SUPPLEMENTARY MATERIAL

BOC is a Modifier Gene in Holoprosencephaly

Mingi Hong, Kshitij Srivastava, Sungjin Kim, Benjamin L. Allen, Daniel J. Leahy, Ping Hu, Erich Roessler, Robert S. Krauss, and Maximilian Muenke

SUPPLEMENTARY METHODS

Patient selection

All human subjects provided informed consent for their research participation and clinical data collection in accordance with NHGRI IRB approved protocols. Our patient collection often included parents of affected individuals, affected siblings, and have been systematically studied using a variety of analysis platforms, over many years, including chip array, targeted capture and exome sequencing in some instances.

Patient screening

High-throughput screening of *BOC* was performed using High Resolution Melting (HRM, as described in Kauvar *et al*. 2010) among 360 patients with HPE. The entire clinical spectrum of HPE was represented in the cohort. In addition, 384 unrelated individuals were screened as controls. These were acquired as de-identified samples from Sigma-Aldrich Corporation and the European Collection of Cell Cultures (Human Random Control Panels 1, 3, and 4). HPE and control samples with melting profiles that deviated from wild type melting curves were directly sequenced for variant confirmation.

Gene analysis

The sequence of the *BOC* gene (NM_033254.3) was obtained from public database programs (http://www.ncbi.nlm.nih.gov) and the Bioinformatics site annotation provided by http://www.genome.ucsc.edu. Guidelines for the naming of the sequence variants conform to the recommendations of the human nomenclature committee (www.hgvs.org/mutnomen). Variants not already described in dbSNP were submitted to ClinVar (https://www.ncbi.nlm.nih.gov/clinvar).

Oligonucleotide primers were designed using Oligo^{TM} 6.8 for the eighteen coding exons (exons 2-20) and exon/intron boundaries of *BOC*. The same primers were used for HRM, direct PCR amplification, and sequencing.

HRM Roche LightCycler® 480: Each sample was amplified in a reaction volume of 7.5 µl, using 10 ng of DNA template, 3.75 µl of Roche High-Resolution Melting Master (containing FastStart Taq DNA polymerase, reaction buffer, dNTP mix, and HRM dye), 0.9 µl of magnesium chloride (Roche), 1.7 µl of molecular grade water (Cellgro), and 0.2 µM of each primer. All reactions were done using a Roche LightCycler® 480 (Roche, Ind.). PCR amplification parameters were: incubation at 95°C for 10 min, followed by 45 cycles of denaturation at 95°C for 10 sec, annealing 53°C for 15 sec, and extension at 72°C for 20 sec.

PCR assay for direct DNA sequencing: Each sample was amplified using a Roche system, in a reaction volume of 25 µl, using 25 ng of DNA template, 2.5 µl of 10X PCR buffer+$MgCl_2$, 0.2 µl of FastStart Taq (5U/ µl), 0.2 mM of dNTP, 20.3 µl of molecular grade water (Cellgro), and 0.2 µM of each primer. All reactions were done using a PTC-225 thermocycler (MJ Research, Mass). PCR amplification parameters were: incubation at 95°C for 4 min, followed by 39 cycles of denaturation at 95°C for 30 sec, annealing at specific exon temperature (see Suppl. Table S1) for 30 sec, and extension at 72°C for 1 min, with a final extension step of 72°C for 7 min. Details of PCR primers are available upon request.

High-Resolution DNA melting acquisition and analysis

High-Resolution DNA Melting (HRM) involves PCR in the presence of saturating DNA binding dyes. Melting profiles and melting temperatures (Tm) are analyzed based on fluorescence data generated during DNA melting. Sequence variations can be assessed based on melting analysis.

The LightCycler® 480 II system (Roche Applied Science, Indianapolis, Ind.) was used to identify putative sequence variants in *BOC.* The samples were tested as eighteen different amplicons, one for each of the coding exons (exons 3-20). First, samples were pre-incubated at 95°C for 10 min, followed by PCR amplification (above). Melting curves were generated using pre-incubation at 95°C for 1 min, pre-hold at 40°C for 1 min, and increasing temperature from 65°C to 95°C at a programmed rate of 0.2°C/sec with fluorescence acquisition of 25 times per 1°C of temperature increase. Within the 384-well plates, either 8 or 16 water-containing blank samples were used as negative controls.

HRM curve analysis was performed according to the manufacturer's recommendations using LightCycler® 480 Software release v.1.5.0 (Idaho Technology Inc., Idaho and Roche Applied Sciences, Ind.). Amplicons were normalized and temperature-shifted fluorescence-over-temperature plots were identified. Samples with melting profiles that deviated from the wild type melting profiles were then column-purified and sequenced bi-directionally to confirm the presence of sequence changes.

DNA Sequencing

Bidirectional DNA sequencing of *BOC* was performed on a selection of samples with variant HRM melting profiles. For HRM variants, sequencing was performed on the PCR product amplified during the HRM analysis following column-purification; sequence variants were then confirmed following PCR re-amplification of the original DNA sample. Sequencing was performed at the DNA Sequencing Facility, National Institute of Neurological Disorders and Stroke (NINDS), NIH. DNA sequences were manually annotated and analyzed using Sequencher

4.7 (Gene Codes Corporation). The reference sequence used for *BOC* was NM_033254.3 (NCBI, 2009/human GRCh37/hg19).

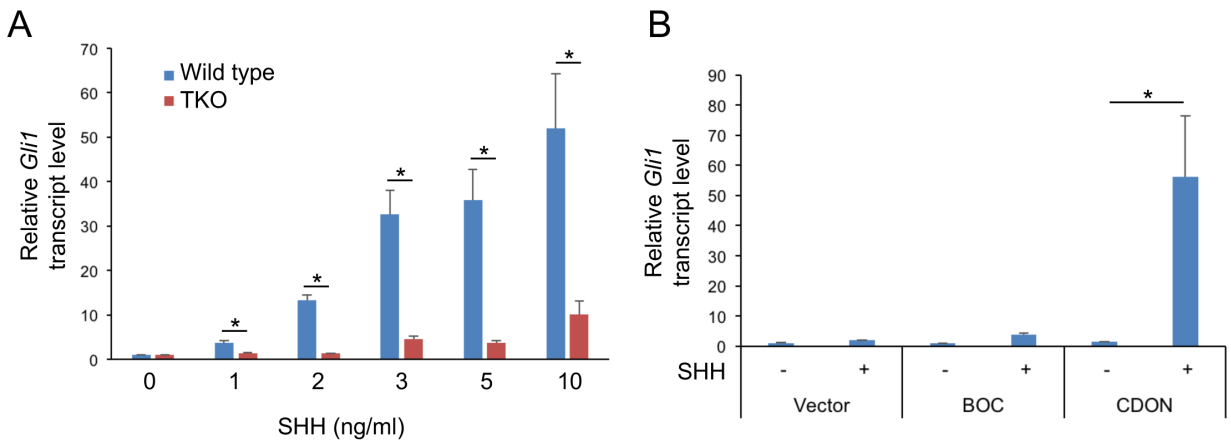
Bioinformatic analysis

A suite of publically available databases and Annovar annotations were interrogated using the hg19 coordinates and basepair changes detected either by Sanger sequencing or targeted capture. Of the eleven measures of amino acid substitution consequences [SIFT, Polyphen2[HVID], Polyphen2[HVAR], LRT, MutationTaster, MutationAssessor, FATHMM, Provean, Fathmm-MKL, MetaSVM, MetaLR] each variant was assigned to benign (green) or damaging (pink) based on the consensus of these tools (see Table 1). The Annovar consensus interpretations were used as a single factor when synthesizing a prediction towards the overall pathogenicity assessment.

Cell culture, transfection and SHH signaling assays

Site-directed variants of *BOC* harboring a C-terminal FLAG epitope tag in the pBabePuro expression vector (Kang et al., 2002) were constructed by GENEWIZ (South Plainfield, NJ) and verified by sequencing. 293T cells were cultured in DMEM with 10% fetal bovine serum (Gibco) and transfected using Effectene (Qiagen). Western blot analysis was performed as described (Kang et al., 2002), with ECL system detection (Thermo Scientific). Antibodies were from the following suppliers: anti-BOC (R&D Systems); anti-β-actin (Abcam); anti-FLAG (Sigma) Control and *Cdon*$^{-/-}$*; Boc*$^{-/-}$*; Gas1*$^{-/-}$ triple mutant MEFs (Mathew et al., 2014) were cultured in DMEM with 10% calf bovine serum (ATCC). For SHH signaling assays, MEFs were reverse transfected with plasmids diluted in Opti-MEM with Lipofectamine 2000 (Invitrogen)

and transferred onto 6-well plates (Bae et al., 2011). Transfection efficiency was about 30%, as determined with a GFP-expressing vector. Transfected MEFs were grown in DMEM with 10% CBS for 24 hours until nearly confluent, then changed to DMEM with 2% CBS. 24 hours after serum starvation, MEFs were treated with recombinant SHH (StemRD) for 48 hours. Total RNA was extracted using Trizol (Invitrogen). The cDNAs were reverse-transcribed from one μg of total RNA with random primers and Superscript III reverse transcriptase (Invitrogen). qPCR reactions for *Gli1* (Bae et al., 2011) were performed on an iCycler iQ Real-Time PCR Detection System (BioRad).



Supp. Fig. S1. Responses of TKO MEFs to SHH

(A) qRT-PCR analysis of *Gli1* expression in wild type and TKO (*Cdon*$^{-/-}$;*Boc*$^{-/-}$;*Gas1*$^{-/-}$) MEFs treated with various amounts of SHH. *Gli1* mRNA levels were normalized to *Gapdh* mRNA levels. Note that wild type MEFs were much more responsive to SHH at all doses. *$p<0.01$ with Student's t-test.

(B) Expression of CDON, but not BOC, restored the response to SHH by TKO cells. Cultures were treated with 5 ng/ml SHH and *Gli1* expression was quantified as in (A). *$p<0.05$ with Student's t-test.

Values in (A) and (B) are means ± SEM, encompassing duplicate determinations from independent experiments performed at least three times.