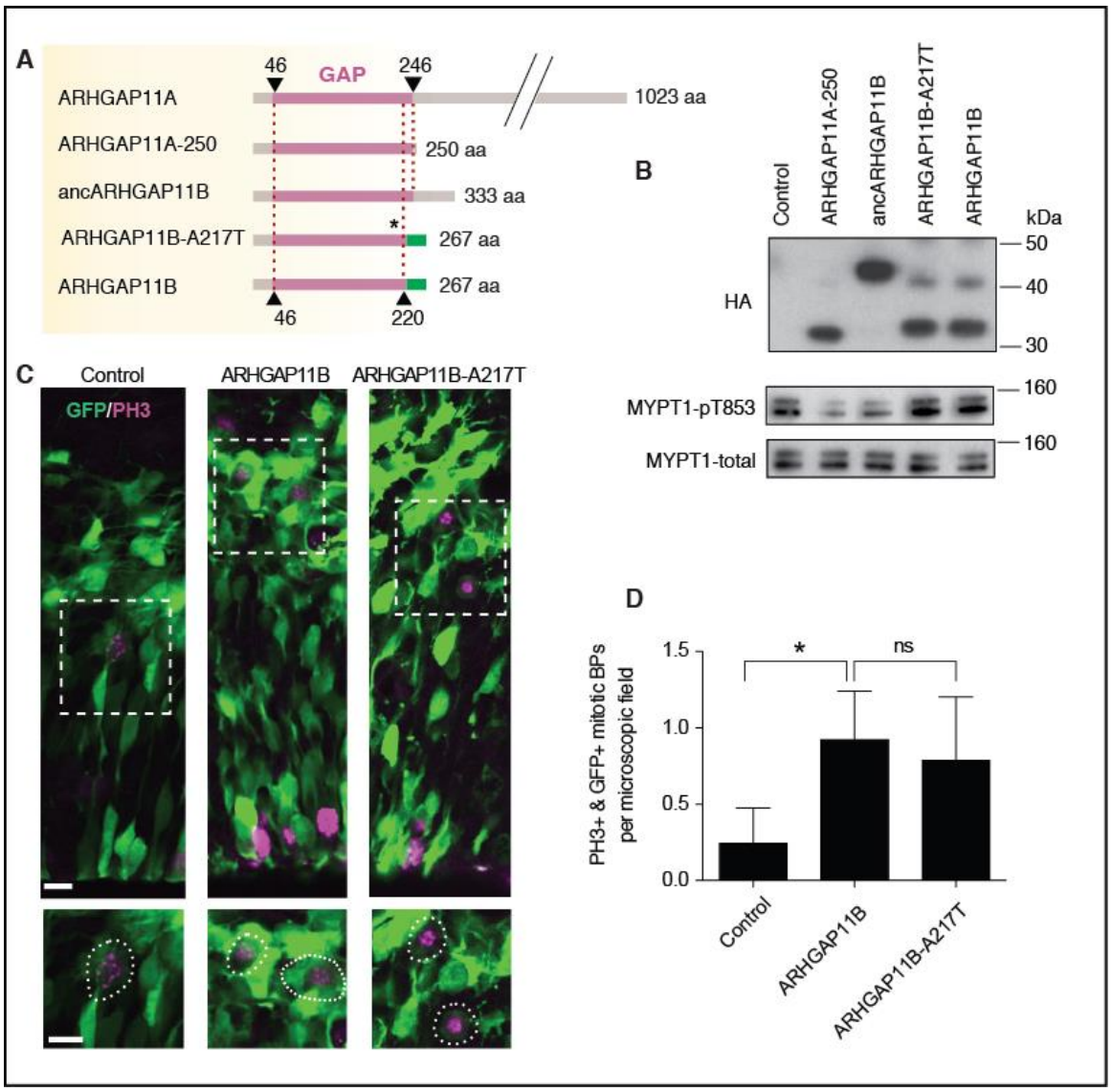

fig. S8. ARHGAP11B-A217T mutant, like modern ARHGAP11B, lacks GAP activity and increases BPs in the mouse developing neocortex. (**A**) Domain structures of ARHGAP11A, ARHGAP11A-250, ancestral (anc) ARHGAP11B (anc11B_short in fig. S5), mutant ARHGAP11B-A217T (see fig. S5), and modern ARHGAP11B. Purple, GAP domain; green, C-terminal domain unique to modern ARHGAP11B. Black arrowheads indicate first and last residues of the full-length or truncated GAP domain. Asterisk indicates the alanine residue at position 217 in modern ARHGAP11B that was reverted to threonine as in ancestral ARHGAP11B. (**B**) Electrophoretic mobility (top) and GAP activity assay (middle and bottom, see also Fig. 2B) of HA-tagged (4 kDa) ARHGAP11A-250 (28+4 kDa), ancARHGAP11B (37+4 kDa), mutant ARHGAP11B-A217T (30+4 kDa) and modern ARHGAP11B (30+4 kDa). COS-7 cells were transfected with pCAGGS-empty-vector (Control), pCAGGS-*HA-ARHGAP11A-250*, pCAGGS-*HA-ancARHGAP11B*, pCAGGS-*HA-ARHGAP11B-A217T* and pCAGGS-*HA-ARHGAP11B*. Total cell lysates prepared 24 h after transfection were subjected to immunoblotting for the HA-tag (B), or threonine 853-phosphorylated (pT853, C top) and total (C bottom)

myosin phosphatase target subunit 1 (MYPT1). The positions of molecular weight markers are indicated. (**C**, **D**) *In utero* electroporation of E13.5 mouse neocortex with pCAGGS-*GFP* together with pCAGGS-empty-vector (Control), pCAGGS-*ARHGAP11B* and pCAGGS-*ARHGAP11B-A217T*, followed by analysis at E14.5. (**C**) GFP (green) and phosphohistone H3 (PH3, magenta) immunofluorescence, labeling targeted and mitotic cells, respectively. Areas in dashed boxes are shown at higher magnification at the bottom. Scale bars, 10 µm. (**D**) Quantification of targeted mitotic BPs at E14.5, identified by the presence of both GFP & PH3 immunofluorescence at a basal location. Data are the mean of 3 embryos per condition from 3 litters, expressed per 100-µm-wide field of cortical wall; error bars, SD; *$p<0.05$; ns, not significant (ANOVA with Bonferroni post-hoc test for multiple comparisons).

Description of data: Of the four amino acid changes within residues 1-220 where ancestral ARHGAP11B differs from modern ARHGAP11B (arginine-8, aspartate-78, threonine-156, threonine-217), threonine-217 is likely to be the most relevant as this residue constitutes a predicted phosphorylation site. We therefore mutated alanine-217 in modern ARHGAP11B to threonine and examined this mutant for RhoGAP activity and its ability to increase the size of the BP pool. Using the same cell transfection-based assay revealing RhoGAP activity by a decreased MYPT1-pT853 level as in Fig. 2, the ARHGAP11B-A217T mutant, like modern ARHGAP11B, did not exhibit RhoGAP activity, in contrast to ancestral ARHGAP11B. This indicates that threonine-217 is not able to restore RhoGAP activity of modern ARHGAP11B. Upon expression in mouse embryonic neocortex, the ARHGAP11B-A217T mutant did not differ from modern ARHGAP11B in the ability to increase the abundance of targeted BPs in mitosis (D). This indicates that threonine-217 does not impair the ability of modern ARHGAP11B to increase the size of the BP pool.