nature genetics

A genome-wide association study identifies susceptibility loci for nonsyndromic sagittal craniosynostosis near *BMP2* and within *BBS9*

Cristina M Justice[1,24], Garima Yagnik[2,24], Yoonhee Kim[1], Inga Peter[3], Ethylin Wang Jabs[3], Monica Erazo[3], Xiaoqian Ye[3], Edmond Ainehsazan[3], Lisong Shi[3], Michael L Cunningham[4], Virginia Kimonis[5], Tony Roscioli[6], Steven A Wall[7], Andrew O M Wilkie[7,8], Joan Stoler[9], Joan T Richtsmeier[10], Yann Heuzé[10], Pedro A Sanchez-Lara[11], Michael F Buckley[12], Charlotte M Druschel[13], James L Mills[14], Michele Caggana[15], Paul A Romitti[16], Denise M Kay[15], Craig Senders[17], Peter J Taub[18], Ophir D Klein[19–21], James Boggan[22], Marike Zwienenberg-Lee[22], Cyrill Naydenov[23], Jinoh Kim[2], Alexander F Wilson[1] & Simeon A Boyadjiev[2]

Sagittal craniosynostosis is the most common form of craniosynostosis, affecting approximately one in 5,000 newborns. We conducted, to our knowledge, the first genome-wide association study for nonsyndromic sagittal craniosynostosis (sNSC) using 130 non-Hispanic case-parent trios of European ancestry (NHW). We found robust associations in a 120-kb region downstream of *BMP2* flanked by rs1884302 ($P = 1.13 \times 10^{-14}$, odds ratio (OR) = 4.58) and rs6140226 ($P = 3.40 \times 10^{-11}$, OR = 0.24) and within a 167-kb region of *BBS9* between rs10262453 ($P = 1.61 \times 10^{-10}$, OR = 0.19) and rs17724206 ($P = 1.50 \times 10^{-8}$, OR = 0.22). We replicated the associations to both loci (rs1884302, $P = 4.39 \times 10^{-31}$ and rs10262453, $P = 3.50 \times 10^{-14}$) in an independent NHW population of 172 unrelated probands with sNSC and 548 controls. Both *BMP2* and *BBS9* are genes with roles in skeletal development that warrant functional studies to further understand the etiology of sNSC.

Skull development is a complex process that involves ongoing interaction between the bones of the skull and cranial soft tissues. The cranial vault is comprised of intramembranous bones joined by sutures of dense fibrous tissue that accommodate the growing brain. Bone is added at these sutures during growth, and the skull eventually ossifies completely. Craniosynostosis, the premature closure of one or more of the cranial vault sutures, is a common congenital anomaly. Approximately 80% of craniosynostosis occurs as an isolated anomaly, called nonsyndromic craniosynostosis (NSC), without major associated malformations[1]. Rare mutations in the *FGFR2*, *TWIST1*, *FREM1*, *LRIT3*, *EFNA4* and *RUNX2* duplications have been reported in a minor fraction of individuals with NSC[2–9]. The remainder (~20%) of craniosynostosis cases are syndromic, occurring with one or more additional major malformations caused by single-gene mutations in one of at least eight genes (*FGFR1*, *FGFR2*, *FGFR3*, *TWIST1*, *EFNB1*, *POR*, *MSX2* and *RAB23*), involving primarily the coronal sutures[10–12].

The most common type of NSC is sNSC, which accounts for 40–58% of all cases[13,14]. The etiology of sNSC is not well understood; however, the published literature suggests that it is a multifactorial condition in which both genetic and environmental factors have a role, as indicated by a higher rate of concordance in monozygotic as compared to dizygotic twins (30% compared to 0%, respectively)[15],

[1]Genometrics Section, Inherited Disease Research Branch, Division of Intramural Research, National Human Genome Research Institute, US National Institutes of Health (NIH), Baltimore, Maryland, USA. [2]Section of Genetics, Department of Pediatrics, University of California Davis, Sacramento, California, USA. [3]Department of Genetics and Genomic Sciences, Mount Sinai School of Medicine, New York, New York, USA. [4]Department of Pediatrics, Division of Craniofacial Medicine, University of Washington and Seattle Children's Research Institute, Seattle, Washington, USA. [5]Department of Pediatrics, Division of Genetics, University of California Irvine, Irvine, California, USA. [6]School of Women's and Children's Health, Sydney Children's Hospital, University of New South Wales, Sydney, New South Wales, Australia. [7]Craniofacial Unit, Oxford University Hospitals National Health Service Trust, John Radcliffe Hospital, Oxford, UK. [8]Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK. [9]Division of Genetics, Children's Hospital Boston, Harvard University, Boston, Massachusetts, USA. [10]Department of Anthropology, Pennsylvania State University, University Park, Pennsylvania, USA. [11]Department of Pathology and Pediatrics, Children's Hospital Los Angeles, University of Southern California, Los Angeles, California, USA. [12]Department of Haematology and Genetics, South Eastern Area Laboratory Services, Sydney, New South Wales, Australia. [13]Congenital Malformations Registry, New York State Department of Health, Albany, New York, USA. [14]Division of Epidemiology, Statistics and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, NIH, Department of Health and Human Services, Bethesda, Maryland, USA. [15]Division of Genetics, Wadsworth Center, New York State Department of Health, Albany, New York, USA. [16]Department of Epidemiology, College of Public Health, University of Iowa, Iowa City, Iowa, USA. [17]Department of Otolaryngology, University of California Davis, Sacramento, California, USA. [18]Division of Plastic and Reconstructive Surgery, Kravis Children's Hospital, Mount Sinai Medical Center, New York, New York, USA. [19]Department of Orofacial Sciences, University of California San Francisco, San Francisco, California, USA. [20]Department of Pediatrics, University of California San Francisco, San Francisco, California, USA. [21]Program in Craniofacial and Mesenchymal Biology, University of California San Francisco, San Francisco, California, USA. [22]Department of Neurological Surgery, University of California Davis, Sacramento, California, USA. [23]Department of Chemistry and Biochemistry, Medical University, Sofia, Bulgaria. [24]These authors contributed equally to this work. Correspondence should be addressed to S.A.B. (simeon.boyd@ucdmc.ucdavis.edu).

Received 3 May; accepted 4 October; published online 18 November 2012; doi:10.1038/ng.2463

an increased male-to-female ratio (3:1)[1,16] and a high risk of recurrence in affected families[17]. Several environmental factors have also been associated with sNSC, including parity, prematurity, intrauterine constraint and maternal tobacco or nitrosatable drug use[18–21].

In an attempt to identify susceptibility loci for sNSC, we recruited 201 case-parent trios and 13 nuclear families with at least one individual with sNSC (68% NHW and 32% mixed-ethnicity families). These participants were enrolled and evaluated through the collaborative effort of the International Craniosynostosis Consortium (https://genetics.ucdmc.ucdavis.edu/icc.cfm). Samples were genotyped on the Illumina 1M Human Omni1-Quad array. The final criteria for the current analyses were based on strict sNSC phenotype (no additional synostoses, congenital anomalies or developmental delay in the proband or any affected sibling) and NHW ethnicity as determined by self-reported ancestry and confirmed by principal component analysis using EIGENSOFT 3.0 (ref. 22). Application of these criteria reduced our discovery population to 130 case-parent trios.

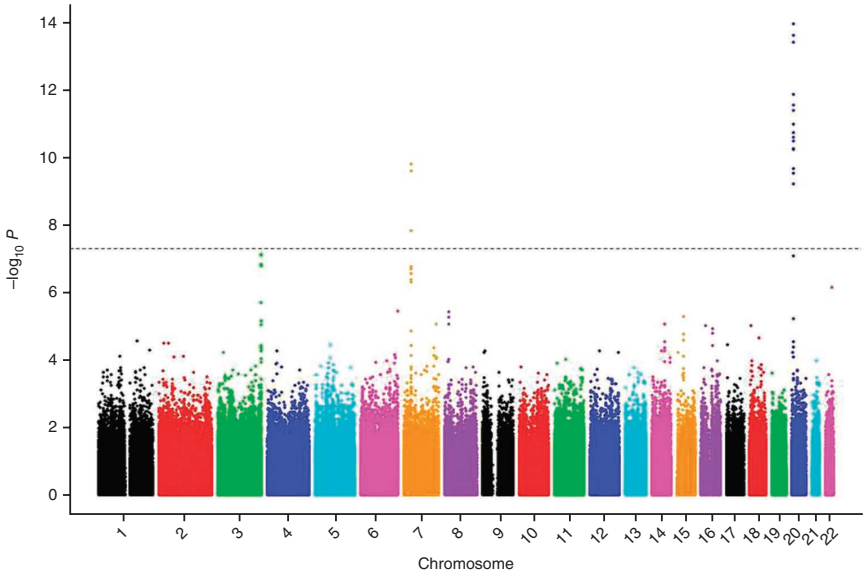

Figure 1 Manhattan plot of the *P* values obtained from the genome-wide TDT of 130 trios ($N = 914{,}402$). The *x* axis corresponds to the genomic position of the autosomes, and the *y* axis shows the $-\log_{10}$ of the *P* value. The horizontal dashed line corresponds to the genome-wide significance threshold of $P \leq 5 \times 10^{-8}$.

After stringent quality-control procedures, we retained genotypes for 915,307 SNPs (914,402 autosomal and 905 pseudoautosomal), which we analyzed using the transmission disequilibrium test (TDT) as implemented in PLINK v1.07 (ref. 23). This resulted in the identification of 21 SNPs, 18 on 20p12.3 and 3 on 7p14.3, that reached the genome-wide significance threshold of $P < 5 \times 10^{-8}$ (**Fig. 1** and **Supplementary Fig. 1**). The 18 SNPs on chromosome 20 were in two linkage disequilibrium (LD) blocks (92 kb and 19 kb on chromosome 20p12.3; **Fig. 2a**, **Table 1** and **Supplementary Fig. 2**), with the most significant SNP, rs1884302 ($P = 1.13 \times 10^{-14}$), located approximately 345 kb downstream of *BMP2*. Conditional analysis using extended TDT in UNPHASED[24] confirmed that rs1884302 led these signals (data not shown). Except for a predicted mRNA, *BC038533*, we did not identify any other known genes, copy number variants, transcripts, microRNAs or regulatory elements in this region. However, we identified at least four highly conserved genome regions with multiple predicted sites for transcription-factor binding using bioinformatic analysis in the vicinity of rs1884302 (**Supplementary Fig. 3**). Exonic resequencing of *BMP2* in our cohort did not identify any variants with obvious deleterious effects. The three SNPs on chromosome 7p14.3 span a region of 167 kb within *BBS9* introns 4 and 15 (**Fig. 2b** and **Supplementary Fig. 4**). RT-PCR sequencing of complementary DNA (cDNA) from eight sNSC calvarial osteoblast cell lines revealed six *BBS9* variants, but three were previously reported SNPs, and all were benign on the basis of bioinformatic protein prediction tools and conservation analysis (Online Methods). We also searched for two-way gene-gene interactions using the 100 most significant TDT genome-wide association study (GWAS) SNPs using a conditional logistic regression model[25]. No results were significant after correcting for multiple testing (1,247 tests, Bonferroni $P \leq 4 \times 10^{-5}$; **Supplementary Table 1**).

Considering the male predominance of sNSC, we performed a separate TDT analysis for 104 NHW trios with a male proband. The 21 SNPs on chromosomes 20 and 7 remained genome-wide significant, with an additional pseudoautosomal marker reaching a low, but not genome-wide statistically significant, *P* value (rs2522623, $P = 5.75 \times 10^{-5}$) and thus did not meet our criteria for replication. This marker is located in *PCDH11X* (encoding protocadherin 11 X-linked), which has not previously been reported to have a functional role in sNSC. Analysis of a larger cohort may show a stronger association to this locus.

From a total of 46 autosomal SNPs that showed significant and suggestive associations with $P < 1 \times 10^{-5}$, we prioritized 25 SNPs for the replication study (**Supplementary Table 2**). To confirm the observed signals, we genotyped a NHW replication population of 172 unrelated cases with sNSC and 548 unaffected controls. We considered our data as a replication when the direction of effect of the alleles surpassing nominal significance was consistent between the discovery and replication datasets. All seven genotyped markers on chromosome 20 and four out of the five markers on 7p14.3 (with one SNP failing genotyping) were successfully replicated with meta-analysis $P < 10^{-9}$ (**Supplementary Tables 2** and **3**). There was little or no evidence of heterogeneity of the ORs in the two different sample sets for the 11 replicated SNPs. Additionally, genotypic TDT analysis indicated the effect of SNPs was consistent with an additive model (**Supplementary Table 4**). We detected that, in most cases, the transmitted minor alleles conferred a 2.36- to 4.38-fold risk of sNSC at the chromosome 20 markers, whereas the transmitted minor alleles at the chromosome 7 markers were associated with a decreased risk of sNSC (**Supplementary Table 3**). We were unable to replicate other signals from the GWAS TDT with $P < 1 \times 10^{-5}$. These associations could be false positives or were missed because of differences between the study design and statistical methods used for the discovery and replication stages. We also relied on self-reported maternal ethnicity instead of ancestry-informative markers for the replication cohort, which limited our ability to precisely assign ethnicity to the case-control study. The effect of a gene-environment interaction or the relatively small sample of the replication cohort may also have affected the findings.

We imputed borderline significant regions on chromosomes 3, 7, 8, 15 and 22 using BEAGLE v3.3.2 (ref. 26) to identify additional

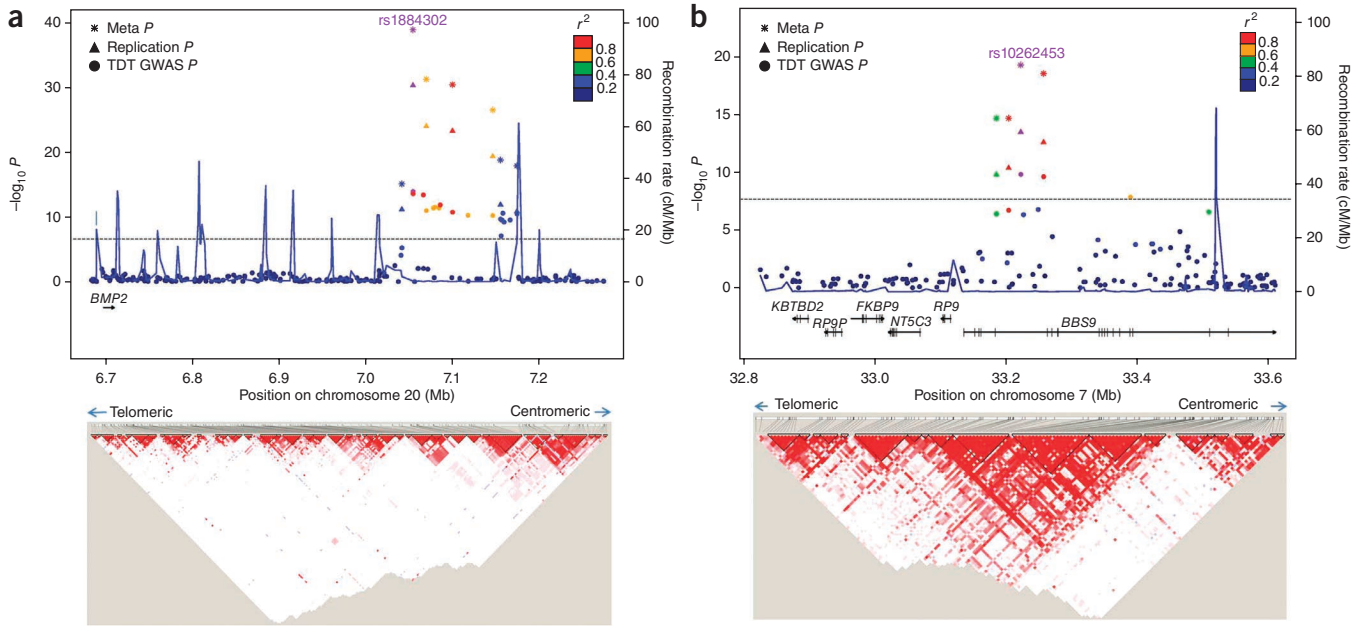

Figure 2 Regional association plots for associations with sNSC at genome-wide significance. (**a**,**b**) *P* values ($-\log_{10}$) of the GWAS (solid circle), replication (triangle) and meta-analysis (asterisk) on the *y* axis are plotted against the genomic positions of each SNP on the *x* axis on chromosomes 20 (**a**) and 7 (**b**) (top). The horizontal dashed line corresponds to the genome-wide significance threshold of $P \leq 5 \times 10^{-8}$. Genes in the region are shown below (not to scale). The LD values (r^2) between the lead SNP and the other SNPs are indicated in different colors. The blue lines indicate the recombination rates in cM per Mb using HapMap controls. LD plots (on the basis of r^2 values) using genotyped subjects from the GWAS study (bottom).

genome-wide significant SNPs. Fifty-two intronic SNPs in *DLG1* reached genome-wide significance (**Supplementary Table 5**), but we previously could not replicate this association in our case-control sample (**Supplementary Table 2**). The lack of replication of a common variant may be a result of our small sample size, differences in study design between our GWAS and replication datasets or rare variants that contribute to risk. We performed the imputation using the 1000 Genomes data as the reference panel, which includes some sequence variants with minor allele frequencies (MAFs) <1%. Given that sNSC has a prevalence of 1 in 5,000, we suspect that rare functional variants not identified by 1000 Genomes may be involved, and sequencing of the associated region will be necessary. We found no statistically significant signals after the imputation at the other loci.

Using these modestly sized populations, we discovered and replicated common genetic variants in at least two loci that contribute to sNSC. The locus on chromosome 20 is located in the vicinity of *BMP2*, which has a role in skeletal development. BMP2 is a member of the TGF-β superfamily that activates SMAD1, SMAD5 and SMAD8 and, together with other osteogenic transcription factors such as RUNX2, ultimately regulates osteoblast development[27].

An interesting observation in humans and animal models is that mutations with opposite effects in the same gene result in craniosynostosis (accelerated ossification) or parietal foramina or large fontanelles (deficient ossification) phenotypes. This holds true for mutations in at least four genes: *MSX2* (refs. 28–30), *TWIST1* (refs. 31,32), *RUNX2* (refs. 6,7,33) and *Nell1* (refs. 34,35). Consistent with this, gain-of-function mutations in *ALX4*, a gene that is associated with parietal foramina, have previously been identified in sNSC probands by our group[36]. Similarly, a variant in *BMP2*, p.Ser37Ala, has been associated with osteoporosis[37]; thus, if p.Ser37Ala results in loss of function, one can speculate that gain-of-function variants causing overexpression of *BMP2* would result in a phenotype of accelerated ossification and, potentially, craniosynostosis. Additionally, a limb-specific regulatory element located 110 kb downstream of *BMP2* is mutated in brachydactyly type A2 (ref. 38). Therefore, we predict that a skull-specific *BMP2* regulatory element is located within or near the association region. Also, noncoding SNPs can affect variation in the expression of one or many genes. Even though none of our top hits was in the expression quantitative trait loci databases, this may be because no human expression quantitative trait loci studies have been conducted in the skull suture tissues.

The GWAS and the replication analyses also showed evidence for association on chromosome 7p14 within an intronic region of *BBS9*. BBS9 is a component of the multiprotein complex BBSome, which has crucial roles in intraflagellar transport and moving cargo molecules in and out of cilia. Loss-of-function mutations in *BBS9* have been found in patients with Bardet-Biedl syndrome (BBS)[39]. BBS arises from defects of primary cilia that function as a platform

Table 1 GWAS, replication study and meta-analysis results for most significant SNP on chromosomes 20 and 7

				GWAS		Replication		Meta-analysis		
Chr.	SNP	Alleles[a]	RAF[b]	OR (95% CI)	*P*	OR (95% CI)	*P*	OR (95% CI)	*P*	P_{het}[c]
20	rs1884302	C/T	0.3	4.58 (2.99–6.97)	1.1×10^{-14}	4.30 (3.33–5.56)	4.4×10^{-31}	4.38 (3.51–5.45)	1.1×10^{-39}	0.81
7	rs10262453	C/A	0.69	0.20 (0.11–0.34)	1.6×10^{-10}	0.26 (0.18–0.37)	3.5×10^{-14}	0.24 (0.17–0.32)	5.6×10^{-20}	0.44

[a]Minor allele/major allele. [b]Risk allele frequency (RAF) calculated from the replication study controls. [c]*P* value for heterogeneity between the discovery and replication cohorts. Underlined bases are overtransmitted alleles. Chr., chromosome.

for many important signaling events involving, at least, PDGFR-α, hedgehog, epidermal growth factor and the 5-HT_6 serotonin receptor[40–44]. Although BBS does not present with suture phenotypes, it does present with a skeletal phenotype (polydactyly) and orofacial anomalies[39]. It is noteworthy that aberrant signaling of PDGFR-α and epidermal growth factor or hedgehog has been implicated in craniofacial development[45–49]. In addition, it has been shown that FGF signaling regulates cilia length and function during the development of zebrafish and *Xenopus*[50]. Thus, it is possible that craniosynostosis-causing gain-of-function mutations in *FGFR* genes also work through aberrant assembly and/or signaling of cilia. It remains to be seen whether gain-of-function mutations or overexpression of *BBS9* would influence suture fusion. Three of the genes in the vicinity of the association on chromosome 7 (*RP9P*, *KBTBD2* and *FKBP9*) are pseudogenes, and the remaining two genes in the region, *RP9* and *NT5C3*, have no known role in skeletal development and cause retinitis pigmentosa and hemolytic anemia, respectively[51,52]. This region is located approximately 14 Mb telomeric of *TWIST1* and is thus unlikely to affect its function.

Our findings implicate variants within or near genes with relevant function in skeletal development, suggesting a major genetic effect for sNSC. A larger sample size will be needed to detect additional susceptibility genes with moderate genetic effects. The lack of Mendelian segregation suggests a more complex pattern of inheritance for sNSC. The causative variants may be of low frequency and in LD with the common alleles associated with sNSC in this study. Another possibility is that one or more environmental factors interacting with genetic variants contribute to sNSC. Less likely, but still possible, are oligogenic inheritance, gene-gene interactions, imprinting, epigenetic changes or incomplete penetrance of the variants.

We identified two biologically plausible candidate genes in our study, *BMP2* and *BBS9*, both of which are involved in skeletal development. Functional studies of these genes, such as RNA sequencing and, possibly, tiling arrays of the regions to look for noncoding or microRNAs, are warranted. The potential contribution of these loci to other types of NSC is worthy of investigation. Our results represent the first major step toward deciphering the genetic etiology of sNSC.

METHODS

Methods and any associated references are available in the online version of the paper.

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

The authors thank all families who contributed to this study. S.A.B. is partially funded through a Children's Miracle Network Endowed Chair and through grants K23 DE00462, R03 DE016342 and R01 DE016886 from the National Institute of Dental and Craniofacial Research (NIDCR)/NIH and M01-RR00052 from the National Center for Research Resources/NIH and was fully supported by Zlatka, Anton and Alec Boyadjiev. Partial funding was also obtained from grants to E.W.J. (US Centers for Disease Control and Prevention (CDC) 5 R01 DD000350), M.L.C. (R01 DE018227), A.O.M.W. (Wellcome Trust 093329), P.A.R. (CDC 5U01DD000492), J.K. (NIDCR/NIH R21DE022419), J.L.M. (Intramural Research Program (IRP) of the NIH, Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD); IRP HHSN267200703431C; NICHD N01-DK-7-3431), P.A.S.-L. (Robert Wood Johnson Foundation 3R37DE012711-13S1 and Children's Hospital Los Angeles–University of Southern California Child Health Research Career Development Program; NIH K12-HD05954), J.T.R. (NIDCR/NIH and the American Recovery and Reinvestment Act R01 DE018500 and 3R01 DE018500-02S1) and I.P. (National Center for Advancing Translational Sciences, NIH UL1TR000067). This project was also supported in part by the Division of Intramural Research Program of the National Human Genome Research Institute, NIH (C.M.J., Y.K. and A.F.W.). Genotyping services were provided by the Center for Inherited Disease Research (CIDR). CIDR is fully funded through a federal contract from the NIH to Johns Hopkins University, contract number HHSN268200782096C. We thank B. Wilson, N. Issac, C. Nauta, E. Goude, E. Cherkez, L. Peters and J. Harrison for patient recruitment, C. Boehm and A. Scott for coordination of discovery-phase genotyping, C. Stevens, A. Stoner, J.L. Liu, A. Gearhart, A. Atkins and E. McGrath for bench work, D. Mortlock for bioinformatic analysis and informative discussion and J. Graham (Cedars-Sinai Hospital, Los Angeles, California, USA), J. Bernstein (Stanford University, Palo Alto, California, USA), J. Marsh (Washington University, St. Louis, Missouri, USA), J. Panchal (University of Oklahoma Health Science Center, Oklahoma, USA), T. Tollefson (University of California Davis, Sacramento, California, USA) and M. Passos-Bueno (University of São Paolo, Brazil) for contributing clinical information and biospecimens for this project.

AUTHOR CONTRIBUTIONS

C.M.J. and G.Y. are the first coauthors of the manuscript. C.M.J., Y.K. and I.P. performed statistical analyses. G.Y., M.E., X.Y., E.A. and L.S. performed experiments. C.M.J., G.Y., Y.K., I.P., M.L.C., V.K., T.R., A.O.M.W., J.S., J.T.R., Y.H., P.A.S.-L., M.F.B., J.K., A.F.W. and S.A.B. analyzed data. C.M.J. and G.Y. wrote the manuscript, with contributions from Y.K., I.P., J.T.R., Y.H., P.A.R., A.F.W. and S.A.B. E.W.J., M.L.C., V.K., S.A.W., J.S., P.A.S.-L., M.F.B., C.M.D., J.L.M., M.C., P.A.R., D.M.K., C.S., P.J.T., O.D.K., J.B., M.Z.-L. and C.N. contributed materials and reagents. E.W.J., A.O.M.W., M.F.B., D.M.K. and P.J.T. contributed to editing of the manuscript. E.W.J., C.M.D., J.L.M., M.C., P.A.R., D.M.K., C.S., P.J.T., O.D.K., J.B., M.Z.-L. and C.N. helped with experimental design. J.K. and A.F.W. supervised the research. S.A.B. was the principal investigator, designed and supervised research and recruited and evaluated participants.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/doifinder/10.1038/ng.2463.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Cohen, M.M. *Craniosynostosis: Diagnosis, Evaluation, and Management* (Oxford University Press, New York, 2000).
2. Johnson, D. *et al.* A novel mutation, Ala315Ser, in FGFR2: a gene-environment interaction leading to craniosynostosis? *Eur. J. Hum. Genet.* **8**, 571–577 (2000).
3. Seto, M.L. *et al.* Isolated sagittal and coronal craniosynostosis associated with TWIST box mutations. *Am. J. Med. Genet. A* **143**, 678–686 (2007).
4. Weber, I. *et al.* Molecular analysis of 74 patients with craniosynostosis. *Eur. J. Hum. Genet.* **9** (suppl. 1), P0409, 179 (2001).
5. Merrill, A.E. *et al.* Cell mixing at a neural crest-mesoderm boundary and deficient ephrin-Eph signaling in the pathogenesis of craniosynostosis. *Hum. Mol. Genet.* **15**, 1319–1328 (2006).
6. Wilkie, A.O. *et al.* Clinical dividends from the molecular genetic diagnosis of craniosynostosis. *Am. J. Med. Genet. A* **143A**, 1941–1949 (2007).
7. Mefford, H.C. *et al.* Copy number variation analysis in single-suture craniosynostosis: multiple rare variants including RUNX2 duplication in two cousins with metopic craniosynostosis. *Am. J. Med. Genet. A* **152A**, 2203–2210 (2010).
8. Vissers, L.E. *et al.* Heterozygous mutations of *FREM1* are associated with an increased risk of isolated metopic craniosynostosis in humans and mice. *PLoS Genet.* **7**, e1002278 (2011).
9. Kim, S.-D. *et al.* Leucine-rich repeat, immunoglobulin-like and transmembrane domain 3 (LRIT3) is a modulator of FGFR1. *FEBS Lett.* **586**, 1516–1521 (2012).
10. Melville, H. *et al.* Genetic basis of potential therapeutic strategies for craniosynostosis. *Am. J. Med. Genet. A* **152A**, 3007–3015 (2010).
11. Wilkie, A.O. *et al.* Prevalence and complications of single-gene and chromosomal disorders in craniosynostosis. *Pediatrics* **126**, e391–e400 (2010).
12. Passos-Bueno, M.R. *et al.* Genetics of craniosynostosis: genes, syndromes, mutations and genotype-phenotype correlations. *Front. Oral Biol.* **12**, 107–143 (2008).
13. Kimonis, V. *et al.* Genetics of craniosynostosis. *Semin. Pediatr. Neurol.* **14**, 150–161 (2007).
14. Kolar, J.C. An epidemiological study of nonsyndromal craniosynostoses. *J. Craniofac. Surg.* **22**, 47–49 (2011).
15. Lajeunie, E. *et al.* Genetic considerations in nonsyndromic midline craniosynostoses: a study of twins and their families. *J. Neurosurg.* **103** (suppl.), 353–356 (2005).
16. Lajeunie, E. *et al.* Syndromal and nonsyndromal primary trigonocephaly: analysis of a series of 237 patients. *Am. J. Med. Genet.* **75**, 211–215 (1998).
17. Lajeunie, E. *et al.* Genetic study of scaphocephaly. *Am. J. Med. Genet.* **62**, 282–285 (1996).
18. Sanchez-Lara, P.A. *et al.* Fetal constraint as a potential risk factor for craniosynostosis. *Am. J. Med. Genet. A* **152A**, 394–400 (2010).
19. Zeiger, J.S. *et al.* Genetic and environmental risk factors for sagittal craniosynostosis. *J. Craniofac. Surg.* **13**, 602–606 (2002).
20. Gardner, J.S. *et al.* Maternal exposure to prescription and non-prescription pharmaceuticals or drugs of abuse and risk of craniosynostosis. *Int. J. Epidemiol.* **27**, 64–67 (1998).

21. Källén, K. Maternal smoking and craniosynostosis. *Teratology* **60**, 146–150 (1999).
22. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet.* **2**, e190 (2006).
23. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
24. Dudbridge, F. Pedigree disequilibrium tests for multilocus haplotypes. *Genet. Epidemiol.* **25**, 115–121 (2003).
25. Cordell, H.J. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum. Mol. Genet.* **11**, 2463–2468 (2002).
26. Browning, B.L. & Browning, S.R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84**, 210–223 (2009).
27. Rosen, V. BMP2 signaling in bone development and repair. *Cytokine Growth Factor Rev.* **20**, 475–480 (2009).
28. Jabs, E.W. *et al.* A mutation in the homeodomain of the human *MSX2* gene in a family affected with autosomal dominant craniosynostosis. *Cell* **75**, 443–450 (1993).
29. Spruijt, L. *et al.* A novel mutation in the *MSX2* gene in a family with foramina parietalia permagna (FPP). *Am. J. Med. Genet. A* **139**, 45–47 (2005).
30. Wilkie, A.O. *et al.* Functional haploinsufficiency of the human homeobox gene MSX2 causes defects in skull ossification. *Nat. Genet.* **24**, 387–390 (2000).
31. Howard, T.D. *et al.* Mutations in TWIST, a basic helix-loop-helix transcription factor, in Saethre-Chotzen syndrome. *Nat. Genet.* **15**, 36–41 (1997).
32. Stankiewicz, P. *et al.* Phenotypic findings due to trisomy 7p15.3-pter including the TWIST locus. *Am. J. Med. Genet.* **103**, 56–62 (2001).
33. Mundlos, S. *et al.* Mutations involving the transcription factor CBFA1 cause cleidocranial dysplasia. *Cell* **89**, 773–779 (1997).
34. Desai, J. *et al.* Nell1-deficient mice have reduced expression of extracellular matrix proteins causing cranial and vertebral defects. *Hum. Mol. Genet.* **15**, 1329–1341 (2006).
35. Zhang, X. *et al.* The role of NELL-1, a growth factor associated with craniosynostosis, in promoting bone regeneration. *J. Dent. Res.* **89**, 865–878 (2010).
36. Yagnik, G. *et al.* ALX4 gain-of-function mutations in nonsyndromic craniosynostosis. *Hum. Mutat.* published online, doi:10.1002/humu.22166 (24 July 2012).
37. Styrkarsdottir, U. *et al.* Linkage of osteoporosis to chromosome 20p12 and association to BMP2. *PLoS Biol.* **1**, E69 (2003).
38. Dathe, K. *et al.* Duplications involving a conserved regulatory element downstream of BMP2 are associated with brachydactyly type A2. *Am. J. Hum. Genet.* **84**, 483–492 (2009).
39. Tobin, J.L. & Beales, P.L. Bardet-Biedl syndrome: beyond the cilium. *Pediatr. Nephrol.* **22**, 926–936 (2007).
40. Marshall, W.F. & Nonaka, S. Cilia: tuning in to the cell's antenna. *Curr. Biol.* **16**, R604–R614 (2006).
41. Schneider, L. *et al.* PDGFRαα signaling is regulated through the primary cilium in fibroblasts. *Curr. Biol.* **15**, 1861–1866 (2005).
42. Haycraft, C.J. *et al.* Gli2 and Gli3 localize to cilia and require the intraflagellar transport protein polaris for processing and function. *PLoS Genet.* **1**, e53 (2005).
43. Ma, R. *et al.* PKD2 functions as an epidermal growth factor–activated plasma membrane channel. *Mol. Cell. Biol.* **25**, 8285–8298 (2005).
44. Brailov, I. *et al.* Localization of 5-HT(6) receptors at the plasma membrane of neuronal cilia in the rat brain. *Brain Res.* **872**, 271–275 (2000).
45. Jenkins, D. *et al.* RAB23 mutations in Carpenter syndrome imply an unexpected role for hedgehog signaling in cranial-suture development and obesity. *Am. J. Hum. Genet.* **80**, 1162–1170 (2007).
46. Ehlen, H.W., Buelens, L.A. & Vortkamp, A. Hedgehog signaling in skeletal development. *Birth Defects Res. C Embryo Today* **78**, 267–279 (2006).
47. Miettinen, P.J. *et al.* Epidermal growth factor receptor function is necessary for normal craniofacial development and palate closure. *Nat. Genet.* **22**, 69–73 (1999).
48. Ding, H. *et al.* A specific requirement for PDGF-C in palate formation and PDGFR-α signaling. *Nat. Genet.* **36**, 1111–1116 (2004).
49. Choi, D.S. *et al.* 5–HT2B receptor-mediated serotonin morphogenetic functions in mouse cranial neural crest and myocardiac cells. *Development* **124**, 1745–1755 (1997).
50. Neugebauer, J.M. *et al.* FGF signalling during embryo development regulates cilia length in diverse epithelia. *Nature* **458**, 651–654 (2009).
51. Keen, T.J. *et al.* Mutations in a protein target of the Pim-1 kinase associated with the RP9 form of autosomal dominant retinitis pigmentosa. *Eur. J. Hum. Genet.* **10**, 245–249 (2002).
52. Marinaki, A.M. *et al.* Genetic basis of hemolytic anemia caused by pyrimidine 5′ nucleotidase deficiency. *Blood* **97**, 3327–3332 (2001).

ONLINE METHODS

Subjects. Informed consent was obtained from all patients and/or their parents. This study was approved by the Institutional Review Boards of the University of California, Davis and all other participating institutions and was conducted in accordance with institutional guidelines. Craniosynostosis was documented by computerized tomography of the skull and/or the protocol of the surgical correction. All patients with sNSC were clinically assessed by a clinical geneticist, and patients with synostosis of additional sutures, associated extracranial congenital anomalies or developmental delays were excluded. Patients for both the initial GWAS and the replication study were recruited within the United States.

The case-control replication sample was comprised of archived, residual newborn blood spots from NHW infants diagnosed with sNSC (case infants) (n = 172) and control infants without a congenital malformation (n = 548). Case infants were delivered by New York State residents from January 1, 1998 through December 31, 2005 and were identified from the New York State Congenital Malformations Registry, a population-based birth defects surveillance program. Control infants were a random sample of live births in New York State delivered during the same time frame with a control-to-case ratio of approximately 3:1. Records for cases and controls were linked to records of the New York State Newborn Screening Program to obtain the respective blood spots and were provided anonymously to study investigators. The use of the blood spots was approved by the Institutional Review Board of the New York State Department of Health and Mount Sinai School of Medicine.

SNP genotyping and quality control in GWAS. Genomic DNA was extracted from whole blood or oral samples (468 whole-blood, 139 mouth-wash, 2 buccal and 24 Oragene samples). A total of 680 samples, including 16 blind duplicates and 13 HapMap controls (11 Utah residents of European ancestry, CEU, and 2 Yoruban controls, YRI) were genotyped on the Illumina 1M Human Omni1-Quad array at the Center for Inherited Disease Research (http://www.cidr.jhmi.edu/). Genotypes for 1,006,422 SNPs and 133,997 copy number variant probes were released. The missing data rate was 0.39%, the blind duplicate rate was 99.99%, and the HapMap sample concordance rate was 99.7%. To check for cryptic relatedness and population stratification, 92,867 independent SNPs were selected on the basis of a pairwise intermarker LD $r^2 < 0.4$, MAF > 0.1 and genotype call rate >0.99. Pairwise identity-by-descent calculations were performed using PLINK v1.0.7 (ref. 53), and inconsistent samples were dropped, reducing the sample size to 646 individuals (213 trios). To confirm the self-reported ancestry of the samples, principal components analysis was performed using EIGENSOFT 3.0 (ref. 22). A total of 145 families (451 individuals) were found to cluster with the 11 HapMap CEU controls, with no other well-defined clusters noted. A strict sNSC phenotype (no additional synostosis, congenital anomalies or developmental delay in the proband or any affected sibling) further reduced our number of NHW trios to 130. SNPs were excluded from analysis if the call rate was less than 98%, they were discordant in more than one duplicate, had a Mendelian inconsistency in more than one trio, were missing in more than 10% of the samples, had MAF = 0.0 and/or had a Hardy-Weinberg equilibrium $P \leq 1 \times 10^{-5}$ with no heterozygotes. This reduced the number of SNPs available for analysis to 914,402 autosomal SNPs and 905 pseudoautosomal SNPs.

SNP selection and quality control in the replication study. Genomic DNA was extracted from 3-mm dried blood spot punches using a sodium hydroxide precipitation protocol described previously[54]. SNPs from the discovery stage were selected if their P value for association was $\leq 1.0 \times 10^{-5}$ in the discovery stage. When multiple SNPs from the same region passed the P value cutoff for replication, SNPs with the lowest P value and/or lowest pairwise LD were selected. Using this screening process, from 46 autosomal SNPs with $P \leq 1.0 \times 10^{-5}$, we selected 25 SNPs to genotype (**Supplementary Table 2**) that tagged all 11 significant regions of interest (one region each on ten different chromosomes and two regions on chromosome 7). SNPs for the replication phase were genotyped using the ABI TaqMan allelic discrimination assay according to the manufacturer's protocol (Applied Biosystems). SNP genotypes were called and discriminated using software S.D.S. v3.4.4. Genotyping was conducted in duplicates, and the concordance rate was 100%. Case and control samples were randomly distributed across the DNA plates. Disease status was unknown at the time of the genotyping experiments. Of the 25 SNPs from the original discovery phase that were selected for replication, two failed multiple TaqMan assay attempts, even when conditions were modified (chromosome 7 rs17724206 failed, and chromosome 22 rs4634034 had call rate <90%). For the remaining SNPs, the genotype call rates were >95%.

Sequencing. *BMP2* sequencing was carried out on both exons for 149 sagittal NSC probands, of which 102 were part of the 130-proband cohort used in the initial discovery phase. Primer sequences for BMP2 are available in **Supplementary Table 6a**. Three variants were identified. Two of them, c.109T>G (p.Ser37Ala) and c.570A>T (p.Arg190Ser), were previously reported (rs2273073 and rs23576, respectively), and the allelic frequencies for the CEU population (obtained from HapMap Rel. 28, Phase II + III) were not significantly different than those found in our sequencing sample. The other SNP, c.393A>T (p.Arg131Ser), was new, present in the proband's father and characterized as nondamaging on the basis of bioinformatic predictions from SIFT (Sorting Tolerant from Intolerant, Fred Hutchinson Cancer Research Center, Seattle, web software, http://blocks.fhcrc.org/sift/SIFT.html) and PPH (PolyPhen, Bork Group and Sunyaev Lab, Harvard, web software, http://genetics.bwh.harvard.edu/pph/).

RT-PCR using control fetal calvarial osteoblast cDNA was attempted with all possible exonic primer combinations for the predicted mRNA *BC038533* but did not produce PCR amplicons. *BBS9* sequencing was performed on cDNA from eight NSC fetal calvarial osteoblast lines. cDNA was created by RT-PCR (using the Invitrogen SuperScript protocol and kit) on RNA extracted from available patients. *BBS9* sequencing primers are available in **Supplementary Table 6b**. Six variants were identified: three (c.1029A>G (p.Gly343Gly), c.2100G>A (p.Leu700Leu) and c.1729A>C (p.Ile577Leu)) were previously reported (rs25195153, rs115809567 and rs34209904, respectively), whereas the remaining three (BBS9 3′ untranslated region (UTR)+108T>A, BBS9 3′ UTR+115C>T and c.1284C>T) were not found in any databases searched (1000 Genomes, HapMap, NCBI or MapBack). All six variants were indicated as nondamaging or benign by SIFT or PPH.

ClustalW (http://www.ebi.ac.uk/Tools/msa/clustalw2/) alignment of the BBS9 protein sequence for conservation analysis was performed in 12 available species: NP_001028776.1 (*Homo sapiens*), NP_848502.1 (*Mus musculus*), XP_002664838.1 (*Danio rerio*), XP_235942.4 (*Rattus norvegicus*), NP_001179782.1 (*Bos taurus*), XP_418843.3 (*Gallus gallus*), NP_001016821.1 (*Xenopus tropicalis*), XP_002957514.1 (*Volvox carteri f. nagariensis*), XP_001107028.1 (*Macaca mulatta*), XP_852259.1 (*Canis lupis familiaris),* XP_001168457.1 (*Pan troglodytes*) and NP_001087412.1 (*Xenopus laevis*). Parental DNA was not available for sequencing analysis.

Imputation. Ungenotyped markers 50 kb upstream and downstream from borderline-significant regions from the TDT GWAS (chromosomes 3, 7, 8, 15 and 22) were imputed for the discovery sample using a trio design as implemented in BEAGLE v3.3.2 (ref. 26). A reference panel of 379 European individuals (GBR, FIN, IBS, CEU and TWI) from the 1000 Genomes Project database was used. We first prephased the trio GWAS data and then imputed SNPs using a phased reference panel. There were 58, 19, 75, 37 and 30 genotyped SNPs and 1,074, 234, 613, 276 and 431 imputed variants on chromosomes 3, 7, 8, 15 and 22, respectively. We carried out TDT tests on imputed variants of high quality ($r^2 \geq 0.3$).

Statistical analyses. GWAS association tests were carried out using TDT analysis as implemented in PLINK v1.07 (ref. 53). Manhattan plots were generated using the R script, and LD plots and r^2 measures were obtained using HaploView v4.06 (ref. 55). All SNP positioning information was gathered from human genome version hg18 (March 2006; NCBI36).

Association mapping in the replication panel was performed by the χ^2 test for allelic association using SAS/Genetics software version 9.2 (SAS Institute, Inc., Cary, NC). The direction of effect of markers surpassing nominal significance in the replication dataset was compared between the discovery and replication datasets, and three markers had opposite effects (chromosome 6 rs4716412, chromosome 7 rs7457112 and chromosome 16 rs4522429).

We searched for two-way gene-gene interactions using the 100 most significant SNPs from the 130 NHW trio TDT analysis. We performed a genotypic TDT for the interaction of each pair of SNPs, excluding results from SNP pairs on the same chromosome because of long-range LD (1,247 tests). For each case-parents trio, we generated 15 pseudo controls matched to each proband as a function of the parental genotypes of the two considered SNPs. 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, we fitted a conditional logistic regression model[25] to test for epistatic interactions that incorporate additive and dominance effects at the two loci. The 'trio' R package v1.4.24 was used to perform these tests[56].

Conditional analysis was performed to identify additional SNPs potentially masked by the primary significant SNP using the conditional extended TDT for trio data[57] as implemented in UNPHASED v3.0.9 (ref. 24). A meta-analysis was performed by combining the trait-marker associations of the discovery TDT and replication case-control study using a fixed-effects model after testing for interstudy heterogeneity[58].

53. Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).
54. Mills, J.L. *et al.* Folate and vitamin B12–related genes and risk for omphalocele. *Hum. Genet.* **131**, 739–746 (2012).
55. Barrett, J.C. *et al.* Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics* **21**, 263–265 (2005).
56. Schwender, H. *et al.* Rapid testing of SNPs and gene-environment interactions in case-parent trio data based on exact analytic parameter estimation. *Biometrics* **68**, 766–773 (2012).
57. Koeleman, B.P. *et al.* Adaptation of the extended transmission/disequilibrium test to distinguish disease associations of multiple loci: the conditional extended transmission/disequilibrium test. *Ann. Hum. Genet.* **64**, 207–213 (2000).
58. Kazeem, G.R. & Farrall, M. Integrating case-control and TDT studies. *Ann. Hum. Genet.* **69**, 329–335 (2005).

doi:10.1038/ng.2463