A genome-wide association study implicates the *BMP7* locus as a risk factor for nonsyndromic metopic craniosynostosis

Cristina M Justice,[1,a] Araceli Cuellar,[2,a] Krithi Bala,[2,a] Jeremy A Sabourin,[1] Michael L Cunningham,[3] Karen Crawford,[4] Julie M Phipps,[4,5] Yan Zhou,[4] Deirdre Cilliers,[5] Jo C Byren,[6] David Johnson,[6] Steven A Wall,[6] Jenny E V Morton,[7,8] Peter Noons,[8] Elizabeth Sweeney,[9] Astrid Weber,[9] Katie E M Rees,[10] Louise C Wilson,[10] Emil Simeonov,[11] Radka Kaneva,[12] Nadezhda Yaneva,[13] Kiril Georgiev,[14] Assen Bussarsky,[14] Craig Senders,[15] Marike Zwienenberg,[16] James Boggan,[16] Tony Roscioli,[17] Gianpiero Tamburrini,[18-19] Marta Barba,[18,20] Kristin Conway,[21] Val C Sheffield,[22] Lawrence Brody,[23] James L Mills,[24] Denise Kay,[25] Robert J Sicko,[25] Peter H Langlois,[26] Rachel K Tittle,[27] Lorenzo D Botto,[28] Mary M Jenkins,[29] Janine M LaSalle,[30] Wanda Lattanzi,[18,20] Andrew O M Wilkie,[4-6] Alexander F Wilson,[1,b] Paul A Romitti,[21,b,c] Simeon A Boyadjiev,[2,b,c] and the National Birth Defects Prevention Study

[a]authors with equal contribution
[b]co-senior authors
[c]co-corresponding authors

Affiliations:
[1]Genometrics Section, Computational and Statistical Genomics Branch, Division of Intramural Research, NHGRI, NIH, Baltimore, MD, USA
[2]Department of Pediatrics, University of California Davis, Sacramento, CA, USA
[3]University of Washington, Department of Pediatrics, Division of Craniofacial Medicine, Seattle Children's Craniofacial Center and Seattle Children's Research Institute, Seattle, WA, USA
[4]MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK
[5]Oxford Centre for Genomic Medicine, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[6]Craniofacial Unit, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[7]West Midlands Regional Clinical Genetics Service and Birmingham Health Partners, Birmingham Women's and Children's Hospitals NHS Foundation Trust, Birmingham, UK
[8]Birmingham Craniofacial Unit, Birmingham Women's and Children's Hospitals NHS Foundation Trust, Birmingham, UK
[9]Department of Clinical Genetics, Liverpool Women's NHS Foundation Trust, Liverpool, England, UK
[10]Clinical Genetics Service, Great Ormond Street Hospital for Children NHS Foundation Trust, London, UK
[11]National Institute of Pediatrics, Sofia Medical University, Sofia, Bulgaria
[12]Molecular Medicine Center, Department of Medical Chemistry and Biochemistry, Medical Faculty, Medical University of Sofia, Sofia, Bulgaria
[13]National Genetic Laboratory, University Hospital of Obstetrics and Gynecology "Maichin Dom", Medical University of Sofia, Sofia, Bulgaria
[14]Department of Neurosurgery, University Hospital 'St. Ivan Rilski', Medical University of Sofia, Sofia, Bulgaria
[15]Department of Otolaryngology, Head and Neck Surgery, University of California Davis, Sacramento, CA, USA
[16]Department of Neurosurgery, University of California Davis, Sacramento, CA, USA

[17]Neuroscience Research Australia, University of New South Wales, Sydney, Australia
[18]Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy
[19]Department of Neuroscience, Section of Neurosurgery, Università Cattolica del Sacro Cuore, Rome, Italy
[20]Department of Life Science and Public Health, Section of Experimental Biology, Università Cattolica del Sacro Cuore, Rome, Italy
[21]Department of Epidemiology, College of Public Health, The University of Iowa, Iowa City, IA, USA
[22]Department of Pediatrics, Division of Medical Genetics, Carver College of Medicine, The University of Iowa, Iowa City, IA, USA
[23]Gene and Environment Interaction Section, NHGRI, NIH, Bethesda, MD, USA
[24]Epidemiology Branch, *Eunice Kennedy Shriver* NICHD, NIH, Bethesda, MD, USA
[25]Division of Genetics, Wadsworth Center, NYS Department of Health, Albany, NY, USA
[26]Birth Defects Epidemiology and Surveillance Branch, Texas Department of State Health Services, Austin, TX, USA
[27]Department of Nutritional Sciences, University of Texas at Austin, Austin, TX, USA
[28]Department of Pediatrics, University of Utah School of Medicine, Salt Lake City, UT, USA
[29]National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta, GA, USA
[30]Department of Medical Microbiology and Immunology, Genome Center, and MIND Institute, University of California Davis, Davis, CA, USA

Co-corresponding Authors
Paul A Romitti, PhD
Department of Epidemiology
The University of Iowa
145 N Riverside Dr, S416 CPHB
Iowa City, IA 52242
Tel 319-335-4912
Fax 319-335-4915
email: paul-romitti@uiowa.edu

Simeon A Boyadjiev, MD
Department of Pediatrics
University of California Davis
4625 2nd Avenue, Research Building II
Sacramento, CA 95817 USA
Tel 916-703-0446
Fax 916-703-0460
email: sboyd@ucdavis.edu

Abstract Word Count=250 (Limit=250)

Main Text Word Count=5,436

Abstract

Our previous genome-wide association study (GWAS) for sagittal nonsyndromic craniosynostosis (sNCS) provided important insights into the genetics of midline CS. In this study, we performed a GWAS for a second midline NCS, metopic NCS (mNCS), using 215 non-Hispanic white case-parent triads. We identified six variants that were genome-wide significant ($P \leq 5 \times 10^{-8}$): rs781716 ($P = 4.71 \times 10^{-9}$; odds ratio [OR] = 2.44) intronic to *SPRY3*; rs6127972 ($P = 4.41 \times 10^{-8}$; OR = 2.17) intronic to *BMP7;* rs62590971 ($P = 6.22 \times 10^{-9}$; OR = 0.34), located ~155 kb upstream from *TGIF2LX*; and rs2522623, rs2573826, and rs2754857, all intronic to *PCDH11X* ($P = 1.76 \times 10^{-8}$, OR = 0.45; $P = 3.31 \times 10^{-8}$, OR = 0.45; $P = 1.09 \times 10^{-8}$, OR=0.44, respectively). We performed a replication study of these variants using an independent non-Hispanic white sample of 194 unrelated mNCS cases and 333 unaffected controls; only the association for rs6127972 ($P = 0.004$, OR = 1.45; meta-analysis $P = 1.27 \times 10^{-8}$, OR = 1.74) was replicated. Our meta-analysis examining single nucleotide polymorphisms common to both our mNCS and sNCS studies showed the strongest association for rs6127972 ($P = 1.16 \times 10^{-6}$). Our imputation analysis identified a linkage disequilibrium block encompassing rs6127972, which contained an enhancer overlapping a CTCF transcription factor binding site (chr20:55,798,821-55,798,917) that was significantly hypomethylated in mesenchymal stem cells derived from fused metopic compared to open sutures from the same probands. This study provides additional insights into genetic factors in midline CS.

Keywords: genome-wide association study; nonsyndromic, metopic craniosynostosis; *BMP7*; *SPRY3*; *TGIF2LX*; *PCDH11X*

Introduction

Craniosynostosis (CS) arises from the premature closure of one or more of the infant cranial vault sutures. This premature closure of the cranial sutures results in intracranial pressure as the infant's brain grows, which can lead to blindness, seizures, and/or brain damage (Gupta et al. 2003; Tamburrini et al. 2005; Thompson et al. 1995). Surgical intervention is required to relieve the intracranial pressure and allow for brain growth. Even after successful surgery, children with CS can experience ongoing medical problems, such as developmental disabilities (Magge et al. 2002) and vision problems (Gupta et al. 2003). Long-term assessment of neurobehavioral outcomes identified learning disabilities (most often language or visual perception deficits) in 47% of affected school-aged children (Kapp-Simon 1998) compared to 10% of unaffected children in the general population (Altarac and Saroha 2007).

Approximately 80% of CS cases are nonsyndromic (NCS) (Cohen and MacLean 2000), where the premature suture fusion is the only major defect. Two common NCS subtypes are sagittal NCS (sNCS) and metopic NCS (mNCS), which affect the midline skull sutures. Estimates for sNCS suggest it occurs in 1.9 - 2.3 per 10,000 live births (Hunter and Rudd 1976; Lajeunie et al. 1996) with a 3:1 male to female ratio (Cohen and MacLean 2000). About 2% of sNCS cases are thought to be familial with an increased recurrence risk of 1% for siblings of affected individuals (Lajeunie et al. 1996). Our previous GWAS for sNCS, consisting of 130 non-Hispanic white (NHW) case-parent triads with sNCS, identified robust associations to loci near *BMP2* (rs1884302; $P=1.1x10^{-39}$; OR=4.38) and within *BBS9* (rs10262453; $P=5.6x10^{-20}$; OR=0.24) (Justice et al. 2012), which were genes not previously reported in CS patients.

Metopic CS, manifesting as trigonocephaly, occurs in about 1 in 15,000 live births (Cohen and MacLean 2000), with most (75%) cases presenting as nonsyndromic (without developmental

delays and/or additional unrelated major structural defects) (Cohen and MacLean 2000; Greenwood et al. 2014). mNCS shows a three-fold excess among males (Lajeunie et al. 1995; Slater et al. 2008), with a family history of metopic synostosis occurring in about 10-15% of mNCS cases (Jehee et al. 2005; Lajeunie et al. 1995). Additional evidence that genetic factors contribute to the etiology of mNCS comes from the difference between concordance ratios (43% vs. 5%) for monozygotic versus dizygotic twins and the increased incidence (6.4%) for CS among first-degree relatives of probands with mNCS (Greenwood et al. 2014; Lajeunie et al. 2005).

Following up on GWAS for sNCS, we performed the first GWAS for mNCS. Case-parent triads were obtained from the International Craniosynostosis Consortium (ICC; https://health.ucdavis.edu/pediatrics/boyd-genetics-lab/icc.html and National Birth Defects Prevention Study (NBDPS) (Reefhuis et al. 2015; Yoon et al. 2001). Using these samples, we investigated genetic variation associated with mNCS. In addition, we conducted a meta-analysis of our mNCS and sNCS GWAS data to identify associated variants common to both types of midline NCS.

Materials and methods

Subjects

Discovery sample

Our discovery sample was comprised of 410 families, of which 262 were ICC case-parent triads, 13 were ICC multiplex families, and 135 were NBDPS case-parent triads. The enrollment criterion for the study was mNCS in the absence of other unrelated birth defects and/or developmental delays. The presence of mNCS was validated by computerized tomography of the skull or surgical reports. Participants provided whole blood or oral (buccal) specimens. Data collection was approved by the Institutional Review Boards (IRBs) of the University of California, Davis and all

participating institutions in accordance with institutional guidelines. Signed informed consent was obtained for all participants included in the discovery sample. Experimental and data analysis protocols are available upon request from the authors; data has been deposited in database of Genotypes and Phenotypes (dbGaP).

Replication sample

We selected an independent non-Hispanic white (NHW) sample of 285 unrelated cases and 855 unaffected controls for replication. Case and control specimens either were mother-child dyads recruited from the ICC or NBDPS or from anonymized residual newborn blood spots provided by the New York State Department of Health where the proband had a diagnosis of mNCS confirmed by the New York State Congenital Malformations Registry. Signed informed consent was obtained for ICC or NBDPS dyads, and IRB approval was obtained from the New York State Department of Health for use of anonymous blood spots.

Genotyping

Discovery sample

We extracted genomic DNA from 688 whole blood and 320 oral specimens and from whole genome amplified DNA from 13 blood and four oral specimens. We performed targeted mutation analysis for *FGFR1*, *FGFR2*, *FGFR3*, and *TWIST1* to exclude mild or atypical syndromic cases (Lattanzi, et al. 2017). The Center for Inherited Disease Research (CIDR) genotyped the specimens using the Infinium Multi-Ethnic Genotyping Array plus DrugDev (MEGA) Array containing 1,881,804 single nucleotide polymorphisms (SNPs), which included 166,523 pharmacogenetic SNPs to assess exposure to medications as a risk factor for mNCS. Genotypes for 1,881,804 SNPs were released for 1,050 specimens, which included 23 blind duplicates and 19 HapMap controls (10 Utah residents of European ancestry, four Han Chinese, two Japanese, three Yorubans). The

missing rate was 0.24%, blind duplicate reproducibility rate was 99.99%, and HapMap concordance rate was 99.70%. We used Plink v1.90b5.2 (Purcell et al. 2007) to detect discrepancies between expected and annotate sex; five specimens annotated as 'unknown" were reclassified to reflect the genetically-inferred sex. We also used Plink v1.90b5.2 (Purcell et al. 2007) to conduct pairwise identity-by-descent analyses. Three contaminated specimens, three identified monozygotic twins, 16 siblings, two case-parent triads with inconsistent parent-child relationships, and 22 cases with additional suture involvement were excluded prior to analysis.

Replication sample

Our replication sample was genotyped using a panel of 120 SNPs. The SNPs selected were: 1) significant at a suggestive genome-wide level in our GWAS ($P < 1 \times 10^{-5}$); 2) part of the Infinium DrugDev Array and had a $P < 1 \times 10^{-4}$ because these SNPs had lower MAFs than most other SNPs in the array; 3) associated with sNCS in our previous GWAS but not with mNCS (at $P < 1 \times 10^{-5}$); or 4) in high linkage disequilibrium (LD) to associated SNPs. Additionally, the replication SNP panel included 48 Ancestry Informative Markers (AIMs) (Kosoy et al. 2009). CIDR performed genotyping of the replication sample using TaqMan OpenArray, and genotypes for 115 SNPs (5 SNPs failed genotyping) for 920 specimens were released. These specimens included 18 blind duplicates (blind duplicate reproducibility rate of 99.4%), 24 HapMap controls (concordance rate of 99.5%), and 15 replicate specimens, comprised of five case-parent triads from the discovery sample. SNPs in the replication panel were dropped if they had a call rate < 90%, were discordant in one or more duplicates, or were discordant in any replicate when compared to the GWAS genotypes. Sixty-six SNPs remained for replication analysis, of which 22 were AIMs.

Data analysis

Discovery sample

We performed principal components analysis (PCA) using the 'prcomp' package in R v3.2.5 on 58,725 SNPs with a minor allele frequency (MAF) > 15% genotyped in all specimens, including the 19 HapMap specimens; SNPs were pruned such that the maximum r^2 within a 50kb sliding window was 0.2. Families with both founders within one standard deviation of PC1 and PC2 of the mean of the 10 HapMap CEU (Utah residents of European ancestry) specimens were retained for analysis. To minimize loss of power, families in which one founder and their affected offspring were within the desired PCA region and the other parent was close to the HapMap CEU cluster were included.

We conducted a genome-wide association test using the allelic transmission disequilibrium test (TDT) as implemented in Plink v1.90b5.2 (Purcell et al. 2007) which analyzed autosomal and chromosome X markers. The TDT was performed on 215 NHW case-parent triads on 650,848 SNPs that met stringent quality-control procedures and had a MAF > 5%. SNPs were dropped if they had a call rate less than 98%, had a Mendelian inconsistency in more than one case-parent triad, were discordant in one or more duplicates, had heterozygous haploid genotypes present, were monomorphic and/or showed deviation from Hardy-Weinberg equilibrium $P \leq 1x10^{-6}$ (calculated using only the unrelated individuals of European ancestry), or had a MAF difference > 0.20 between males and females on the X chromosome. All SNP map position information was based on Human February 2009 GRCh37/hg19 Assembly. Manhattan and qq-plots were generated using 'qqman" package R v3.23.0 and LD plots were obtained using Haploview v4.234.

The TDT under additive, dominant, and recessive modes of inheritance was performed using the 'trio' R package v3.23.0 (Schwender et al. 2014). A search for two-way gene-gene interactions was carried out using the 'trio' R package v3.23.0 (Schwender et al. 2014) using Cordell's method (Cordell 2002) on 822 selected SNPs. A genotypic TDT for the possible

interaction between each pair of SNPs (581,322 tests) was performed, excluding results from SNP pairs on the same chromosome due to long-range LD. As recommended in Cordell's method (Cordell 2002), for each case-parent triad, 15 pseudo controls matched to each case were generated as a function of the parental genotypes of the SNP pair tested. Pseudo controls for each case were comprised of one of the possible two-locus genotypes that was not transmitted to the case. Using these 15 pseudo controls and matched cases, a conditional logistic regression model was fitted to test for epistatic interactions that incorporate additive effects at the two loci.

We performed pathway analysis using the SNPs associated with mNCS at $P < 1 \times 10^{-5}$, for which autosomal SNPs were annotated to genes using wANNOVAR (Chang and Wang 2012); X-linked SNPs were annotated to genes using UCSC Genome Browser (GRCh37/hg19, www.genome.ucsc.edu). Using Ingenuity Pathway Analysis (IPA, QIAGEN, Redwood City, www.qiagen.com/ingenuity), we implemented canonical pathway analysis to identify the most significant pathways from the IPA library of canonical pathways in the annotated genes, and a Benjamini-Hochberg multiple comparison adjustment was applied (Benjamini and Hochberg 1995). In order to corroborate our results, we used a second pathway analysis program iGSEAvGWASv1.1 (http://gsea4gwas.psych.ac.cn/docs/documents.jsp).

A pre-phasing approach was used to impute unobserved SNPs. Genomic strand information was used to identify and flip the strand of SNPs where the TOP (indicates the A or T allele containing strand) alleles were not aligned to the plus ("+") strand of the human genome reference assembly. Data were phased using SHAPEIT2 (Delaneau et al. 2013), inputting the filtered, chromosome-specific Plink files and receiving the best guess haplotypes as output. These best guess haplotypes were fed directly into the minimac3 software (Das et al. 2016) on the University of Michigan Imputation server, v.1.0.3. The Haplotype Reference Consortium (HRC)

reference panel (McCarthy et al. 2016), which contains 64,976 haplotypes and 39,235,157 sites, was used to filter SNPs with a minor allele count of at least five. The imputation of the pseudo-autosomal regions on chromosomes X and Y, PAR1, and PAR2 could not be carried out on the imputation server; instead, the PAR regions of the X chromosome regions were fed directly into the IMPUTE2 imputation software (Howie et al. 2011) using the 1000 Genomes Project Phase 3 reference panel. Imputed variants were filtered by imputation quality score Rsq > 0.3.

Replication sample

To analyze our replication sample, we performed PCA using Plink v1.90b5.2 (Purcell et al. 2007) in two stages. The first stage used all 214 cases and 560 controls (including the six HapMap specimens) and identified 201 cases and 533 controls (including three HapMap specimens) within 2.5 standard deviations of the mean of PC1 and PC2. To ensure the selected set of 201 cases and 530 controls (excluding the three HapMap specimens) was homogeneous, PCA was performed again with the 22 AIMs remaining after quality control procedures to identify any remaining structure. Because some population structure was still present, cases and controls within 2.5 standard deviations of the mean of PC1 and PC2 were again selected, yielding 194 cases and 502 controls for analysis.

In order to match the 3:1 male:female ratio observed in the 194 cases, a random selection of female controls that would match this ratio was performed ten times. No significant changes were detected when running association analyses on these ten sets of female controls; thus, one of these selected groups of female controls was randomly selected for the final analysis. This exclusion reduced the number of controls analyzed to 333. For the replication analysis, we applied a χ^2 test for allelic association using Plink v1.90b5.2 (Purcell et al. 2007). This test does not adjust for covariates, such as sex; thus, the sample with a reduced number of female controls was used. Our

association results (*P* value) for the 10 sets of samples for which females were randomly dropped to match the 3:1 ratio found in the cases are presented in Table S1. This approach was selected rather than applying logistic regression analysis, which could adjust for sex, because the meta-analysis method that was used applied a χ^2 test (Kazeem and Farrall 2005).

Meta-analyses

Discovery and replication samples

We conducted a meta-analysis combining the results from our discovery TDT and replication case-control study using a fixed-effects model, which tested for interstudy heterogeneity (Kazeem and Farrall 2005). We also conducted a meta-analysis of our previous sNCS and mNCS GWAS data. In the meta-analysis, imputation of the mNCS data to match the genotyped sNCS specimens was considered, but because of discrepancies between imputed and genotyped variants due to differences in allele frequencies between the reference panels and the small sample size available, it was preferable to only use the genotyped variants in both studies. The sNCS GWAS was genotyped using the Illumina 1 M Human Omni1-Quad array (Justice et al. 2012), and LiftOver (Kent et al. 2002) was used to convert all hg18 positions to hg19 positions. Chromosome and base pair (bp) position SNPs from the sNCS GWAS were compared to those on the Illumina MEGA array used in the mNCS GWAS; after removing SNPs with three alleles and those with a MAF < 1%, 306,233 SNPs were retained for the meta-analysis.

Functional analysis

Cell culture

Mesenchymal stem cells (MSCs) and HeLa cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. Culture medium was collected for enzyme-linked immunosorbent assay (ELISA), and MSCs were lysed using Pierce RIPA Buffer (Thermo

Scientific) with protease and phosphatase inhibitors in addition to EDTA and frozen until needed for western blotting.

Real-time quantitative polymerase chain reaction

RNA was isolated from MSCs derived from fused and open sutures using the Zymo RNA Mini-Prep Kit. Real-time quantitative polymerase chain reaction (qPCR) for *BMP7* gene expression analysis was performed using the TaqMan RNA to Ct 1-Step Kit (ThermoFisher). The PCR program began with a 15-minute reverse transcription step at 48º C and a 10-minute activation of the deoxyribonucleic acid (DNA) polymerase at 95º C. This was followed by 40 cycles of denaturation for 15 seconds at 95º C and annealing/extending for one minute at 60º C. TaqMan gene expression assays (ThermoFisher) were used for either *BMP7* (Hs00233476_m1) or the housekeeping gene *GAPDH* (Hs99999905_m1). All qPCR reactions were performed in triplicate, and the amplified signals from *BMP7* were normalized to those obtained from *GAPDH* in the same reactions.

Exon analysis

DNA was extracted from blood, saliva, or mouthwash of 183 mNCS patients according to the manufacturer's protocol using the Gentra Puregene Blood Kit (QIAGEN). The KAPA2G Robust HotStart PCR Kit (Kapa Biosystems) was used to amplify the coding exons of *BMP7* gene using the following PCR primers:

BMP7_exon1_Forward: CGTCTGCAGCAAGTGACC

BMP7_exon1_Reverse: CTGCGATTTCAGCCAGGAG

BMP7_exon2_Forward: GATGCTTGGACTCAGAGCC

BMP7_exon2_Reverse: GTGCCAATCTGACCCATCC

BMP7_exon3_Forward: GATGTTCCCACTTGTCGGG

BMP7_exon3_Reverse: TGAAGTCCAGGAGCACAGG

BMP7_exon4_Forward: AACAGTACCTGGCCTAGAGT

BMP7_exon4_Reverse: GGATTTGGGGGTTTTCTTCC

BMP7_exon5_Forward: CCGTCTGTGCTTCATTGCT

BMP7_exon5_Reverse: AGCGAGGCCACTTGATACT

BMP7_exon6_Forward: TGCTCAGAAGGCATGGTCT

BMP7_exon6_Reverse: ATGACATGGCAATGGGCTG

BMP7_exon7_Forward: TAGAACAGGGAGTGCTTGG

BMP7_exon7_Reverse: AAAGTTGGGGATAGGGAGG

PCR products were purified using ExoSap-IT (Affymetrix) and sequenced by Sanger sequencing. Electropherograms were analyzed with the SnapGene software by two independent investigators.

Western-blot analyses

Equal amounts of protein were heated to 70°C for 10 minutes with NuPAGE LDS Sample Buffer (4×) (Life Technologies), ran on NuPAGE Novex 4–12% Bis-Tris gels (Life Technologies), and transferred to polyvinylidene difluoride membranes (Life Technologies). Membranes were blocked for one hour at room temperature with 5% nonfat milk in Tris Buffered Saline with Tween (TBST; 20 mM Tris-HCl, pH 7.5, 137 mM NaCl, 0.1% Tween 20). Following this, membranes were incubated with primary antibody overnight in blocking solution at 4°C with slight agitation. Primary antibodies for BMP7 (Abcam), phospho-Smad1/5/8 (Cell Signaling), Smad1 (Cell Signaling) and β-tubulin (Cell Signaling) were diluted according to manufacturer's recommendations. The membranes were washed with TBST and incubated with secondary antibody for one hour at room temperature. The secondary antibody (polyclonal goat anti-rabbit immunoglobulins/horseradish peroxidase, ThermoFisher) was diluted according to manufacturer's

recommendations in blocking solution. Membranes were washed with TBST and developed with SuperSignal West Pico PLUS Chemiluminescent Substrate (Thermo Scientific). Experiments were performed in triplicate, and .tiff images were analyzed using ImageJ (https://imagej.nih.gov/ij/). Bands were quantified and intensities were normalized with β-tubulin.

Enzyme-linked immunosorbent assay

Levels of secreted BMP7 were measured by an indirect enzyme-linked immunosorbent assay (ELISA). MSCs and HeLa cells were seeded at 1 x 10^6 cells in 100 mm plates and incubated for 24 hours. The media was removed and replaced with serum-free media. The plates were incubated for four hours and the serum-free media was collected. Ninety-six-well microtiter plates were coated with lectin and blocked for two hours at room temperature with 5% nonfat milk in phosphate-buffered saline (PBS). After removing the blocking buffer, the plates were washed with PBS with 0.05% Tween 20. Specimens were added, and the plates were incubated for two hours at room temperature. After washing out unbound substances, a BMP7 monoclonal detective antibody (R&D Systems) was added to the wells and incubated overnight. Again, unbound substances were washed out and the secondary antibody (polyclonal goat anti-mouse, ThermoFisher) was added to the wells. Any unbound antibody-enzyme reagent was washed out and a substrate solution was added to the wells. This caused color development in proportion to the amount of BMP7 present in the specimen. The absorbance of each well was read with an ELISA plate reader (BioTek Synergy HT) at 450 nm. Using a standard curve, the amount of protein secreted by the MSCs and HeLa cells was calculated, and the latter value was used to normalize the data to compare across the different plates.

Dual luciferase assay

We generated 667-bp fragments (chr20:55,796,885-55,797,557, hg19) with the different alleles (G and T) of rs6127972 by PCR using DNA from homozygotes with either allele using primer set: 5'-GAGGGGTGGGCAGGGATAA-3' and 5'-GTTCCGCTTGGGGTCCTC-3'. These fragments were placed in the *SacI*/*MluI* site of pRLuc-promoter vector (SwitchGear Genomics) upstream of the *BMP7* promoter using the In-Fusion HD Cloning Plus CE kit (Takara Bio USA, Inc.) with primer set: 5'-ACTGGCCGGTACCTGGAGGGGTGGGCAGGGATAAG-3' and 5'-GCTTCCTGGAACGCGGTTCCGCTTGGGGTCCTCTC-3'. The resulting constructs (pRLuc-BMP7) contained either the G (common) allele or T (risk) allele. The inserted portions of the resulting constructs were sequenced to verify the nucleic acid sequences and the identity and location of the SNP. For the luciferase assay, 5×10^3 HeLa cells were seeded per well in 96-well plates and transfected with an empty vector or with pRLuc-BMP7 containing the risk (T) or common (G) allele using FuGENE HD reagent (Active Motif). pTK-CLuc construct (SwitchGear Genomics) was co-transfected as a normalizing internal control vector. All transfections were carried out in triplicate. After 24 hours of incubation, luciferase activity was measured using the LightSwitch Dual Assay System (SwitchGear Genomics). Each transfection set included the empty control vector, risk allele, and common allele and was performed independently 12 times. An independent dual-luciferase assay was conducted using randomly selected 667 bp fragments (chr10:45,032,341-45,033,007, hg19) with either the G or T allele of rs1857502. Fragments were placed in the SacI/MluI site of pRLuc-promoter vector upstream of the BMP7 promoter with primer set: 5'- ACTGGCCGGTACCTGCACCATCCAGTCTGTGTC -3' and 5'-GCTTCCTGGAACGCGTTCAATAATCATATCATTGGAGA-3'. Each transfection set included the empty control vector, G allele, and T allele and was performed independently 10 times.

Pyrosequencing assay

Genomic DNA isolated from 24 MSC lines derived from fused metopic and control open sutures from the same proband (12 of each type) was bisulfite converted using the Zymo EZ DNA-Methylation Lightning Kit. PCR and sequencing primers were designed using the PyroMark Assay Design Software 2.0 (Qiagen). Two sets of pyrosequencing PCR and sequencing primers were designed to cover seven CpG sites (chr20:55,798,821-55,798,917, chr20:55,798,975-55,799,053, hg19). These sites each overlap an enhancer predicted to be active in osteoblasts (chromHMM imputed data) and a CTCF transcription factor binding site. Primers were also designed to cover four CpG sites (chr20:55,796,275-55,796,368, hg19) in a nearby region not overlapping the enhancer and CTCF binding site to serve as a control. The specimens were run in triplicate for all assays. The PyroMark PCR Kit (Qiagen) was used to amplify the bisulfite converted DNA, using the following primers:

CpG1-4-PCR-Forward: 5'- AGAAGTTTTAATTATAGGGTGGAGAT-3'

CpG1-4-PCR-Reverse: 5'-CTTACCCAATCCTCTCCTAAAAATAC-3'

CpG5-7-PCR-Forward: 5'- TTAGGAGAGGATTGGGTAAGGA-3'

CpG5-7-PCR-Reverse: 5'- ACCTTTCTAAAAAACTCCCTAA-3'

Control-PCR-Forward: 5'-ATTGTTTTTGTTGGGTTTTATTTAGAT-3'

Control-PCR-Reverse: 5'- AACCAATAACCTACCCAACCTATC-3'

The specimens were sequenced using the PyroMark Biotage Q96 (Qiagen). The following sequencing primers were used:

CpG1-4-Sequencing: 5'- TTTTAATTATAGGGTGGAGATA-3

CpG5-7-Sequencing: 5'- GGATTGGGTAAGGAG – 3'

Control-Sequencing: 5'- GTTTAGTAAGTGTTTATATGTTG-3'

Results

Six SNPs exceeded the genome-wide significance threshold of $P \leq 5x10^{-8}$ (Figure 1, Figure 2, Figure S1): rs781716 intronic to *SPRY3* ($P = 4.71 \times 10^{-8}$), rs6127972 intronic to *BMP7* ($P = 4.411 \times 10^{-8}$), rs62590971 located 155 kb upstream from *TGIF2LX* ($P = 6.22 \times 10^{-9}$), and three SNPs intronic to *PCDH11X* (rs2522623, $P = 1.76 \times 10^{-8}$; rs2573826, $P = 3.31 \times 10^{-8}$; rs2754857, $P = 1.09 \times 10^{-8}$). Because Plink TDT analysis assumes a multiplicative mode of inheritance by performing an allelic TDT analysis, we also conducted a genotypic TDT analysis using the 'trio' R package v3.23.0 (Schwender et al. 2014) to test for additive, dominant, and recessive modes of inheritance. Our results indicated the effect of the significant SNPs was consistent with an additive model (Table S2).

We tested for two-way gene-gene interactions using 822 selected SNPs. Of these, 808 SNPs were located in CS candidate genes (Lattanzi et al. 2017), 13 SNPs had a $P < 1 \times 10^{-5}$ for the TDT GWAS and were not located in CS candidate gene regions, and one SNP, rs1884302 (345 kb downstream of *BMP2*) identified in our previous GWAS for sagittal NCS (sNCS) (Justice et al. 2012), was not significant in our current study (Table S3). No significant interaction effect was observed after correcting for multiple testing (581,322 tests, Bonferroni $P < 9 \times 10^{-8}$), with the most significant association ($P = 1.16 \times 10^{-5}$) observed for an interaction between rs876688 (intronic to *TGFBR2*) and rs4637716 (intronic to *BBS9*). Pathway analysis conducted by selecting SNPs associated with mNCS in the 215 NHW case-parent triads at $P < 1 \times 10^{-5}$ did not show any pathways to be enriched at a Benjamini-Hochberg false discovery rate of 0.05.

We extracted the imputed regions (± 100 kb) on each chromosome (2, 5, 11, 20) where there were suggestive genome-wide significant associations ($P < 1 \times 10^{-5}$) and the PAR2 pseudo-autosomal region on the X chromosome. Genome-wide significant associations for SNPs were

identified only on chromosome 20 intronic to *BMP7*. The most significant association was for rs162319 ($P = 2.86 \times 10^{-8}$, Figure S2), which was comparable to the genotyped rs6127972 ($P = 4.41 \times 10^{-8}$).

Our replication analyses included markers on the X chromosome, which are susceptible to type 1 errors when allelic frequencies differ between males and females in an unbalanced sample (different number of females and males in cases vs. controls) (Loley et al. 2011). As such, a random sample of females was selected and dropped from the control sample to match the 3:1 male:female ratio observed in the 194 cases and reduced the number of controls analyzed to 333. The only SNP which was genome-wide significant in our discovery sample and replicated in our case-control sample was rs6127972 ($P = 0.004$, OR = 1.45).

In our meta-analysis combining the results from our discovery TDT and the replication case-control study, the only genotyped SNP to reach genome-wide significance was rs6127972 ($P = 1.27 \times 10^{-8}$). SNPs identified at a suggestive genome-wide significance level ($P < 1 \times 10^{-5}$) included rs10254116, rs10262453, and rs4723276 in *BBS9*; rs34360385 ~ 282kb and ~627kb from *PABC5* and *TGI2LX,* respectively; and rs230217, rs6014954, rs230218, and rs17404303, in close proximity to rs6127972 (Table S4), with the imputed SNPs rs4723276, rs230217, rs6014954, and rs230218 reaching genome-wide significance ($P < 5 \times 10^{-8}$).

Combining data for genotyped SNPs (MAF > 1%) common to our GWAS for mNCS and our previous GWAS for sNCS, we performed a meta-analysis with the 215 NHW mNCS and 130 NHW sNCS case-parent triads (Justice et al. 2012) using METAL (Willer et al. 2010). No SNP reached genome-wide significance, with the top association identified for rs6127972 ($P = 1.16 \times 10^{-6}$). For SNPs rs1884302 and rs10262453, the most significant associations found in our previous sNCS GWAS, only rs10262453 reached a significance level of $P < 1 \times 10^{-5}$ in both our mNCS

GWAS and sNCS GWAS. Opposite directions of effect were observed for rs10262453, with the A allele being over-transmitted for sNCS, but the C allele being over-transmitted for mNCS (Table S5). This finding may indicate a functional role for this variant.

We sequenced all seven exons of *BMP7* in 183 NHW mNCS cases, of which 118 were included in our discovery sample. No variation from the reference coding sequence of the gene was observed. Because no deleterious *BMP7* mutations (loss of function or missense variants) identified here or by a previous study (Timberlake et al. 2017), we explored if *BMP7* is abnormally expressed in mNCS cases due to a regulatory effect of rs6127972 or a nearby noncoding regulatory element. To test this hypothesis, we measured *BMP7* expression and BMP7 protein secretion in MSCs isolated from both open and fused suture tissue from the same mNCS probands as previously described (Lattanzi et al. 2013). Using qPCR, expression levels of *BMP7* were below the reference threshold cycle (Livak and Schmittgen 2001). BMP7 protein levels were detected by Western blot and quantified by ELISA on 16 pairs of MSCs derived from fused metopic and open sutures obtained from the same mNCS probands. We observed no significant difference in secreted BMP7 protein levels when comparing within each pair, as well as between fused and open sutures. When considering the rs6127972 genotype, we did not find a significant difference in protein levels. In summary, no significant difference in *BMP7* expression or secretion was observed between T (over-transmitted allele) or G (common allele) at the rs6127972 locus.

We also performed luciferase assays to assess the regulatory activity of rs6127972, the non-imputed SNP observed to have the strongest association with mNCS. To conduct the assays, we generated 667-bp fragments (chr20:55,796,885-55,797,557, hg19) with either the G allele or T allele of rs6127972 and containing a nearby DNaseI hypersensitivity cluster (chr20:55,797,206-55,797,555, hg19). These fragments were cloned 5′ upstream of a *BMP7* promoter reporter

construct (Figure 3). The expression of the luciferase reporter for both fragments with either with the risk (T) or common (G) allele at rs6127972 was significantly lower compared to the empty control vector (Figure 3). Although the trend for expression levels was consistent with the T allele (over- transmitted in our mNCS sample) having the lowest expression, the two allele fragments did not differ significantly in their modulation of promoter activity ($P = 0.05$). After adjustment for multiple tests, both fragments compared to the empty control were significant (G allele $P = 3.78 \times 10^{-5}$; T allele $P = 5.11 \times 10^{-6}$), suggesting the region around rs6127972 may act as a repressor element. The control luciferase experiment with the randomly selected fragments of 667 bp flanking rs1857502 did no change the expression of the luciferase reporter. This corroborates our observation of inhibitory effect of rs6127972 as locus specific (Figure 3b).

Our imputation analysis identified a linkage disequilibrium block (chr20:55,790,147-55,807,110, hg19) encompassing rs6127972 and rs162319 (the most significant association to mNCS) containing several regulatory elements. These regulatory elements included several annotated enhancers. One of these enhancers overlaps a CTCF transcription factor binding site, and we found this region (chr20:55,798,821-55,798,917) to be significantly hypomethylated in MSCs derived from fused metopic sutures compared to those from open sutures from the same proband, suggesting stronger binding at this site (Figure 4). Further investigation of the other SNPs in the linkage disequilibrium block and regulatory elements in the region may be warranted.

Discussion

Our GWAS of 215 NHW case-parent triads identified several loci associated with mNCS at a genome-wide significance level, but only the association with rs6127972 was replicated in an independent case-control sample. Of the SNPs that did not replicate, rs781716 is intronic to

SPRY3, which is a modulator of FGF signaling (Panagiotaki et al. 2010) and has been suggested as a candidate for autism (Ning et al. 2015); rs62590971 is ~155 kb upstream from *TGIF2LX*, which plays a role in spermatogenesis (Wang et al. 2008); and the remaining three SNPs (rs2522623, rs2573826, rs2754857) are all located intronic to *PCDH11X*, which has been suggested to play a role in cognition related disorders (Veerappa et al. 2013). Whole exome sequencing (WES) of 136 cases with mNCS and 237 cases with sNCS did not report rare variants in *SPRY3*, *TGIF2LX,* or *PCDH11X* (Timberlake et al. 2017), and a WES of 191 cases with either mNCS or sNCS did not report any mutations in *SPRY3* (Timberlake et al. 2016).

Our discovery sample of 215 NHW case-parent triads had a 5:1 ratio of male to female cases, and when we restricted analysis of X chromosome SNPs to those with allele frequency differences < 0.15 between males and females (quality control threshold for analysis was < 0.2), only the association with rs62590971 was significant (allele frequency difference between males and females equal to 0.099). Differences in allelic frequency, especially in unbalanced samples like ours, can lead to apparent over-transmission of one or the other allele to the affected offspring (Loley et al. 2011). Lack of any strong association signal in our replication sample to these pseudo-autosomal regions suggests that the associations in our discovery sample might have been a result of allele frequency differences between males and females. Based on this finding and available data from the literature, *SPRY3*, *TGIF2LX*, and *PCDH11X* have not been currently identified as genetic factors involved in the development of mNCS.

The one SNP, rs6127972, replicated in our mNCS GWAS is intronic to *BMP7*, a member of the BMP superfamily. This superfamily is involved in skeletal development (Beederman et al. 2013; Salazar et al. 2016) by inducing bone formation (Asahina et al. 1996; Fujii et al. 1999) and may have therapeutic potential for orthopedic applications (Salazar et al. 2016). Developmental

abnormalities in the skull were identified in *Bmp7*-deficient mutant mice, including fusion of the basisphenoid with the occipital bone.(Luo et al. 1995). Enhanced BMP signaling through BMP type IA receptor in neural crest cells produces CS in mice (Komatsu et al. 2013). Rapamycin rescues BMP-mediated midline CS through the inhibition of mTOR signaling in mice (Kramer et al. 2018). Importantly, our sNCS GWAS documented a strong and reproducible association with rs1884302 located 345 kb downstream of *BMP2* (Justice et al. 2012), and our functional analysis demonstrated that this locus may act as an enhancer element (Justice et al. 2017). Loss of function mutations were identified in *SMAD6*, a negative regulator of the BMP signaling pathway, in cases with midline CS, and a two-locus inheritance with synergistic effect of *SMAD6* loss of function mutations with the C allele of rs1884302 was suggested (Timberlake et al. 2016). Our findings for mNCS further implicate the BMP signaling pathway in the development of CS.

In summary, our GWAS identified a variant in *BMP7*, rs6127972, which was significantly and reproducibly associated with mNCS. Results from our luciferase assays suggest that the intronic *BMP7* region containing this SNP could act as a repressor element, as both the risk and common allele reduced expression of the *BMP7* promoter that was not observed in our control luciferase assay. Even with these results, it is possible that the 667 bp fragment may regulate another gene in the region, including an uncharacterized gene ~6kb away LOC102723590; however, given our phenotype (premature ossification of a suture), we considered BMP7 as a candidate due to its role in bone development. Previous work has shown *BMP7* to be involved in skeletal development (Beederman et al. 2013; Salazar et al. 2016), and along with the findings from our study and previous reports (Justice et al. 2012; Kramer et al. 2018; Whitton et al. 2016), suggests some role for BMP signaling in CS. Additionally, our meta-analysis of GWAS data for mNCS and sNCS identified rs10262453 (intronic to *BBS9*) appears to be involved with midline

NCS. Future efforts to examine whether the allele at rs10262453, or the regulatory region surrounding it, interacts with *BMP2* and/or *BMP7*, or acts independently by regulating another gene involved in suture formation are warranted.

Financial disclosure statement: On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgements The authors thank all families who contributed to this study. This project was supported by the: National Institute of Dental and Craniofacial Research (NIDCR)/National Institutes of Health (NIH) R01 DE016886 (S.A.B., P.A.R., A.C., K.B.); NIDCR/NIH DE018277 (M.L.C.); Centers for Disease Control and Prevention cooperative agreements under PA #96043, PA #02081, FOA #DD09-001, FOA #DD13-003, and NOFO #DD18-001 to the Centers for Birth Defects Research and Prevention participating in the National Birth Defects Prevention Study and/or the Birth Defects Study To Evaluate Pregnancy exposureS and the Iowa Center for Birth Defects Research and Prevention U01 DD001035 and U01 DD001223 (P.A.R.); Wellcome Investigator Award 102731 (A.O.M.W.); the Università Cattolica del Sacro Cuore (W.L.); and in part, by the Division of Intramural Research Program of the National Human Genome Research Institute, NIH (C.M.J., J.A.S., A.F.W.). Additional support was provided by the Intramural Research Program, National Institute of Child Health and Human Development, NIH: HHSN275201100001I, HHSN27500005. We thank Samantha Edwards, Gill Roberts, Danielle Stevenson, Vivienne Sutton, Elizabeth Tidey, Paolo Frassanito, Luca Massimi and Massimo Caldarelli for help with specimen collection. Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the NIH

to The Johns Hopkins University, contract number HHSN268200782096C. The findings and conclusions in the report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

References

Altarac M, Saroha E (2007) Lifetime prevalence of learning disability among US children. Pediatrics 119 Suppl 1: S77-83. doi: 10.1542/peds.2006-2089L

Asahina I, Sampath TK, Hauschka PV (1996) Human osteogenic protein-1 induces chondroblastic, osteoblastic, and/or adipocytic differentiation of clonal murine target cells. Exp Cell Res 222: 38-47. doi: 10.1006/excr.1996.0005

Beederman M, Lamplot JD, Nan G, Wang J, Liu X, Yin L, Li R, Shui W, Zhang H, Kim SH, Zhang W, Zhang J, Kong Y, Denduluri S, Rogers MR, Pratt A, Haydon RC, Luu HH, Angeles J, Shi LL, He TC (2013) BMP signaling in mesenchymal stem cell differentiation and bone formation. J Biomed Sci Eng 6: 32-52. doi: 10.4236/jbise.2013.68A1004

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B 57: 289-300.

Chang X, Wang K (2012) wANNOVAR: annotating genetic variants for personal genomes via the web. J Med Genet 49: 433-6. doi: 10.1136/jmedgenet-2012-100918

Cohen MM, MacLean RE (2000) Craniosynostosis : diagnosis, evaluation, and management, 2nd edn. Oxford University Press, New York

Cordell HJ (2002) Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. Hum Mol Genet 11: 2463-8.

Das S, Forer L, Schonherr S, Sidore C, Locke AE, Kwong A, Vrieze SI, Chew EY, Levy S, McGue M, Schlessinger D, Stambolian D, Loh PR, Iacono WG, Swaroop A, Scott LJ, Cucca F, Kronenberg F, Boehnke M, Abecasis GR, Fuchsberger C (2016) Next-generation genotype imputation service and methods. Nat Genet 48: 1284-1287. doi: 10.1038/ng.3656

Delaneau O, Zagury JF, Marchini J (2013) Improved whole-chromosome phasing for disease and population genetic studies. Nat Methods 10: 5-6. doi: 10.1038/nmeth.2307

Fujii M, Takeda K, Imamura T, Aoki H, Sampath TK, Enomoto S, Kawabata M, Kato M, Ichijo H, Miyazono K (1999) Roles of bone morphogenetic protein type I receptors and Smad proteins in osteoblast and chondroblast differentiation. Mol Biol Cell 10: 3801-13. doi: 10.1091/mbc.10.11.3801

Greenwood J, Flodman P, Osann K, Boyadjiev SA, Kimonis V (2014) Familial incidence and associated symptoms in a population of individuals with nonsyndromic craniosynostosis. Genet Med 16: 302-10. doi: 10.1038/gim.2013.134

Gupta PC, Foster J, Crowe S, Papay FA, Luciano M, Traboulsi EI (2003) Ophthalmologic findings in patients with nonsyndromic plagiocephaly. J Craniofac Surg 14: 529-32. doi: 10.1097/00001665-200307000-00026

Howie B, Marchini J, Stephens M (2011) Genotype imputation with thousands of genomes. G3 (Bethesda) 1: 457-70. doi: 10.1534/g3.111.001198

Hunter AG, Rudd NL (1976) Craniosynostosis. I. Sagittal synostosis: its genetics and associated clinical findings in 214 patients who lacked involvement of the coronal suture(s). Teratology 14: 185-93. doi: 10.1002/tera.1420140209

Jehee FS, Johnson D, Alonso LG, Cavalcanti DP, de Sa Moreira E, Alberto FL, Kok F, Kim C, Wall SA, Jabs EW, Boyadjiev SA, Wilkie AO, Passos-Bueno MR (2005) Molecular screening for microdeletions at 9p22-p24 and 11q23-q24 in a large cohort of patients with trigonocephaly. Clin Genet 67: 503-10. doi: 10.1111/j.1399-0004.2005.00438.x

Justice CM, Kim J, Kim SD, Kim K, Yagnik G, Cuellar A, Carrington B, Lu CL, Sood R, Boyadjiev SA, Wilson AF (2017) A variant associated with sagittal nonsyndromic craniosynostosis alters the regulatory function of a non-coding element. Am J Med Genet A 173: 2893-2897. doi: 10.1002/ajmg.a.38392

Justice CM, Yagnik G, Kim Y, Peter I, Jabs EW, Erazo M, Ye X, Ainehsazan E, Shi L, Cunningham ML, Kimonis V, Roscioli T, Wall SA, Wilkie AO, Stoler J, Richtsmeier JT, Heuze Y, Sanchez-Lara PA, Buckley MF, Druschel CM, Mills JL, Caggana M, Romitti PA, Kay DM, Senders C, Taub PJ, Klein OD, Boggan J, Zwienenberg-Lee M, Naydenov C, Kim J, Wilson AF, Boyadjiev SA (2012) A genome-wide association study identifies susceptibility loci for nonsyndromic sagittal craniosynostosis near BMP2 and within BBS9. Nat Genet 44: 1360-4. doi: 10.1038/ng.2463

Kapp-Simon KA (1998) Mental development and learning disorders in children with single suture craniosynostosis. Cleft Palate Craniofac J 35: 197-203. doi: 10.1597/1545-1569_1998_035_0197_mdaldi_2.3.co_2

Kazeem GR, Farrall M (2005) Integrating case-control and TDT studies. Ann Hum Genet 69: 329-35. doi: 10.1046/j.1529-8817.2005.00156.x

Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, Haussler D (2002) The human genome browser at UCSC. Genome Res 12: 996-1006. doi: 10.1101/gr.229102

Komatsu Y, Yu PB, Kamiya N, Pan H, Fukuda T, Scott GJ, Ray MK, Yamamura K, Mishina Y (2013) Augmentation of Smad-dependent BMP signaling in neural crest cells causes craniosynostosis in mice. J Bone Miner Res 28: 1422-33. doi: 10.1002/jbmr.1857

Kosoy R, Nassir R, Tian C, White PA, Butler LM, Silva G, Kittles R, Alarcon-Riquelme ME, Gregersen PK, Belmont JW, De La Vega FM, Seldin MF (2009) Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America. Hum Mutat 30: 69-78. doi: 10.1002/humu.20822

Kramer K, Yang J, Swanson WB, Hayano S, Toda M, Pan H, Kim JK, Krebsbach PH, Mishina Y (2018) Rapamycin rescues BMP mediated midline craniosynostosis phenotype through reduction of mTOR signaling in a mouse model. Genesis 56: e23220. doi: 10.1002/dvg.23220

Lajeunie E, Crimmins DW, Arnaud E, Renier D (2005) Genetic considerations in nonsyndromic midline craniosynostoses: a study of twins and their families. J Neurosurg 103: 353-6. doi: 10.3171/ped.2005.103.4.0353

Lajeunie E, Le Merrer M, Bonaiti-Pellie C, Marchac D, Renier D (1995) Genetic study of nonsyndromic coronal craniosynostosis. Am J Med Genet 55: 500-4. doi: 10.1002/ajmg.1320550422

Lajeunie E, Le Merrer M, Bonaiti-Pellie C, Marchac D, Renier D (1996) Genetic study of scaphocephaly. Am J Med Genet 62: 282-5. doi: 10.1002/(SICI)1096-8628(19960329)62:3<282::AID-AJMG15>3.0.CO;2-G

Lattanzi W, Barba M, Novegno F, Massimi L, Tesori V, Tamburrini G, Galgano S, Bernardini C, Caldarelli M, Michetti F, Di Rocco C (2013) Lim mineralization protein is involved in the premature calvarial ossification in sporadic craniosynostoses. Bone 52: 474-84. doi: 10.1016/j.bone.2012.09.004

Lattanzi, W., Barba, M., Di Pietro, L. & Boyadjiev, S. A. Genetic advances in craniosynostosis. *Am J Med Genet A* **173**, 1406-1429, doi:10.1002/ajmg.a.38159 (2017).

Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-8. doi: 10.1006/meth.2001.1262

Loley C, Ziegler A, Konig IR (2011) Association tests for X-chromosomal markers--a comparison of different test statistics. Hum Hered 71: 23-36. doi: 10.1159/000323768

Luo G, Hofmann C, Bronckers AL, Sohocki M, Bradley A, Karsenty G (1995) BMP-7 is an inducer of nephrogenesis, and is also required for eye development and skeletal patterning. Genes Dev 9: 2808-20.

Magge SN, Westerveld M, Pruzinsky T, Persing JA (2002) Long-term neuropsychological effects of sagittal craniosynostosis on child development. J Craniofac Surg 13: 99-104. doi: 10.1097/00001665-200201000-00023

McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, Kang HM, Fuchsberger C, Danecek P, Sharp K, Luo Y, Sidore C, Kwong A, Timpson N, Koskinen S, Vrieze S, Scott LJ, Zhang H, Mahajan A, Veldink J, Peters U, Pato C, van Duijn CM, Gillies CE, Gandin I, Mezzavilla M, Gilly A, Cocca M, Traglia M, Angius A, Barrett JC, Boomsma D, Branham K, Breen G, Brummett CM, Busonero F, Campbell H, Chan A, Chen S, Chew E, Collins FS, Corbin LJ, Smith GD, Dedoussis G, Dorr M, Farmaki AE, Ferrucci L, Forer L, Fraser RM, Gabriel S, Levy S, Groop L, Harrison T, Hattersley A, Holmen OL, Hveem K, Kretzler M, Lee JC, McGue M, Meitinger T, Melzer D, Min JL, Mohlke KL, Vincent JB, Nauck M, Nickerson D, Palotie A, Pato M, Pirastu N, McInnis M, Richards JB, Sala C, Salomaa V, Schlessinger D, Schoenherr S, Slagboom PE, Small K, Spector T, Stambolian D, Tuke M, Tuomilehto J, Van den Berg LH, Van Rheenen W, Volker U, Wijmenga C, Toniolo D, Zeggini E, Gasparini P, Sampson MG, Wilson JF, Frayling T, de Bakker PI, Swertz MA, McCarroll S, Kooperberg C, Dekker A, Altshuler D, Willer C, Iacono W, Ripatti S, et al. (2016) A reference panel of 64,976 haplotypes for genotype imputation. Nat Genet 48: 1279-83. doi: 10.1038/ng.3643

Ning Z, McLellan AS, Ball M, Wynne F, O'Neill C, Mills W, Quinn JP, Kleinjan DA, Anney RJ, Carmody RJ, O'Keeffe G, Moore T (2015) Regulation of SPRY3 by X chromosome and PAR2-linked promoters in an autism susceptibility region. Hum Mol Genet 24: 5126-41. doi: 10.1093/hmg/ddv231

Panagiotaki N, Dajas-Bailador F, Amaya E, Papalopulu N, Dorey K (2010) Characterisation of a new regulator of BDNF signalling, Sprouty3, involved in axonal morphogenesis in vivo. Development 137: 4005-15. doi: 10.1242/dev.053173

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, Sham PC (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 81: 559-75. doi: 10.1086/519795

Reefhuis J, Gilboa SM, Anderka M, Browne ML, Feldkamp ML, Hobbs CA, Jenkins MM, Langlois PH, Newsome KB, Olshan AF, Romitti PA, Shapira SK, Shaw GM, Tinker SC, Honein MA, National Birth Defects Prevention S (2015) The National Birth Defects Prevention Study: A review of the methods. Birth Defects Res A Clin Mol Teratol 103: 656-69. doi: 10.1002/bdra.23384

Salazar VS, Gamer LW, Rosen V (2016) BMP signalling in skeletal development, disease and repair. Nat Rev Endocrinol 12: 203-21. doi: 10.1038/nrendo.2016.12

Schwender H, Li Q, Neumann C, Taub MA, Younkin SG, Berger P, Scharpf RB, Beaty TH, Ruczinski I (2014) Detecting disease variants in case-parent trio studies using the bioconductor software package trio. Genet Epidemiol 38: 516-22. doi: 10.1002/gepi.21836

Slater BJ, Lenton KA, Kwan MD, Gupta DM, Wan DC, Longaker MT (2008) Cranial sutures: a brief review. Plast Reconstr Surg 121: 170e-8e. doi: 10.1097/01.prs.0000304441.99483.97

Tamburrini G, Caldarelli M, Massimi L, Santini P, Di Rocco C (2005) Intracranial pressure monitoring in children with single suture and complex craniosynostosis: a review. Childs Nerv Syst 21: 913-21. doi: 10.1007/s00381-004-1117-x

Thompson DN, Harkness W, Jones B, Gonsalez S, Andar U, Hayward R (1995) Subdural intracranial pressure monitoring in craniosynostosis: its role in surgical management. Childs Nerv Syst 11: 269-75. doi: 10.1007/bf00301758

Timberlake AT, Choi J, Zaidi S, Lu Q, Nelson-Williams C, Brooks ED, Bilguvar K, Tikhonova I, Mane S, Yang JF, Sawh-Martinez R, Persing S, Zellner EG, Loring E, Chuang C, Galm A, Hashim PW, Steinbacher DM, DiLuna ML, Duncan CC, Pelphrey KA, Zhao H, Persing JA, Lifton RP (2016) Two locus inheritance of non-syndromic midline craniosynostosis via rare SMAD6 and common BMP2 alleles. Elife 5. doi: 10.7554/eLife.20125

Timberlake AT, Furey CG, Choi J, Nelson-Williams C, Yale Center for Genome A, Loring E, Galm A, Kahle KT, Steinbacher DM, Larysz D, Persing JA, Lifton RP (2017) De novo mutations in inhibitors of Wnt, BMP, and Ras/ERK signaling pathways in non-syndromic midline craniosynostosis. Proc Natl Acad Sci U S A 114: E7341-E7347. doi: 10.1073/pnas.1709255114

Veerappa AM, Saldanha M, Padakannaya P, Ramachandra NB (2013) Genome-wide copy number scan identifies disruption of PCDH11X in developmental dyslexia. Am J Med Genet B Neuropsychiatr Genet 162B: 889-97. doi: 10.1002/ajmg.b.32199

Wang Y, Wang L, Wang Z (2008) Transgenic analyses of TGIF family proteins in Drosophila imply their role in cell growth. J Genet Genomics 35: 457-65. doi: 10.1016/S1673-8527(08)60063-6

Whitton A, Hyzy SL, Britt C, Williams JK, Boyan BD, Olivares-Navarrete R (2016) Differential spatial regulation of BMP molecules is associated with single-suture craniosynostosis. J Neurosurg Pediatr 18: 83-91. doi: 10.3171/2015.12.PEDS15414

Willer CJ, Li Y, Abecasis GR (2010) METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 26: 2190-1. doi: 10.1093/bioinformatics/btq340

Yoon PW, Rasmussen SA, Lynberg MC, Moore CA, Anderka M, Carmichael SL, Costa P, Druschel C, Hobbs CA, Romitti PA, Langlois PH, Edmonds LD (2001) The National Birth Defects Prevention Study. Public Health Rep 116 Suppl 1: 32-40. doi: 10.1093/phr/116.S1.32

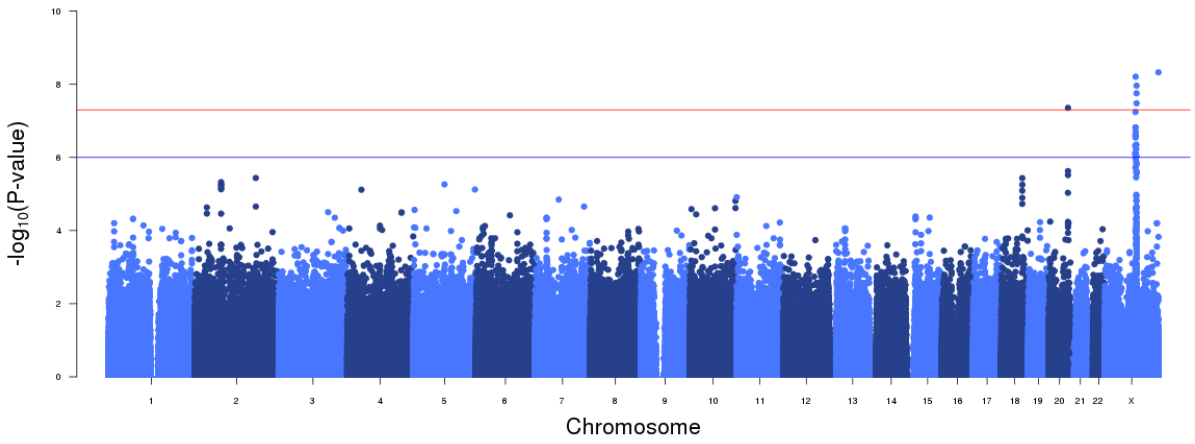
Figure 1. Plot of GWAS TDT *P*-values from analysis of 215 mNCS case-parent triads (649,669 SNPs). The values on the *x*-axis are the genomic positions of the markers, and the values on the *y*-axis are -$\log_{10}$ of the *P*-values. The blue line represents the suggestive significance threshold of $P \leq 1x10^{-5}$, whereas the red line represents a genome-wide significance threshold of $P \leq 5x10^{-8}$.

Figure 2. Regional association plot and LD plot for association of rs6127972 to mNCS. (a) *P*-values (-$\log_{10}$) of the GWAS are plotted against the genomic position of each SNP on chromosome 20 associated region (± 150 kb from rs6127972), with genes in the region depicted below. The LD (r^2) between rs6127972 and other SNPs are indicated with different colors. Recombination rate in cM per Mb using HapMap controls. (**b**) LD plot (r^2) based on genotyped subjects (215 NHW case-parent triads).

Figure 3. Luciferase assay of rs6127972 intronic to *BMP7*. (a) A 667 bp DNA fragment with either the common G allele or risk T allele at our GWAS significant SNP, rs6127972, were cloned at the RE position, 5′ upstream of a BMP7 promoter reporter construct. Expression of the luciferase reporter for both fragments was significantly lower compared to the empty control (G allele $P=3.78x10^{-5}$; T allele $P=5.11x10^{-6}$). There was no statistically significant difference in the expression of the luciferase reporter between the G allele and T allele fragments ($P=0.038$). (b) A randomly selected 667 bp DNA fragment with either G allele or T allele at rs1857502 (control fragment) was cloned at the RE position, 5′ upstream of a BMP7 promoter reporter construct. Expression of the luciferase reporter for both fragments did not change compared to the empty control (G allele P=0.188; T allele P=0.041). (*) represents significance at a level of 0.05 compared to the empty vector control. Bars represent mean ± SE.

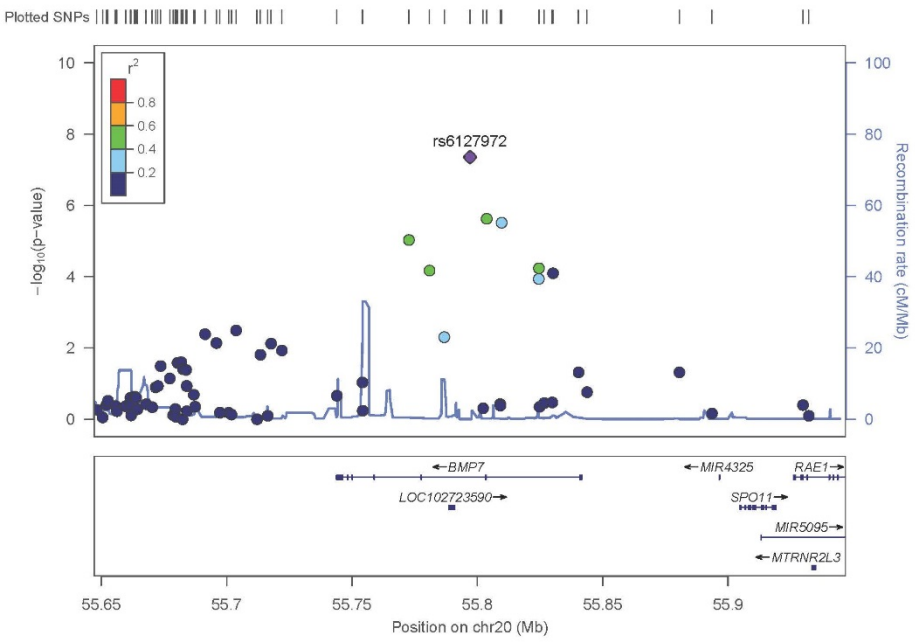
Figure 4. Average percent methylation of DNA isolated from 12 pairs of MSCs derived from fused metopic and open sutures from the same proband at three CpG sites. Of the seven CpG sites covered by the two sets of primers, three of the sites had 100% methylation across all specimens, and one site was extremely variable, so these sites were not analyzed further. Three sites (chr20:55,798,831-55,798,832,

chr20:55,798,857-55,798,858, chr20:55,798,873-55,798,874) were consistent across replicates and analyzed to determine any genotype- or phenotype-specific differences. No significant genotypic-specific differences were identified. There was a significant difference in methylation between the fused metopic and open sutures at the three analyzed sites, with *P*-values of 6.4×10^{-6}, 1.39×10^{-5}, and 0.012, respectively.

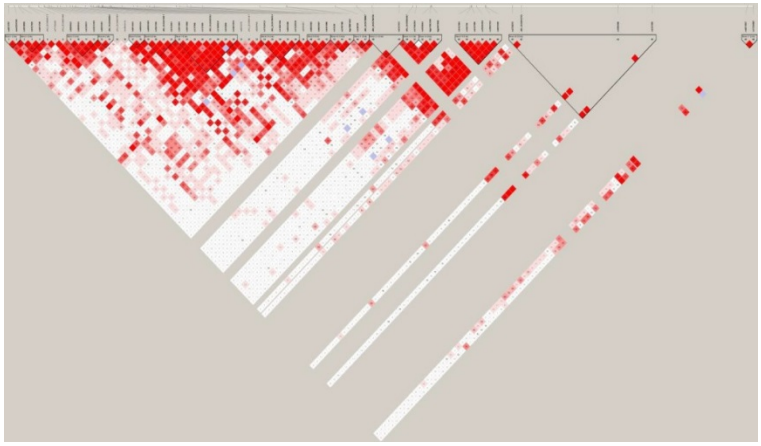


GWAS, genome-wide association study; mNCS, metopic nonsyndromic craniosynostosis; SNP, single nucleotide polymorphism; TDT, transmission disequilibrium test.

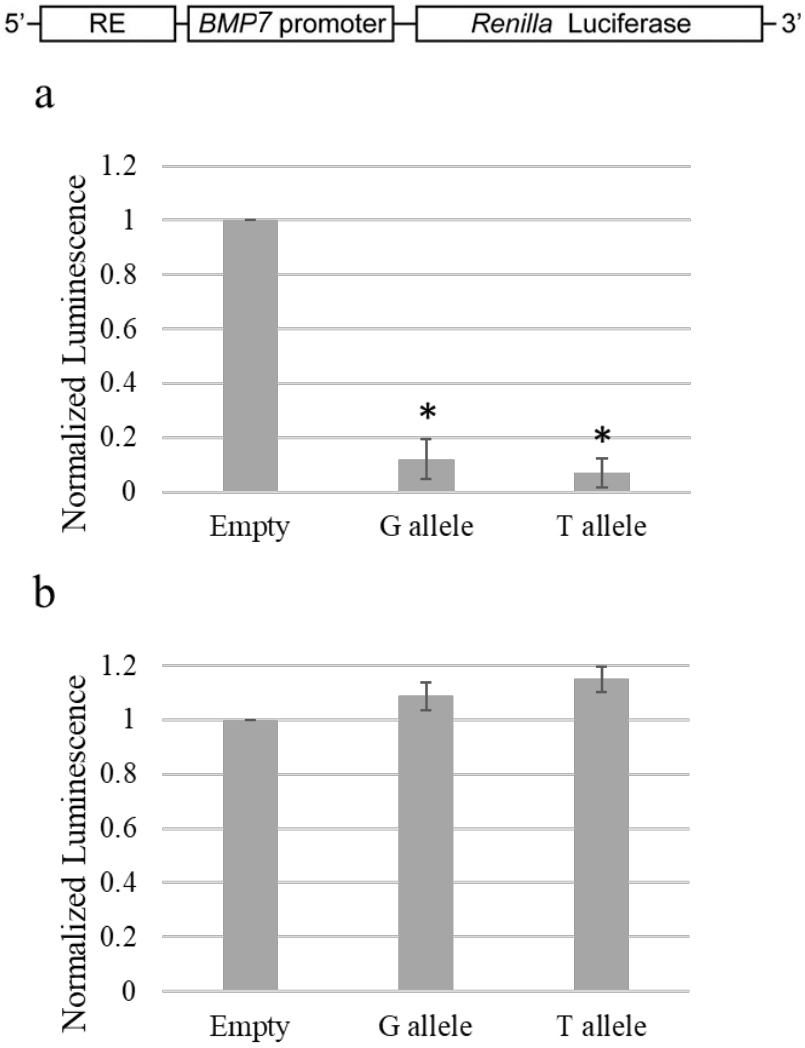
a.



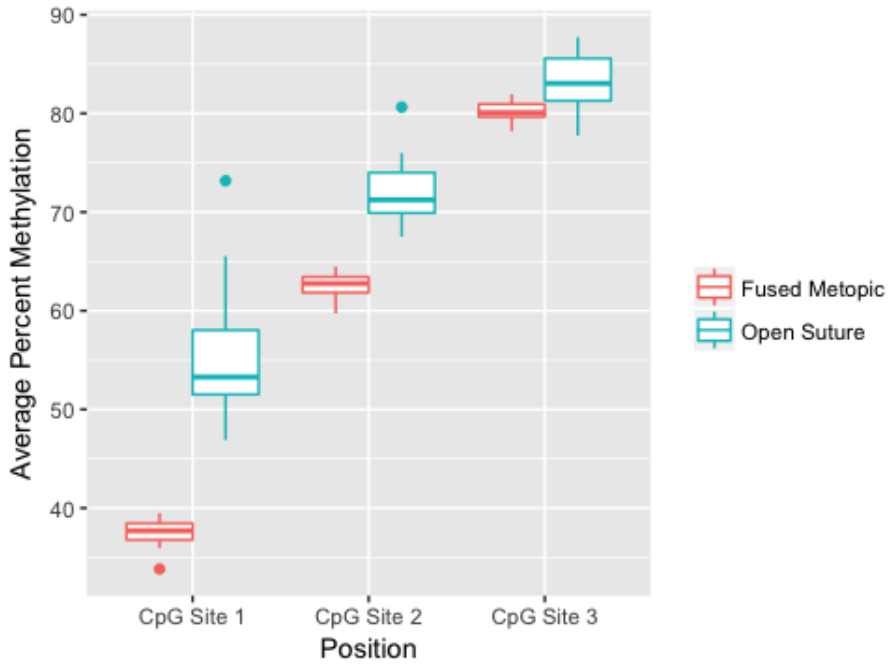
b.



cM centimorgan; GWAS, genome-wide association study; LD, linkage disequilibrium; Mb, megabase; mNCS, metopic nonsyndromic craniosynostosis; NHW, non-Hispanic White; SNP, single nucleotide polymorphism.



(*) represents significance at a level of $P < 0.05$ compared to the empty vector control. Bars represent mean ± SE. bp, base pair; GWAS, genome-wide association study; RE, regulatory element; SNP, single nucleotide polymorphism.



MSCs, mesenchymal stem cells. The control sites were consistent across replicates and showed no significant differences in percent methylation between the fused and open sutures.