ANALYSIS OF WHOLE GENOME SEQUENCING IN A COHORT OF INDIVIDUALS WITH PHACE SYNDROME
SUGGESTS DYSREGULATION OF RAS/PI3K SIGNALING

Short title: WHOLE GENOME SEQUENCING ANALYSIS IN PHACE SYNDROME

Authors: Elizabeth S. Partan (1,2), Francine Blei (3), Sarah L. Chamlin (4,5,6), Olivia M. T. Davies (7), Beth A. Drolet (8), Ilona J. Frieden (9,10), Ioannis Karakikes (11), Chien-Wei Lin (12), Anthony J. Mancini (4,5,6), Denise Metry (13), Anthony Oro (14), Nicole S. Stefanko (4), Laksshman Sundaram (15), Monika Tutaj (16), Alexander E. Urban (17,18), Kevin C. Wang (14, 19), Xiaowei Zhu (17,18), Nara Sobreira* (1, 20), Dawn H. Siegel* (21)

Affiliations

1. McKusick-Nathans Department of Genetic Medicine, Johns Hopkins Medicine, Baltimore, MD
2. Predoctoral Training Program in Human Genetics, Johns Hopkins School of Medicine, Baltimore, MD
3. Department of Pediatrics / Vascular Anomalies Program, NYU Langone Health, Grossman School of Medicine, New York, NY
4. Department of Dermatology, Northwestern University Feinberg School of Medicine, Chicago, IL
5. Department of Pediatrics, Northwestern University Feinberg School of Medicine, Chicago, IL
6. Division of Dermatology, Ann & Robert H. Lurie Children’s Hospital of Chicago, Chicago, IL
7. Medical College of Wisconsin, Milwaukee, WI
8. Department of Dermatology, University of Wisconsin School of Medicine and Public Health, Madison, WI
9. Department of Dermatology, University of California San Francisco, San Francisco, CA
10. Department of Pediatrics, University of California San Francisco, San Francisco, CA
11. Department of Cardiothoracic Surgery, Stanford University, Palo Alto, CA
12. Division of Biostatistics, Institute for Health and Equity, Medical College of Wisconsin, Milwaukee, WI
13. Department of Dermatology, Baylor College of Medicine, Houston, TX
14. Department of Dermatology, Stanford University, Palo Alto, CA
15. Department of Computer Science, Stanford School of Engineering, Palo Alto, CA
16. Department of Biomedical Engineering, Medical College of Wisconsin, Milwaukee, WI
17. Department of Psychiatry and Behavioral Sciences, Stanford University, Palo Alto, CA
18. Department of Genetics, Stanford University, Palo Alto, CA
19. Veterans Affairs Healthcare System, Palo Alto, CA
20. Department of Pediatrics, Johns Hopkins Medicine, Baltimore, MD
21. Department of Dermatology, Medical College of Wisconsin, Milwaukee, WI

* These authors contributed equally to this work
Work was performed in Baltimore, MD, USA; Milwaukee, WI, USA; and Palo Alto, CA, USA

Corresponding author: Dawn Siegel • 414-955-2818 • dsiegel4@stanford.edu

Abbreviations used: single nucleotide variant (SNV), copy number variant (CNV), extracellular matrix (ECM)
ABSTRACT

The acronym PHACE stands for the co-occurrence of posterior brain fossa malformations, hemangiomas, arterial anomalies, cardiac defects, and eye abnormalities. The majority of patients have a segmental hemangioma and at least one developmental structural anomaly. The etiology and pathogenesis are unknown. Here we discuss the candidate causative genes identified in a de novo analysis of whole genome sequencing of germline samples from 98 unrelated trios in which the probands had PHACE, all sequenced as part of the Gabriella Miller Kids First Pediatric Research Program. A g:Profiler pathway analysis of the genes with rare, de novo variants suggested dysregulation of the RAS/MAPK and PI3K/AKT pathways that regulate cell growth, migration, and angiogenesis. These findings, along with the developmental anomalies and the vascular birthmark, support including PHACE within the RASopathy family of syndromes.

INTRODUCTION

The RASopathies represent a collection of rare genetic conditions caused by mutations in individual genes within the RAS/MAP kinase (MAPK) pathway. RAS/MAPK signaling is upstream of many cellular pathways; therefore, depending on the individual mutation and affected tissue, single-gene perturbations can cause numerous syndromes, including Cardio-Facio-Cutaneous (CFC), Costello (CS), Legius (LS), Neurofibromatosis type 1 (NF1), Noonan (NS), Noonan-like (NS with Multiple Lentigines, NSML; NS with Loose Anagen Hair, NSLH) and Capillary Malformation-Arteriovenous Malformation (CM-AVM) (Grant et al. 2018; Stevenson et al. 2016). These syndromes share many clinical features including distinct facial features, developmental delays,
cardiac defects, growth delays, and structural brain anomalies. While these individual syndromes are rare, as a group the RASopathies are among the most common genetic conditions in the world. The vast majority of mosaic neurocutaneous syndromes have been shown to be caused by somatic variants in RAS pathway genes, including the association of HRAS and KRAS variants with sebaceous nevus syndrome and NRAS variants with neurocutaneous melanosis (Chacon-Camacho et al. 2019; Kinsler et al. 2013).

The PHACE (OMIM 606519) acronym represents the sporadic vascular syndrome exhibiting the co-occurrence of posterior brain fossa malformations, hemangiomas, arterial anomalies, cardiac defects, and eye abnormalities (Frieden et al. 1996). The majority of patients have a segmental infantile hemangioma and at least one other feature (Garzon et al. 2016; Metry et al. 2009). Dysplasia and stenosis of the arteries of the head and neck is the second most common feature (occurring in >85%), followed by structural brain anomalies (including hypoplastic cerebellum), cardiovascular anomalies (including aortic coarctation or ventricular septal defect), and eye anomalies (Metry et al. 2009; Siegel 2017). The significant overlap in developmental anomaly phenotypes between PHACE and RASopathy patients suggests a common mechanistic underpinning of sporadic PHACE and germline RASopathy variants (Ruggieri et al. 2020).

The etiology and pathogenesis of PHACE are unknown, although consideration of the variety of body systems affected suggests that the age of onset is between 3 and 12 weeks of gestation during early vasculogenesis (Frieden et al. 1996; Metry et al. 2006; Siegel 2017). There is a 6:1 female:male preponderance, but no significant difference in the severity of the phenotype between males and females (Metry et al. 2009; Wan et al. 2017). PHACE is a sporadic disorder; there are no reported familial cases and at least four female patients with PHACE have given birth to clinically unaffected children (Martel et al. 2015; Stefanko et al. 2019). PHACE
is believed to have a genetic cause because there is no evidence of a common exposure during gestation or early infancy for children diagnosed with PHACE (Wan et al. 2017). Autosomal dominant inheritance caused by \textit{de novo}, germline pathogenic variants or mosaicism caused by somatic mosaic pathogenic variants have been suggested as possible modes of inheritance (Mitchell et al. 2012; Siegel 2017).

Here we present the \textit{de novo} analysis of whole genome sequencing (WGS) of germline samples from 98 unrelated trios in which the probands had PHACE, all sequenced as part of the Gabriella Miller Kids First project (X01HL140519-01). Additionally, we analyzed the DECIPHER database (Firth et al. 2009) for genes affected by copy number variants (CNVs) identified in individuals with the search terms “hemangioma” and “facial hemangioma” and used this analysis to prioritize genes with \textit{de novo} variants identified among our cohort of 98 patients. Among the rare, \textit{de novo} variants, we prioritized variants in 10 genes in 18 patients; six of these genes were in the RAS pathway. Our results support including PHACE within the RASopathy family of syndromes.

\section*{RESULTS}

\textbf{Germline whole genome sequencing and \textit{de novo} analysis pipeline}

We identified a total of 16,107 \textit{de novo} single nucleotide variants (SNVs) in our cohort of 98 patients; on average, this is 164 SNVs per patient. Our analysis found 15,752 rare (minor allele frequency [MAF] < 1\%), \textit{de novo} variants, including 219 coding variants (119 missense/nonsense/stoploss/splicing and 100 synonymous) and 15,533 noncoding variants. No \textit{de novo} variant was found in more than two patients, and no gene had nonsynonymous, \textit{de novo} coding variants in more than one patient.
**Variant prioritization**

Among the rare, *de novo* variants, we prioritized 18 candidate SNVs in 10 genes in 18 patients, and two indels in these same genes (Supplemental Table 1; see Methods for a description of prioritization metrics). Briefly, variants were prioritized if they were rare or novel (population frequency < 1% or 0), affected a gene known to cause a phenotype similar to PHACE in humans or mice, or affected genes disrupted by CNVs in a previous study of PHACE (Siegel et al. 2013).

**Genes associated with vascular human phenotypes**

Using OMIM annotation, we identified a missense, heterozygous *STAMBP*-p.Ile301Asn variant in Patient 16. Missense and frameshift homozygous or compound heterozygous variants in this gene cause autosomal recessive microcephaly-capillary malformation syndrome (OMIM 614261), which is characterized by capillary malformation, microcephaly, optic atrophy, cleft palate, tapered fingers, and seizures. *STAMBP*-p.Ile301Asn is not found in gnomAD and is predicted to be pathogenic by 18 *in silico* algorithms. Additionally, 88% of non-VUS missense variants in this gene are pathogenic. Patient 16 was heterozygous for the *STAMBP*-p.Ile301Asn variant and a second candidate variant in this gene was not identified despite careful analysis of the WGS data including a search for deletions or duplications.

**Genes associated with vascular mouse phenotypes**

We then identified coding and noncoding variants in six genes associated with lethal knockout mouse models due to vascular abnormalities. Two variants, *RASA3*-p.Val85Met in Patient 4 and *THBS2*-p.Asp859Asn in Patient 14, were exonic. The *RASA3*-p.Val85Met variant is predicted to be pathogenic by 11 algorithms (CADD, DEOGEN2, EIGEN, FATHMM-MKL, FATHMM-XF,
LRT, Mutation assessor, MutationTaster, PrimateAI, SIFT, and SIFT4G) and to affect splicing. The THBS2–p.Asp859Asn variant is found in the C-terminal type 3 repeats and likely disrupts intraprotein interactions and calcium binding (Figure 1); this variant is predicted to be pathogenic by 19 algorithms (BayesDel-addAF, BayesDel-noAF, CADD, DEOGEN2, EIGEN, EIGEN PC, FATHMM, FATHMM-MKL, FATHMM-XF, LIST-S2, MVP, MetaLR, MetaSVM, Mutation assessor, MutationTaster, PROVEAN, REVEL, SIFT, and SIFT4G). The other variants were noncoding and included two intronic SNVs in BCAS3, three intronic SNVs in DLC1, one intronic SNV in GLRX3, one intronic SNV in PIK3CA, and one intronic SNV in THBS2 (see Supplemental Table 1). Most of the identified variants are novel. Variants in BCAS3, DLC1, GLRX3, and PIK3CA are predicted to affect splicing. Variants in BCAS3, DLC1, PIK3CA, and THBS2 are predicted to be pathogenic by at least one algorithm. The BCAS3 SNV in intron 23 is predicted to fall within a binding site for GATA1 and the THBS2 intronic SNV is predicted to fall in a binding site for IKZF1. All three variants in DLC1 and the variants in GLRX3 and PIK3CA are predicted to be in open chromatin in human vascular endothelial cells (HUVECs) by Encode and therefore accessible to proteins regulating gene expression: four are predicted to be transcribed, and one (DLC1 intron 5) is predicted to be in an enhancer.

**Genes previously implicated in CNV analysis**

We next looked for variants in genes affected by CNVs identified in a cohort of patients with PHACE described by Siegel et al. (2013) among our cohort. This analysis identified the aforementioned intronic variant in PIK3CA again. We also prioritized two intronic variants in AFF2, an intronic variant in EPHA3, and four intronic variants in EXOC4. All four genes were affected by a CNV in a single patient in the prior study. The EPHA3 intronic variant is predicted to be pathogenic by EIGEN (0.58), to affect splicing,
and to fall within binding sites for CTCF, TCF12, and RAD21. All four EXOC4 variants are predicted to affect splicing, and the EXOC4 variant in intron 17 is predicted to be pathogenic by EIGEN (0.05) and GWAVA (0.5) and to fall within binding sites for GATA2 and GATA3.

**Indels in prioritized genes**

Finally, we investigated our cohort for rare, *de novo* indels affecting any of the 12 prioritized genes and identified two more variants: a 2-bp intronic deletion in BCAS3 in Patient 20 and an intronic 8-bp insertion in THBS2 in Patient 12, just upstream of the previously identified intronic SNV in THBS2 in the same patient.

**DECIPHER analyses**

Next, we analyzed published CNVs using DECIPHER to identify previously reported individuals with a phenotypic description containing the term “hemangioma” and CNVs containing our prioritized genes. This analysis identified five individuals in DECIPHER with deletions affecting THBS2; a single individual each with a deletion affecting BCAS3, GLRX3, and RASA3; and a single individual each with a duplication affecting DLC1 and PIK3CA.

We expanded our DECIPHER analysis to identify individuals with “facial hemangioma” and the genes disrupted by their CNVs. Next, we searched for rare, *de novo* variants in our cohort in the genes disrupted in two or more DECIPHER individuals. THBS2 was only disrupted by deletions in DECIPHER. Another two genes were each disrupted by two CNVs, one deletion and one duplication, in DECIPHER individuals with facial hemangioma. ALDH3A1 has a novel p.Val373Met variant in our Patient 8; this SNV is in the first codon of exon eight, is predicted to affect splicing by Human Splicing Finder (HSF), and is predicted to be
pathogenic by CADD (23.3) and FATHMM-MKL (0.48). ATP7B has a p.Thr482Ala variant in our Patient 10; this SNV is predicted to affect splicing by HSF.

We performed an enrichment analysis for the DECIPHER CNVs disrupting ALDH3A1, ATP7B, and THBS2. Only deletions disrupting THBS2 were significantly more common among individuals with facial hemangioma (FET, p=3.7e-2, OR=7.17) or hemangioma (FET, p=3.9e-3, OR=5.24) in DECIPHER than among the individuals without these features in DECIPHER, respectively.

**Pathway analysis**

Genes containing rare, de novo, non-intergenic SNVs (n=4,320 genes) were analyzed through g:Profiler to look for common pathways disrupted in our cohort (Table 1). The five gene ontology (GO) terms with the most significant p-values were all related to development and morphogenesis. The most significant term, nervous system development (adjusted p=6.22e-24), included our candidate genes AFF2, DLC1, EPHA3, PIK3CA, and THBS2. We also observed an enrichment for cell adhesion (adjusted p=1.34e-10, including BCAS3, DLC1, EPHA3, PIK3CA, and THBS2), which is critical for angiogenesis. The three KEGG terms with the most significant p-values implicated RAP signaling (adjusted p=1.61e-6, including PIK3CA), focal adhesion (adjusted p=2.43e-6, including PIK3CA and THBS2), and phospholipase D signaling (adjusted p=8.13e-6, including PIK3CA).

**Expression analysis**

We explored the expression of the candidate genes RASA3, AFF2, DLC1, EPHA3, PIK3CA, and THBS2 across diverse cell types in the human developing heart (Figure 2) by incorporating chromatin accessibility (scATAC-seq) (Ameen & Sundaram et. al.,...
unpublished) and gene expression (scRNA-seq) datasets (Asp et al. 2019; Miao et al. 2020; Suryawanshi et al. 2020). We observed that AFF2, EPHA3, PIK3CA, and THBS2 are co-expressed in the vasculature (pre-smooth muscle cells and vascular smooth muscle cells) in the fetal heart, whilst AFF2, EPHA3, and PIK3CA are co-expressed in the endothelium. Furthermore, RASA3, DLC1, and THBS2 are co-expressed in pericytes, the mural cells surrounding blood vessels and embedded within the basement membrane of the vasculature and adjacent to endothelial cells (Armulik et al. 2005), and RASA3 and THBS2 are co-expressed in perivascular fibroblasts, which may serve as pericyte precursors (Rajan et al. 2020). Together, these data suggest that the expression of the candidate genes is enriched in endothelial cells and mural cells (pericytes and vascular smooth muscle cells) in the blood vessel wall.

DISCUSSION

Our whole genome sequencing analyses support a disruption in RAS pathway signaling as the primary pathogenic mechanism behind PHACE. The RAS pathway genes identified in our analysis are associated with abnormal angiogenesis, including changes to cytoskeletal organization, cell migration, and adhesion, in relevant knockout mouse models. DLC1 binds RASA1 (a paralog of RASA3) at focal adhesions, and Dlc1 knockout causes embryonic lethal placenta vascular malformations (Durkin et al. 2005; Yang et al. 2009). EPHA3 regulates cytoskeletal organization, cell-cell adhesion, and cell migration in angiogenesis, and Epha3 knockout causes perinatal lethal cardiac defects (Clifford et al. 2008; Vaidya et al. 2003; Vearing et al. 2005). PIK3CA regulates cell migration and adhesion and has been associated with vascular anomalies, and Pik3ca knockout causes embryonic lethal vascular defects and hemorrhage (Bi et al. 1999; Graupera et al. 2008; Lelievre et al. 2005). RASA3 regulates cell-cell and cell-matrix adhesion and cell
migration, and Rasa3 knockout causes embryonic lethal hemorrhage following vascular lumenization defects (Iwashita et al. 2007; Molina-Ortiz et al. 2018). THBS2 binds heparin in the extracellular matrix (ECM), mediates cell-matrix interactions, and inhibits cell adhesion and angiogenesis, and Thbs2 knockout causes increased vascularity and premature death (Bornstein et al. 2000; Kyriakides et al. 1998; Noh et al. 2003; Streit et al. 1999; Yang et al. 2000).

The g:Profiler pathway analysis of the genes with rare, de novo SNVs in these patients also implicated these same candidate genes. KEGG pathway analysis identified RAP signaling, focal adhesion, and phospholipase D (PLD) signaling as the most significantly affected pathways. RASA3 downregulates RAP1 GTPase signaling (Battram et al. 2017; Molina-Ortiz et al. 2018). BCAS3, DLC1, PIK3CA, and THBS2 all participate in the regulation of focal adhesion (Durkin et al. 2005; Elzie and Murphy-Ullrich 2004; Jain et al. 2012; Matsuoka et al. 2012; Yang et al. 2009). PLD is required for actin cytoskeleton remodeling in cell migration and angiogenesis, participates in RAS activation, and is activated by both RAS/MAPK/ERK and PI3K/AKT/mTOR signaling (Bruntz et al. 2014). RASA3 downregulates RAS signaling through MEK/ERK and also binds PI3K and regulates cell migration though PI3K-dependent integrin signaling (Battram et al. 2017; Castellano and Downward 2011). PIK3CA forms part of the PI3K complex and regulates endothelial cell migration during vascular development (Graupera et al. 2008). STAMBP binds GRB2 to regulate both the RAS/MAPK and PI3K/AKT/mTOR pathways (McDonell et al. 2013; Tsang et al. 2006). THBS2 signals through FAK activation of the MAPK/ERK and PI3K/AKT pathways to stimulate actin reorganization, decrease cell adhesion, and increase migration (Goicoechea et al. 2000; Greenwood et al. 1998; Orr et al. 2002). DLC1 and EPHA3 both bind RASA1, a close paralog of RASA3 (Haupaix et al. 2013; Yang et al. 2009); the EPHA3 intronic variant in particular is also in a binding site for TCF12, which
downregulates ERK signaling (Yi et al. 2017). DLC1 is also phosphorylated by AKT, which abrogates its tumor suppression activity (Ko et al. 2010). The identification of variants in these genes among the probands with PHACE suggests that defects in matricellular signaling affecting endothelial cell adhesion and migration through the RAS and PI3K pathways may contribute to the pathogenesis of PHACE (Figure 3).

Our analysis of rare, de novo germline variants combined with the DECIPHER analysis of CNVs found in individuals with hemangioma identified THBS2 as a strong candidate causative gene for PHACE. THBS2 had missense and intronic variants in our cohort and was deleted in five individuals with hemangioma in DECIPHER; an enrichment analysis showed that deletions disrupting THBS2 were significantly more common among individuals with hemangioma than in individuals without this feature in DECIPHER. While no human phenotype has been associated with this gene, knockout of this gene in a mouse causes blood vessel overgrowth and premature death (Bornstein et al. 2000; Kyriakides et al. 1998; Yang et al. 2000). THBS2 is highly expressed in arterial tissues (aorta, coronary, and tibial) described in GTEx and in the developing vasculature in our scRNA analysis. THBS2 interacts with the ECM, mediates cell-matrix interactions, inhibits cell adhesion, inhibits angiogenesis, and acts as a tumor suppressor (Bornstein et al. 2000; Kyriakides et al. 1998; Noh et al. 2003; Streit et al. 1999; Yang et al. 2000). The THBS2 residue p.Asp859, which had a missense variant in our cohort, is found in the C-terminal type 3 repeats and participates in calcium binding (Figure 1). This calcium binding is necessary for correct folding of the C terminus, which itself must correctly trimerize in order to support the integrin-binding and cell motility functions of the THBS2 protein (Anilkumar et al. 2002; Carlson et al. 2008; Kvansakul et al. 2004). Therefore, a variant in this position could affect the protein function by disrupting protein folding and trimerization.
Our analysis identified RASA3 as a second strong candidate causative gene for PHACE. RASA3 had a missense variant in our cohort and was deleted in one DECIPHER individual with capillary hemangioma. Germline pathogenic variants in a related gene, RASA1, are known to cause capillary malformation–arteriovenous malformation (Eerola et al. 2003). No human phenotype has been associated with RASA3, but knockout of this gene in a mouse causes hemorrhage and embryonic lethality due to vascular abnormalities (Iwashita et al. 2007; Molina-Ortiz et al. 2018). RASA3 also regulates angiogenesis via the RAS/RAP pathway (Molina-Ortiz et al. 2018), which was implicated in our pathway analysis, and was highly expressed in pericytes in our scRNA analysis.

We also identified seven noncoding variants in four RAS pathway genes (DLC1, EPHA3, PIK3CA, and THBS2). These variants were predicted to affect splicing, functional DNA, transcription factor binding, and/or open chromatin in vascular cells (Table 2). The genes DLC1 and PIK3CA are highly intolerant to loss-of-function coding mutations (probability of loss-of-function intolerance scores > 0.9) (Lek et al. 2016) and have enhanced expression in PHACE-relevant tissues: DLC1 in neurons and retinal amacrine cells and PIK3CA in aorta and coronary artery (Genotype-Tissue Expression (GTEx) Project 2021). Single cell RNA expression analyses also showed that our candidate RAS pathway genes have high expression in neural crest, smooth muscle, and endothelial cells in the developing heart (Figure 2). The predicted disruption of expression and/or splicing of these genes by our identified noncoding variants may therefore contribute to other PHACE features. For instance, an intronic SNV, rs1905014-T, in DLC1 is significantly associated with optic disc size (Han et al. 2019). However, the effects of the noncoding variants on splicing and/or functional DNA need to be investigated further.
In conclusion, we have found strong candidate coding and noncoding variants in genes in the RAS/MAPK and PI3K/AKT pathways in individual patients with PHACE. These variants, paired with the vascular birthmark (infantile hemangioma) and the structural brain, heart, arterial, and eye anomalies in PHACE, support the inclusion of PHACE as a RASopathy syndrome. It is striking that the majority of these variants were predicted to have effects that may lead to altered gene expression during development. However, because of the complex cellular interactions between vascular cell types, further \textit{in vivo} studies are necessary to confirm pathogenicity of the variants identified and to understand their effect on the pathways that are likely disrupted by these variants.

\textbf{SUBJECTS AND METHODS}

\textbf{Subjects}

Patients with PHACE and their parents were recruited and sequenced as part of the Gabriella Miller Kids First Pediatric Research Program (X01HL140519-01). The study is approved by the Medical College of Wisconsin IRB #PRO00034867 and patients gave written, informed consent. Of the 98 patients discussed here, 19 were male, and 85 were white (Table 3). The average age of the cohort was 5.8 years at the time of enrollment. Ninety patients had facial hemangioma.

\textbf{Germline whole genome sequencing and de novo analysis pipeline}

WGS was performed on germline DNA from 98 unrelated patients with PHACE and their parents. WGS libraries were sequenced on the Illumina NovaSeq 6000 platform using 150-bp paired ends and the NovaSeq 6000 S4 Reagent Kit (Illumina, San Diego, California). Alignment was performed using DRAGEN (v.SW: 01.011.269.3.2.8) (Ma et al. 2015); data conversion from...
CRAM to BAM files with samtools (v.1.9); data pre-processing with Picard RevertSam, MarkDuplicates, SortSam, and MergeBamAlignment (v2.18.2) (Broad Institute 2019) and base quality recalibration with GATK BQSR (v.4.0.3.0). Germline variant calling and single-sample gVCF generation were performed using GATK HaplotypeCaller (v.3.5-0-g36282e4). Genotype analysis of trios/families, de novo variant discovery, variant quality recalibration, joint variant calling on all samples, and recalculation of genotype call accuracy on the population level were performed with GATK tools (Van der Auwera et al. 2013): HaplotypeCaller (v.3.5-0-g36282e4), VariantAnnotator (v.3.8-0-ge9d806836), VQSR, CalculateGenotypePosteriors, and VariantFiltration (v.4.0.12.0).

All outputs were released to the DRC Portal and dbGaP, where the files can be accessed and downloaded (github.com/kids-first/kf-alignment-workflow).

High confidence de novo SNVs and indels were called for variants present only in the proband with GQ scores ≥ 20 for all trio members. Resulting variants were lifted to GRCh37/hg19 using UCSC’s liftOver tool and annotated using Annovar (v2013_09_11) (Wang et al. 2010). The liftover was necessary because the TraP and GWAVA annotations (see Variant prioritization) are not available under hg38.

**Variant prioritization**

Variants were subset to include SNVs with gnomAD (Karczewski et al. 2020) minor allele frequency (MAF) < 1% (hereafter defined as rare). Next, we prioritized SNVs that:

1) were not present in gnomAD (novel);
2) affected a gene already known to cause a phenotype overlapping that of PHACE in OMIM (omim.org) or GWAS annotation (Buniello et al. 2019);

3) affected a gene where the mouse model phenotype (informatics.jax.org) overlapped that of PHACE;

4) affected genes disrupted by CNVs described by Siegel et al. (2013); or

5) affected genes with coding variants in two or more patients.

Prioritized coding variants were further evaluated for pathogenicity prediction scores using VarSome (Kopanos et al. 2019).

Prioritized noncoding variants were further evaluated for scores designed to predict effects on splicing [TraP (Gelfman et al. 2017), Human Splicing Finder (HSF) (Desmet et al. 2009)]; functional DNA [CADD (Rentzsch et al. 2019), DANN (Quang et al. 2015), GWAVA (Ritchie et al. 2014), EIGEN (Ionita-Laza et al. 2016), FATHMM (Shihab et al. 2015)]; transcription factor binding sites (TFBS) (Euskirchen et al. 2007); and open chromatin in vascular cells in Encode (Ernst et al. 2011).

Finally, we investigated the de novo indels identified in the genes prioritized using metrics 1-5.

**DECIPHER analyses**

We first selected the pathogenic or likely pathogenic deletions and duplications described in DECIPHER (Firth et al. 2009) in individuals with hemangioma. We compared the genes affected by these CNVs with the genes prioritized in our cohort (described above) to identify other individuals in DECIPHER with hemangiomas and CNVs affecting our prioritized genes.

We then selected all pathogenic or likely pathogenic CNVs described in DECIPHER in individuals with facial hemangioma. The “facial hemangioma” feature was chosen to represent PHACE because it is the most common manifestation of this syndrome. We
identified 47 DECIPHER individuals with facial hemangioma, with a total of 55 CNVs (17 gains, 38 losses) disrupting 5,703 genes; 1,191 of these genes were disrupted in at least two DECIPHER individuals. This subset of 1,191 genes was compared to the list of rare, de novo variants identified in our cohort of patients with PHACE to prioritize further candidates.

An enrichment analysis was performed for each gene identified by the second DECIPHER analysis. The DECIPHER pathogenic or likely pathogenic deletions disrupting the gene of interest in individuals with facial hemangioma were classified as “cases”. The DECIPHER pathogenic or likely pathogenic deletions disrupting the same gene in individuals without facial hemangioma were classified as “controls”. For each gene we built a contingency table comparing the number of cases to controls and performed a Fisher’s exact test (FET) using a cutoff of p-value < 0.05 for statistical significance. For significant genes we also calculated an odds ratio.

**Sanger sequencing**

Prioritized candidate variants were validated by Sanger sequencing. DNA from the proband and each parent was amplified using Accuprime Taq polymerase (Invitrogen; 12339-016) using standard thermal cycling conditions with annealing at 60°C and variant-specific primer sequences (Supplemental Table 2). Small aliquots of each PCR product were run on a 1% agarose gel to verify amplification, and the remainder of each PCR product was cleaned up using a mixture of exonuclease 1 (Affymetrix; 70073X) and shrimp alkaline phosphatase (Affymetrix; 78390) and sequenced using the same primers.
Protein structure analysis

The experimentally derived structure of THBS2 was obtained from RCSB PDB (Berman et al. 2000) structure 1YO8 (Carlson et al. 2005). The effect of the identified p.Asp859Asn variant in THBS2 was examined using DynaMut (Rodrigues et al. 2018), which predicts the effects of missense mutations on known PDB structures.

Pathway analysis

All 4,320 genes with rare, de novo SNVs were analyzed with gProfiler (Raudvere et al. 2019) to determine if they had similar functions or affected the same pathways. Multiple testing-adjusted p-values (using the default g:SCS method in gProfiler) were reported, and 0.05 was used as the cutoff for statistical significance.

Expression analysis

The chromatin accessibility atlas and gene expression data of the human developing heart was derived from nuclei isolated from human fetal hearts (Ameen & Sundaram et al., unpublished). Briefly, the Chromium 10X platform (Satpathy et al. 2019) was used to generate scATAC-seq data from three primary human fetal heart samples at 6, 8, and 19 weeks post-gestation. The scRNA-seq data from fetal hearts were analyzed from published studies (Asp et al. 2019; Miao et al. 2020; Suryawanshi et al. 2020) using the batch correction algorithm Harmony (Korsunsky et al. 2019), and then annotated using known cell-type-specific marker genes.

DATA AVAILABILITY
Datasets related to this article are accessible through the Kids First Data Resource Portal (kidsfirstdrc.org) and/or dbGaP (ncbi.nlm.nih.gov/gap) accession phs001785.v1.p1.

**ORCIDS**

Elizabeth S. Partan: 0000-0002-3995-6742

Francine Blei: 0000-0001-6726-2361

Sarah L. Chamlin: 0000-0002-0845-2306

Olivia M. T. Davies: 0000-0001-5428-7279

Beth A. Drolet: 0000-0002-0844-7195

Ilona J. Frieden: 0000-0001-7305-5940

Ioannis Karakikes: 0000-0002-4348-600X

Chien-Wei Lin: 0000-0003-4023-7339

Anthony J. Mancini: 0000-0003-0903-7734

Denise Metry: 0000-0003-1991-1683

Anthony Oro: 0000-0002-6261-138X

Nicole S. Stefanko: 0000-0002-3609-3144

Lakshman Sundaram: 0000-0002-8953-3940

Monika Tutaj: 0000-0002-0378-4002
Alexander E. Urban: 0000-0001-9772-933X

Kevin C. Wang: 0000-0002-5305-6791

Xiaowei Zhu: 0000-0002-5731-3177

Nara Sobreira: 0000-0002-5228-5613

Dawn H. Siegel: 0000-0001-6546-2693

CONFLICT OF INTEREST

Francine Blei has received honoraria and non-drug related educational grants from Pierre Fabre. Sarah Chamlin is a speaker for Sanofi Genzyme and Regeneron Pharmaceuticals. Beth Drolet is a consultant and on the clinical advisory board for Venthera. Ilona Frieden serves on the Data Safety monitoring boards for Pfizer, is a consultant and advisory board member for Novartis, and is a consultant for Venthera. Dawn Siegel has received royalties from UpToDate, is a consultant for Arqule/Merck, and received a SID SUN mid-career investigator award.

ACKNOWLEDGMENTS

We thank all the families for their participation in this study, and the PHACE Syndrome Community and Pediatric Dermatology Research Alliance (PeDRA) for support of this project. We acknowledge the support of the Children’s Research Institute at Children’s Wisconsin and the Genomic Sciences and Precision Medicine Center at Medical College of Wisconsin.

The results analyzed and shown here are based in part upon data generated by Gabriella Miller Kids First Pediatric Research Program (Kids First) projects phs001785.v1.p1. Kids First was supported by the Common Fund of the Office of the Director of the
National Institutes of Health (commonfund.nih.gov/KidsFirst). The Hudson-Alpha Institute for Biotechnology was awarded a U24HD090744-01 to sequence structural birth defect cohort samples submitted by investigators through the Kids First program (1X01HL140519-01).

This study was supported by the NICHD, NHLBI, and NIAMS under award numbers 1X01HL14519, R03HD098526, and R01AR064258, respectively. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Additional support was received from a SID SUN grant, PHACE Foundation Canada, and the Children’s Research Institute.

This study makes use of data generated by the DECIPHER community. A full list of centers who contributed to the generation of the data is available from decipher.sanger.ac.uk and via email from decipher@sanger.ac.uk. Funding for the project was provided by Wellcome. The individuals who carried out the original analysis and collection of the data bear no responsibility for the further analysis or interpretation of it in this paper.

AUTHOR CONTRIBUTIONS (CREDIT)

Elizabeth S. Partan: Conceptualization, Data curation, Investigation, Visualization, Writing – original draft, Writing – review & editing

Francine Blei: Data curation, Writing – review & editing

Sarah L. Chamlin: Data curation, Writing – review & editing

Olivia M. T. Davies: Data curation, Writing – review & editing
Beth A. Drolet: Data curation, Writing – review & editing

Ilona J. Frieden: Data curation, Writing – review & editing

Ioannis Karakikes: Writing – review & editing

Chien-Wei Lin: Formal analysis, Writing – review & editing

Anthony J. Mancini: Data curation, Writing – review & editing

Denise Metry: Data curation, Writing – review & editing

Anthony Oro: Writing – review & editing

Nicole S. Stefanko: Data curation, Writing – review & editing

Lakshman Sundaram: Investigation, Visualization

Monika Tutaj: Data curation, Formal analysis, Writing – review & editing

Alexander E. Urban: Writing – review & editing

Kevin C. Wang: Writing – review & editing

Xiaowei Zhu: Writing – review & editing

Nara Sobreira: Conceptualization, Funding acquisition, Project administration, Supervision, Writing – review & editing

Dawn H. Siegel: Conceptualization, Data curation, Funding acquisition, Project administration, Supervision, Writing – review & editing
REFERENCES


Yi S, Yu M, Yang S, Miron RJ, Zhang Y. Tcf12, a member of basic helix-loop-helix transcription factors, mediates bone marrow mesenchymal stem cell osteogenic differentiation in vitro and in vivo. Stem Cells. 2017;35(2):386–97
**TABLES**

Table 1: g:Profiler pathway analysis of genes with rare, *de novo* variants

A) Gene ontology (GO) terms

<table>
<thead>
<tr>
<th>Term ID</th>
<th>Term name</th>
<th>Adjusted p-value</th>
<th># genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0007399</td>
<td>nervous system development</td>
<td>6.22E-24</td>
<td>646</td>
</tr>
<tr>
<td>GO:0009002</td>
<td>cell morphogenesis</td>
<td>1.88E-23</td>
<td>329</td>
</tr>
<tr>
<td>GO:0032989</td>
<td>cellular component morphogenesis</td>
<td>6.85E-20</td>
<td>260</td>
</tr>
<tr>
<td>GO:0048666</td>
<td>neuron development</td>
<td>8.51E-20</td>
<td>343</td>
</tr>
<tr>
<td>GO:0032990</td>
<td>cell part morphogenesis</td>
<td>1.15E-19</td>
<td>236</td>
</tr>
<tr>
<td>GO:0016043</td>
<td>cellular component organization</td>
<td>1.84E-19</td>
<td>1413</td>
</tr>
<tr>
<td>GO:0031175</td>
<td>neuron projection development</td>
<td>4.82E-19</td>
<td>309</td>
</tr>
<tr>
<td>GO:0030182</td>
<td>neuron differentiation</td>
<td>5.53E-19</td>
<td>401</td>
</tr>
<tr>
<td>GO:0120039</td>
<td>plasma membrane bounded cell projection morphogenesis</td>
<td>6.95E-19</td>
<td>228</td>
</tr>
<tr>
<td>GO:0120036</td>
<td>plasma membrane bounded cell projection organization</td>
<td>1.54E-18</td>
<td>428</td>
</tr>
<tr>
<td>GO:0048858</td>
<td>cell projection morphogenesis</td>
<td>1.60E-18</td>
<td>228</td>
</tr>
<tr>
<td>GO:0048812</td>
<td>neuron projection morphogenesis</td>
<td>2.63E-18</td>
<td>223</td>
</tr>
<tr>
<td>GO:0048699</td>
<td>generation of neurons</td>
<td>1.10E-17</td>
<td>433</td>
</tr>
<tr>
<td>GO:0030030</td>
<td>cell projection organization</td>
<td>2.90E-17</td>
<td>432</td>
</tr>
<tr>
<td>GO:0071840</td>
<td>cellular component organization or biogenesis</td>
<td>7.38E-17</td>
<td>1432</td>
</tr>
<tr>
<td>GO:0043167</td>
<td>ion binding</td>
<td>1.10E-16</td>
<td>1378</td>
</tr>
<tr>
<td>GO:0048468</td>
<td>cell development</td>
<td>1.23E-16</td>
<td>563</td>
</tr>
<tr>
<td>GO:0009004</td>
<td>cell morphogenesis involved in differentiation</td>
<td>1.31E-16</td>
<td>239</td>
</tr>
<tr>
<td>GO:0022008</td>
<td>neurogenesis</td>
<td>4.92E-16</td>
<td>449</td>
</tr>
<tr>
<td>GO:0048867</td>
<td>cell morphogenesis involved in neuron differentiation</td>
<td>3.46E-15</td>
<td>199</td>
</tr>
<tr>
<td>GO:0051128</td>
<td>regulation of cellular component organization</td>
<td>6.13E-15</td>
<td>599</td>
</tr>
<tr>
<td>GO:0019899</td>
<td>enzyme binding</td>
<td>9.62E-15</td>
<td>566</td>
</tr>
<tr>
<td>GO:0010975</td>
<td>regulation of neuron projection development</td>
<td>1.34E-14</td>
<td>176</td>
</tr>
<tr>
<td>GO:0034330</td>
<td>cell junction organization</td>
<td>4.11E-14</td>
<td>217</td>
</tr>
<tr>
<td>GO:0048856</td>
<td>anatomical structure development</td>
<td>5.63E-14</td>
<td>1312</td>
</tr>
<tr>
<td>GO:0051960</td>
<td>regulation of nervous system development</td>
<td>6.34E-14</td>
<td>280</td>
</tr>
<tr>
<td>GO:0120035</td>
<td>regulation of plasma membrane bounded cell projection organization</td>
<td>8.37E-14</td>
<td>217</td>
</tr>
<tr>
<td>GO:0031344</td>
<td>regulation of cell projection organization</td>
<td>2.28E-13</td>
<td>218</td>
</tr>
<tr>
<td>GO:0007275</td>
<td>multicellular organism development</td>
<td>4.63E-13</td>
<td>1211</td>
</tr>
<tr>
<td>GO:0045664</td>
<td>regulation of neuron differentiation</td>
<td>6.17E-13</td>
<td>211</td>
</tr>
<tr>
<td>GO:0022604</td>
<td>regulation of cell morphogenesis</td>
<td>6.51E-13</td>
<td>166</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-----------------</td>
<td>-------------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0009653</td>
<td>anatomical structure morphogenesis</td>
<td>7.53E-13</td>
<td>669</td>
</tr>
<tr>
<td>GO:0048731</td>
<td>system development</td>
<td>5.70E-12</td>
<td>1101</td>
</tr>
<tr>
<td>GO:0032502</td>
<td>developmental process</td>
<td>5.86E-12</td>
<td>1397</td>
</tr>
<tr>
<td>GO:0043168</td>
<td>anion binding</td>
<td>4.70E-11</td>
<td>666</td>
</tr>
<tr>
<td>GO:0007264</td>
<td>small GTPase mediated signal transduction</td>
<td>9.61E-11</td>
<td>162</td>
</tr>
<tr>
<td>GO:0050767</td>
<td>regulation of neurogenesis</td>
<td>1.28E-10</td>
<td>243</td>
</tr>
<tr>
<td>GO:0007155</td>
<td>cell adhesion</td>
<td>1.34E-10</td>
<td>375</td>
</tr>
<tr>
<td>GO:0034329</td>
<td>cell junction assembly</td>
<td>1.51E-10</td>
<td>142</td>
</tr>
<tr>
<td>GO:0022610</td>
<td>biological adhesion</td>
<td>1.70E-10</td>
<td>376</td>
</tr>
<tr>
<td>GO:0060284</td>
<td>regulation of cell development</td>
<td>2.17E-10</td>
<td>273</td>
</tr>
<tr>
<td>GO:0051130</td>
<td>positive regulation of cellular component organization</td>
<td>3.48E-10</td>
<td>320</td>
</tr>
<tr>
<td>GO:0016358</td>
<td>dendrite development</td>
<td>4.05E-10</td>
<td>94</td>
</tr>
<tr>
<td>GO:0023051</td>
<td>regulation of signaling</td>
<td>4.09E-10</td>
<td>820</td>
</tr>
<tr>
<td>GO:0046872</td>
<td>metal ion binding</td>
<td>1.06E-09</td>
<td>930</td>
</tr>
<tr>
<td>GO:0051056</td>
<td>regulation of small GTPase mediated signal transduction</td>
<td>1.09E-09</td>
<td>112</td>
</tr>
<tr>
<td>GO:0007010</td>
<td>cytoskeleton organization</td>
<td>1.28E-09</td>
<td>356</td>
</tr>
<tr>
<td>GO:0050839</td>
<td>cell adhesion molecule binding</td>
<td>1.78E-09</td>
<td>163</td>
</tr>
<tr>
<td>GO:0065008</td>
<td>regulation of biological quality</td>
<td>2.06E-09</td>
<td>899</td>
</tr>
<tr>
<td>GO:0030029</td>
<td>actin filament-based process</td>
<td>2.57E-09</td>
<td>226</td>
</tr>
<tr>
<td>GO:0010769</td>
<td>regulation of cell morphogenesis involved in differentiation</td>
<td>3.24E-09</td>
<td>108</td>
</tr>
<tr>
<td>GO:0010646</td>
<td>regulation of cell communication</td>
<td>3.33E-09</td>
<td>806</td>
</tr>
<tr>
<td>GO:0043169</td>
<td>cation binding</td>
<td>3.43E-09</td>
<td>943</td>
</tr>
<tr>
<td>GO:0050773</td>
<td>regulation of dendrite development</td>
<td>3.93E-09</td>
<td>66</td>
</tr>
<tr>
<td>GO:0050808</td>
<td>synapse organization</td>
<td>6.33E-09</td>
<td>133</td>
</tr>
<tr>
<td>GO:0007409</td>
<td>axonogenesis</td>
<td>6.46E-09</td>
<td>150</td>
</tr>
<tr>
<td>GO:0006996</td>
<td>organelle organization</td>
<td>9.75E-09</td>
<td>877</td>
</tr>
<tr>
<td>GO:0032559</td>
<td>adenyl ribonucleotide binding</td>
<td>1.99E-08</td>
<td>383</td>
</tr>
<tr>
<td>GO:0031346</td>
<td>positive regulation of cell projection organization</td>
<td>2.79E-08</td>
<td>127</td>
</tr>
<tr>
<td>GO:0030554</td>
<td>adenyl nucleotide binding</td>
<td>4.80E-08</td>
<td>383</td>
</tr>
<tr>
<td>GO:0051020</td>
<td>GTPase binding</td>
<td>5.13E-08</td>
<td>162</td>
</tr>
<tr>
<td>GO:0019904</td>
<td>protein domain specific binding</td>
<td>5.21E-08</td>
<td>198</td>
</tr>
<tr>
<td>GO:0044093</td>
<td>positive regulation of molecular function</td>
<td>5.34E-08</td>
<td>436</td>
</tr>
<tr>
<td>GO:0005524</td>
<td>ATP binding</td>
<td>6.16E-08</td>
<td>368</td>
</tr>
<tr>
<td>GO:0030036</td>
<td>actin cytoskeleton organization</td>
<td>6.68E-08</td>
<td>198</td>
</tr>
<tr>
<td>GO:0061564</td>
<td>axon development</td>
<td>8.00E-08</td>
<td>157</td>
</tr>
<tr>
<td>GO:0010976</td>
<td>positive regulation of neuron projection development</td>
<td>8.83E-08</td>
<td>100</td>
</tr>
<tr>
<td>GO:0051962</td>
<td>positive regulation of nervous system development</td>
<td>9.77E-08</td>
<td>165</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0099536</td>
<td>synaptic signaling</td>
<td>1.00E-07</td>
<td>200</td>
</tr>
<tr>
<td>GO:0007169</td>
<td>transmembrane receptor protein tyrosine kinase signaling pathway</td>
<td>2.31E-07</td>
<td>207</td>
</tr>
<tr>
<td>GO:0017016</td>
<td>Ras GTPase binding</td>
<td>2.65E-07</td>
<td>127</td>
</tr>
<tr>
<td>GO:0099537</td>
<td>trans-synaptic signaling</td>
<td>3.45E-07</td>
<td>193</td>
</tr>
<tr>
<td>GO:0097367</td>
<td>carbohydrate derivative binding</td>
<td>4.49E-07</td>
<td>521</td>
</tr>
<tr>
<td>GO:0051179</td>
<td>localization</td>
<td>4.95E-07</td>
<td>1402</td>
</tr>
<tr>
<td>GO:0031267</td>
<td>small GTPase binding</td>
<td>5.62E-07</td>
<td>129</td>
</tr>
<tr>
<td>GO:0005085</td>
<td>guanyl-nucleotide exchange factor activity</td>
<td>7.05E-07</td>
<td>78</td>
</tr>
<tr>
<td>GO:0008092</td>
<td>cytoskeletal protein binding</td>
<td>7.54E-07</td>
<td>252</td>
</tr>
<tr>
<td>GO:0035556</td>
<td>intracellular signal transduction</td>
<td>8.74E-07</td>
<td>632</td>
</tr>
<tr>
<td>GO:0022607</td>
<td>cellular component assembly</td>
<td>9.90E-07</td>
<td>658</td>
</tr>
<tr>
<td>GO:0032879</td>
<td>regulation of localization</td>
<td>1.04E-06</td>
<td>643</td>
</tr>
<tr>
<td>GO:0030165</td>
<td>PDZ domain binding</td>
<td>1.11E-06</td>
<td>40</td>
</tr>
<tr>
<td>GO:0045666</td>
<td>positive regulation of neuron differentiation</td>
<td>1.32E-06</td>
<td>120</td>
</tr>
<tr>
<td>GO:0043087</td>
<td>regulation of GTPase activity</td>
<td>1.85E-06</td>
<td>141</td>
</tr>
<tr>
<td>GO:0048522</td>
<td>positive regulation of cellular process</td>
<td>2.05E-06</td>
<td>1190</td>
</tr>
<tr>
<td>GO:0044087</td>
<td>regulation of cellular component biogenesis</td>
<td>2.11E-06</td>
<td>252</td>
</tr>
<tr>
<td>GO:0097485</td>
<td>neuron projection guidance</td>
<td>2.18E-06</td>
<td>94</td>
</tr>
<tr>
<td>GO:0098916</td>
<td>anterograde trans-synaptic signaling</td>
<td>2.20E-06</td>
<td>188</td>
</tr>
<tr>
<td>GO:0007268</td>
<td>chemical synaptic transmission</td>
<td>2.20E-06</td>
<td>188</td>
</tr>
<tr>
<td>GO:0007416</td>
<td>synapse assembly</td>
<td>2.75E-06</td>
<td>67</td>
</tr>
<tr>
<td>GO:0140096</td>
<td>catalytic activity, acting on a protein</td>
<td>2.88E-06</td>
<td>512</td>
</tr>
<tr>
<td>GO:0032553</td>
<td>ribonucleotide binding</td>
<td>3.02E-06</td>
<td>446</td>
</tr>
<tr>
<td>GO:0005088</td>
<td>Ras guanyl-nucleotide exchange factor activity</td>
<td>3.18E-06</td>
<td>48</td>
</tr>
<tr>
<td>GO:0032555</td>
<td>purine ribonucleotide binding</td>
<td>3.67E-06</td>
<td>442</td>
</tr>
<tr>
<td>GO:0007411</td>
<td>axon guidance</td>
<td>4.05E-06</td>
<td>93</td>
</tr>
<tr>
<td>GO:0043085</td>
<td>positive regulation of catalytic activity</td>
<td>4.06E-06</td>
<td>352</td>
</tr>
<tr>
<td>GO:0050804</td>
<td>modulation of chemical synaptic transmission</td>
<td>4.09E-06</td>
<td>125</td>
</tr>
<tr>
<td>GO:0005509</td>
<td>calcium ion binding</td>
<td>4.56E-06</td>
<td>191</td>
</tr>
<tr>
<td>GO:0099177</td>
<td>regulation of trans-synaptic signaling</td>
<td>4.80E-06</td>
<td>125</td>
</tr>
<tr>
<td>GO:0017076</td>
<td>purine nucleotide binding</td>
<td>5.98E-06</td>
<td>443</td>
</tr>
<tr>
<td>GO:0005515</td>
<td>protein binding</td>
<td>6.42E-06</td>
<td>2726</td>
</tr>
<tr>
<td>GO:0004672</td>
<td>protein kinase activity</td>
<td>7.71E-06</td>
<td>163</td>
</tr>
<tr>
<td>GO:0005543</td>
<td>phospholipid binding</td>
<td>8.46E-06</td>
<td>130</td>
</tr>
<tr>
<td>GO:0016773</td>
<td>phosphotransferase activity, alcohol group as acceptor</td>
<td>8.79E-06</td>
<td>186</td>
</tr>
<tr>
<td>GO:0048813</td>
<td>dendrite morphogenesis</td>
<td>1.23E-05</td>
<td>57</td>
</tr>
<tr>
<td>GO:0008047</td>
<td>enzyme activator activity</td>
<td>1.41E-05</td>
<td>146</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0009966</td>
<td>regulation of signal transduction</td>
<td>1.45E-05</td>
<td>697</td>
</tr>
<tr>
<td>GO:0006928</td>
<td>movement of cell or subcellular component</td>
<td>1.72E-05</td>
<td>514</td>
</tr>
<tr>
<td>GO:0035639</td>
<td>purine ribonucleoside triphosphate binding</td>
<td>1.78E-05</td>
<td>424</td>
</tr>
<tr>
<td>GO:0007417</td>
<td>central nervous system development</td>
<td>2.20E-05</td>
<td>263</td>
</tr>
<tr>
<td>GO:0003779</td>
<td>actin binding</td>
<td>2.23E-05</td>
<td>125</td>
</tr>
<tr>
<td>GO:0004674</td>
<td>protein serine/threonine kinase activity</td>
<td>2.32E-05</td>
<td>127</td>
</tr>
<tr>
<td>GO:1900006</td>
<td>positive regulation of dendrite development</td>
<td>2.51E-05</td>
<td>35</td>
</tr>
<tr>
<td>GO:000166</td>
<td>nucleotide binding</td>
<td>2.83E-05</td>
<td>486</td>
</tr>
<tr>
<td>GO:1901265</td>
<td>nucleoside phosphate binding</td>
<td>3.01E-05</td>
<td>486</td>
</tr>
<tr>
<td>GO:0007167</td>
<td>enzyme linked receptor protein signaling pathway</td>
<td>3.04E-05</td>
<td>270</td>
</tr>
<tr>
<td>GO:0050769</td>
<td>positive regulation of neurogenesis</td>
<td>4.05E-05</td>
<td>140</td>
</tr>
<tr>
<td>GO:0050770</td>
<td>regulation of axonogenesis</td>
<td>4.21E-05</td>
<td>66</td>
</tr>
<tr>
<td>GO:003824</td>
<td>catalytic activity</td>
<td>4.34E-05</td>
<td>1198</td>
</tr>
<tr>
<td>GO:1901888</td>
<td>regulation of cell junction assembly</td>
<td>4.39E-05</td>
<td>68</td>
</tr>
<tr>
<td>GO:0006793</td>
<td>phosphorus metabolic process</td>
<td>4.54E-05</td>
<td>731</td>
</tr>
<tr>
<td>GO:0050807</td>
<td>regulation of synapse organization</td>
<td>5.70E-05</td>
<td>72</td>
</tr>
<tr>
<td>GO:0060589</td>
<td>nucleoside-triphosphatase regulator activity</td>
<td>6.63E-05</td>
<td>102</td>
</tr>
<tr>
<td>GO:0007420</td>
<td>brain development</td>
<td>6.98E-05</td>
<td>201</td>
</tr>
<tr>
<td>GO:0006796</td>
<td>phosphate-containing compound metabolic process</td>
<td>7.81E-05</td>
<td>724</td>
</tr>
<tr>
<td>GO:0043547</td>
<td>positive regulation of GTPase activity</td>
<td>8.09E-05</td>
<td>118</td>
</tr>
<tr>
<td>GO:0007015</td>
<td>actin filament organization</td>
<td>8.49E-05</td>
<td>125</td>
</tr>
<tr>
<td>GO:0031589</td>
<td>cell-substrate adhesion</td>
<td>8.93E-05</td>
<td>107</td>
</tr>
<tr>
<td>GO:0016301</td>
<td>kinase activity</td>
<td>9.70E-05</td>
<td>204</td>
</tr>
<tr>
<td>GO:0050803</td>
<td>regulation of synapse structure or activity</td>
<td>1.07E-04</td>
<td>74</td>
</tr>
<tr>
<td>GO:0017124</td>
<td>SH3 domain binding</td>
<td>1.17E-04</td>
<td>48</td>
</tr>
<tr>
<td>GO:0099173</td>
<td>postsynapse organization</td>
<td>1.31E-04</td>
<td>56</td>
</tr>
<tr>
<td>GO:0036094</td>
<td>small molecule binding</td>
<td>1.59E-04</td>
<td>562</td>
</tr>
<tr>
<td>GO:0099003</td>
<td>vesicle-mediated transport in synapse</td>
<td>2.05E-04</td>
<td>64</td>
</tr>
<tr>
<td>GO:0007265</td>
<td>Ras protein signal transduction</td>
<td>2.14E-04</td>
<td>102</td>
</tr>
<tr>
<td>GO:0065009</td>
<td>regulation of molecular function</td>
<td>2.14E-04</td>
<td>658</td>
</tr>
<tr>
<td>GO:0033036</td>
<td>macromolecule localization</td>
<td>2.22E-04</td>
<td>689</td>
</tr>
<tr>
<td>GO:0016772</td>
<td>transferase activity, transferring phosphorus-containing groups</td>
<td>2.31E-04</td>
<td>235</td>
</tr>
<tr>
<td>GO:0032535</td>
<td>regulation of cellular component size</td>
<td>2.34E-04</td>
<td>112</td>
</tr>
<tr>
<td>GO:0060322</td>
<td>head development</td>
<td>2.38E-04</td>
<td>208</td>
</tr>
<tr>
<td>GO:0030695</td>
<td>GTPase regulator activity</td>
<td>2.70E-04</td>
<td>91</td>
</tr>
<tr>
<td>GO:0007267</td>
<td>cell-cell signaling</td>
<td>3.04E-04</td>
<td>407</td>
</tr>
<tr>
<td>GO:0046578</td>
<td>regulation of Ras protein signal transduction</td>
<td>3.20E-04</td>
<td>65</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>------------</td>
<td>--------------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0010720</td>
<td>positive regulation of cell development</td>
<td>3.53E-04</td>
<td>155</td>
</tr>
<tr>
<td>GO:0005096</td>
<td>GTPase activator activity</td>
<td>3.71E-04</td>
<td>83</td>
</tr>
<tr>
<td>GO:0007215</td>
<td>glutamate receptor signaling pathway</td>
<td>3.83E-04</td>
<td>41</td>
</tr>
<tr>
<td>GO:0032501</td>
<td>multicellular organisinal process</td>
<td>4.13E-04</td>
<td>1568</td>
</tr>
<tr>
<td>GO:0048638</td>
<td>regulation of developmental growth</td>
<td>4.16E-04</td>
<td>104</td>
</tr>
<tr>
<td>GO:0044085</td>
<td>cellular component biogenesis</td>
<td>4.75E-04</td>
<td>683</td>
</tr>
<tr>
<td>GO:0008289</td>
<td>lipid binding</td>
<td>4.80E-04</td>
<td>194</td>
</tr>
<tr>
<td>GO:0005089</td>
<td>Rho guanyl-nucleotide exchange factor activity</td>
<td>4.82E-04</td>
<td>27</td>
</tr>
<tr>
<td>GO:0098609</td>
<td>cell-cell adhesion</td>
<td>4.87E-04</td>
<td>218</td>
</tr>
<tr>
<td>GO:0099504</td>
<td>synaptic vesicle cycle</td>
<td>6.06E-04</td>
<td>60</td>
</tr>
<tr>
<td>GO:0031325</td>
<td>positive regulation of cellular metabolic process</td>
<td>6.53E-04</td>
<td>728</td>
</tr>
<tr>
<td>GO:0048518</td>
<td>positive regulation of biological process</td>
<td>6.85E-04</td>
<td>1278</td>
</tr>
<tr>
<td>GO:0007610</td>
<td>behavior</td>
<td>6.94E-04</td>
<td>160</td>
</tr>
<tr>
<td>GO:0010770</td>
<td>positive regulation of cell morphogenesis involved in differentiation</td>
<td>7.11E-04</td>
<td>55</td>
</tr>
<tr>
<td>GO:0035091</td>
<td>phosphatidylinositol binding</td>
<td>7.32E-04</td>
<td>78</td>
</tr>
<tr>
<td>GO:0050793</td>
<td>regulation of developmental process</td>
<td>7.36E-04</td>
<td>605</td>
</tr>
<tr>
<td>GO:0090066</td>
<td>regulation of anatomical structure size</td>
<td>7.48E-04</td>
<td>142</td>
</tr>
<tr>
<td>GO:0019900</td>
<td>kinase binding</td>
<td>7.78E-04</td>
<td>188</td>
</tr>
<tr>
<td>GO:0008104</td>
<td>protein localization</td>
<td>8.26E-04</td>
<td>609</td>
</tr>
<tr>
<td>GO:0018193</td>
<td>peptidyl-amino acid modification</td>
<td>8.42E-04</td>
<td>305</td>
</tr>
<tr>
<td>GO:0044089</td>
<td>positive regulation of cellular component biogenesis</td>
<td>8.53E-04</td>
<td>147</td>
</tr>
<tr>
<td>GO:0051239</td>
<td>regulation of multicellular organismal process</td>
<td>8.55E-04</td>
<td>721</td>
</tr>
<tr>
<td>GO:0061572</td>
<td>actin filament bundle organization</td>
<td>9.13E-04</td>
<td>56</td>
</tr>
<tr>
<td>GO:0030343</td>
<td>regulation of organelle organization</td>
<td>9.80E-04</td>
<td>306</td>
</tr>
<tr>
<td>GO:0050790</td>
<td>regulation of catalytic activity</td>
<td>1.10E-03</td>
<td>520</td>
</tr>
<tr>
<td>GO:0019901</td>
<td>protein kinase binding</td>
<td>1.14E-03</td>
<td>169</td>
</tr>
<tr>
<td>GO:00051173</td>
<td>positive regulation of nitrogen compound metabolic process</td>
<td>1.16E-03</td>
<td>690</td>
</tr>
<tr>
<td>GO:0016740</td>
<td>transferase activity</td>
<td>1.27E-03</td>
<td>513</td>
</tr>
<tr>
<td>GO:0003013</td>
<td>circulatory system process</td>
<td>1.33E-03</td>
<td>166</td>
</tr>
<tr>
<td>GO:0040011</td>
<td>locomotion</td>
<td>1.60E-03</td>
<td>444</td>
</tr>
<tr>
<td>GO:0036211</td>
<td>protein modification process</td>
<td>1.61E-03</td>
<td>876</td>
</tr>
<tr>
<td>GO:0006464</td>
<td>cellular protein modification process</td>
<td>1.61E-03</td>
<td>876</td>
</tr>
<tr>
<td>GO:0097435</td>
<td>supramolecular fiber organization</td>
<td>1.78E-03</td>
<td>182</td>
</tr>
<tr>
<td>GO:2000026</td>
<td>regulation of multicellular organismal development</td>
<td>1.78E-03</td>
<td>491</td>
</tr>
<tr>
<td>GO:0048583</td>
<td>regulation of response to stimulus</td>
<td>1.79E-03</td>
<td>899</td>
</tr>
<tr>
<td>GO:0005178</td>
<td>integrin binding</td>
<td>1.82E-03</td>
<td>50</td>
</tr>
<tr>
<td>GO:0051963</td>
<td>regulation of synapse assembly</td>
<td>1.92E-03</td>
<td>39</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0051049</td>
<td>regulation of transport</td>
<td>2.04E-03</td>
<td>417</td>
</tr>
<tr>
<td>GO:0030154</td>
<td>cell differentiation</td>
<td>2.20E-03</td>
<td>896</td>
</tr>
<tr>
<td>GO:0048869</td>
<td>cellular developmental process</td>
<td>2.20E-03</td>
<td>911</td>
</tr>
<tr>
<td>GO:0048589</td>
<td>developmental growth</td>
<td>2.26E-03</td>
<td>173</td>
</tr>
<tr>
<td>GO:0048588</td>
<td>developmental cell growth</td>
<td>2.26E-03</td>
<td>74</td>
</tr>
<tr>
<td>GO:0033674</td>
<td>positive regulation of kinase activity</td>
<td>2.57E-03</td>
<td>158</td>
</tr>
<tr>
<td>GO:0051234</td>
<td>establishment of localization</td>
<td>2.59E-03</td>
<td>1089</td>
</tr>
<tr>
<td>GO:0006468</td>
<td>protein phosphorylation</td>
<td>2.63E-03</td>
<td>447</td>
</tr>
<tr>
<td>GO:0051017</td>
<td>actin filament bundle assembly</td>
<td>2.76E-03</td>
<td>54</td>
</tr>
<tr>
<td>GO:0051345</td>
<td>positive regulation of hydrolase activity</td>
<td>2.94E-03</td>
<td>191</td>
</tr>
<tr>
<td>GO:0042391</td>
<td>regulation of membrane potential</td>
<td>2.96E-03</td>
<td>121</td>
</tr>
<tr>
<td>GO:0022603</td>
<td>regulation of anatomical structure morphogenesis</td>
<td>3.22E-03</td>
<td>282</td>
</tr>
<tr>
<td>GO:0008361</td>
<td>regulation of cell size</td>
<td>3.28E-03</td>
<td>60</td>
</tr>
<tr>
<td>GO:0016310</td>
<td>phosphorylation</td>
<td>3.46E-03</td>
<td>535</td>
</tr>
<tr>
<td>GO:0005488</td>
<td>binding</td>
<td>3.68E-03</td>
<td>3124</td>
</tr>
<tr>
<td>GO:0009893</td>
<td>positive regulation of metabolic process</td>
<td>3.71E-03</td>
<td>815</td>
</tr>
<tr>
<td>GO:1990138</td>
<td>neuron projection extension</td>
<td>4.17E-03</td>
<td>57</td>
</tr>
<tr>
<td>GO:0010604</td>
<td>positive regulation of macromolecule metabolic process</td>
<td>4.40E-03</td>
<td>757</td>
</tr>
<tr>
<td>GO:0098742</td>
<td>cell-cell adhesion via plasma-membrane adhesion molecules</td>
<td>4.54E-03</td>
<td>81</td>
</tr>
<tr>
<td>GO:0046873</td>
<td>metal ion transmembrane transporter activity</td>
<td>4.71E-03</td>
<td>116</td>
</tr>
<tr>
<td>GO:0006935</td>
<td>chemotaxis</td>
<td>5.27E-03</td>
<td>165</td>
</tr>
<tr>
<td>GO:0051641</td>
<td>cellular localization</td>
<td>5.39E-03</td>
<td>722</td>
</tr>
<tr>
<td>GO:0031646</td>
<td>positive regulation of nervous system process</td>
<td>5.44E-03</td>
<td>28</td>
</tr>
<tr>
<td>GO:0032970</td>
<td>regulation of actin filament-based process</td>
<td>5.97E-03</td>
<td>112</td>
</tr>
<tr>
<td>GO:0042330</td>
<td>taxis</td>
<td>6.39E-03</td>
<td>165</td>
</tr>
<tr>
<td>GO:0034220</td>
<td>ion transmembrane transport</td>
<td>6.77E-03</td>
<td>272</td>
</tr>
<tr>
<td>GO:0015020</td>
<td>glucuronosyltransferase activity</td>
<td>6.93E-03</td>
<td>17</td>
</tr>
<tr>
<td>GO:0019902</td>
<td>phosphatase binding</td>
<td>7.35E-03</td>
<td>60</td>
</tr>
<tr>
<td>GO:0044260</td>
<td>cellular macromolecule metabolic process</td>
<td>7.38E-03</td>
<td>1639</td>
</tr>
<tr>
<td>GO:0060627</td>
<td>regulation of vesicle-mediated transport</td>
<td>7.97E-03</td>
<td>140</td>
</tr>
<tr>
<td>GO:0048814</td>
<td>regulation of dendrite morphogenesis</td>
<td>8.38E-03</td>
<td>36</td>
</tr>
<tr>
<td>GO:0048041</td>
<td>focal adhesion assembly</td>
<td>8.46E-03</td>
<td>33</td>
</tr>
<tr>
<td>GO:0048639</td>
<td>positive regulation of developmental growth</td>
<td>8.67E-03</td>
<td>60</td>
</tr>
<tr>
<td>GO:0051094</td>
<td>positive regulation of developmental process</td>
<td>8.79E-03</td>
<td>334</td>
</tr>
<tr>
<td>GO:0051493</td>
<td>regulation of cytoskeleton organization</td>
<td>9.17E-03</td>
<td>144</td>
</tr>
<tr>
<td>GO:0048013</td>
<td>ephrin receptor signaling pathway</td>
<td>9.39E-03</td>
<td>34</td>
</tr>
<tr>
<td>GO:0044877</td>
<td>protein-containing complex binding</td>
<td>9.50E-03</td>
<td>287</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0042578</td>
<td>phosphoric ester hydrolase activity</td>
<td>9.92E-03</td>
<td>102</td>
</tr>
<tr>
<td>GO:0048523</td>
<td>negative regulation of cellular process</td>
<td>1.00E-02</td>
<td>1019</td>
</tr>
<tr>
<td>GO:1902414</td>
<td>protein localization to cell junction</td>
<td>1.02E-02</td>
<td>39</td>
</tr>
<tr>
<td>GO:0010810</td>
<td>regulation of cell-substrate adhesion</td>
<td>1.03E-02</td>
<td>66</td>
</tr>
<tr>
<td>GO:0034332</td>
<td>adherens junction organization</td>
<td>1.12E-02</td>
<td>28</td>
</tr>
<tr>
<td>GO:0017048</td>
<td>Rho GTPase binding</td>
<td>1.12E-02</td>
<td>51</td>
</tr>
<tr>
<td>GO:0060341</td>
<td>regulation of cellular localization</td>
<td>1.22E-02</td>
<td>235</td>
</tr>
<tr>
<td>GO:0004115</td>
<td>3',5'-cyclic-AMP phosphodiesterase activity</td>
<td>1.92E-03</td>
<td>1056</td>
</tr>
<tr>
<td>GO:0023056</td>
<td>positive regulation of signaling</td>
<td>1.33E-02</td>
<td>417</td>
</tr>
<tr>
<td>GO:0006810</td>
<td>transport</td>
<td>1.63E-02</td>
<td>9</td>
</tr>
<tr>
<td>GO:1990782</td>
<td>protein tyrosine kinase binding</td>
<td>1.68E-02</td>
<td>35</td>
</tr>
<tr>
<td>GO:0007156</td>
<td>homophilic cell adhesion via plasma membrane adhesion molecules</td>
<td>1.77E-02</td>
<td>71</td>
</tr>
<tr>
<td>GO:0051347</td>
<td>positive regulation of transferase activity</td>
<td>1.84E-02</td>
<td>71</td>
</tr>
<tr>
<td>GO:0010771</td>
<td>negative regulation of cell morphogenesis involved in differentiation</td>
<td>1.92E-02</td>
<td>171</td>
</tr>
<tr>
<td>GO:0052696</td>
<td>flavonoid glucuronidation</td>
<td>1.95E-02</td>
<td>128</td>
</tr>
<tr>
<td>GO:0034765</td>
<td>regulation of ion transmembrane transport</td>
<td>2.01E-02</td>
<td>414</td>
</tr>
<tr>
<td>GO:0006810</td>
<td>regulation of actin cytoskeleton organization</td>
<td>2.08E-02</td>
<td>414</td>
</tr>
<tr>
<td>GO:0040007</td>
<td>cell-matrix adhesion</td>
<td>2.14E-02</td>
<td>142</td>
</tr>
<tr>
<td>GO:0051347</td>
<td>macromolecule modification</td>
<td>2.19E-02</td>
<td>904</td>
</tr>
<tr>
<td>GO:0006050</td>
<td>developmental growth involved in morphogenesis</td>
<td>2.21E-02</td>
<td>904</td>
</tr>
<tr>
<td>GO:0040007</td>
<td>positive regulation of cell communication</td>
<td>2.34E-02</td>
<td>233</td>
</tr>
<tr>
<td>GO:0051347</td>
<td>growth</td>
<td>2.35E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0000160</td>
<td>blood circulation</td>
<td>2.36E-02</td>
<td>142</td>
</tr>
<tr>
<td>GO:0034765</td>
<td>regulation of postsynapse organization</td>
<td>2.36E-02</td>
<td>35</td>
</tr>
<tr>
<td>GO:0006836</td>
<td>neurotransmitter transport</td>
<td>2.49E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0008081</td>
<td>phosphoric diester hydrolase activity</td>
<td>2.54E-02</td>
<td>65</td>
</tr>
<tr>
<td>GO:0000160</td>
<td>regulation of cell growth</td>
<td>2.56E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0008066</td>
<td>regulation of actin cytoskeleton organization</td>
<td>2.68E-02</td>
<td>65</td>
</tr>
<tr>
<td>GO:0040007</td>
<td>regulation of growth</td>
<td>2.74E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0000160</td>
<td>cell-substrate junction assembly</td>
<td>3.00E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0051258</td>
<td>protein polymerization</td>
<td>3.23E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0061387</td>
<td>regulation of extent of cell growth</td>
<td>3.42E-02</td>
<td>65</td>
</tr>
<tr>
<td>GO:0004114</td>
<td>regulation of postsynapse organization</td>
<td>3.55E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0004114</td>
<td>phosphodiesterase activity</td>
<td>3.61E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0008066</td>
<td>glucuronidation activity</td>
<td>3.61E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0016049</td>
<td>cellular response to endogenous stimulus</td>
<td>3.93E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0045296</td>
<td>regulation of supramolecular fiber organization</td>
<td>4.07E-02</td>
<td>36</td>
</tr>
<tr>
<td>GO:0016049</td>
<td>growth</td>
<td>4.16E-02</td>
<td>88</td>
</tr>
<tr>
<td>GO:0045665</td>
<td>cadherin binding</td>
<td>4.16E-02</td>
<td>88</td>
</tr>
<tr>
<td>GO:0045665</td>
<td>cell growth</td>
<td>4.29E-02</td>
<td>67</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------------------------------------------</td>
<td>------------------</td>
<td>---------</td>
</tr>
<tr>
<td>GO:0150115</td>
<td>cell-substrate junction organization</td>
<td>4.30E-02</td>
<td>37</td>
</tr>
<tr>
<td>GO:0008038</td>
<td>neuron recognition</td>
<td>4.33E-02</td>
<td>22</td>
</tr>
<tr>
<td>GO:0051336</td>
<td>regulation of hydrolase activity</td>
<td>4.41E-02</td>
<td>292</td>
</tr>
<tr>
<td>GO:0017137</td>
<td>Rab GTPase binding</td>
<td>4.43E-02</td>
<td>52</td>
</tr>
<tr>
<td>GO:0030010</td>
<td>establishment of cell polarity</td>
<td>4.50E-02</td>
<td>46</td>
</tr>
<tr>
<td>GO:005201</td>
<td>extracellular matrix structural constituent</td>
<td>4.60E-02</td>
<td>51</td>
</tr>
<tr>
<td>GO:0051640</td>
<td>organelle localization</td>
<td>4.69E-02</td>
<td>160</td>
</tr>
<tr>
<td>GO:0045595</td>
<td>regulation of cell differentiation</td>
<td>4.95E-02</td>
<td>423</td>
</tr>
</tbody>
</table>

**B) KEGG pathways**

<table>
<thead>
<tr>
<th>Term ID</th>
<th>Term name</th>
<th>Adjusted p-value</th>
<th># genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>KEGG:04015</td>
<td>Rap1 signaling pathway</td>
<td>1.61E-06</td>
<td>72</td>
</tr>
<tr>
<td>KEGG:04510</td>
<td>focal adhesion</td>
<td>2.43E-06</td>
<td>69</td>
</tr>
<tr>
<td>KEGG:04072</td>
<td>phospholipase D signaling pathway</td>
<td>8.13E-06</td>
<td>54</td>
</tr>
<tr>
<td>KEGG:04360</td>
<td>axon guidance</td>
<td>1.38E-05</td>
<td>62</td>
</tr>
<tr>
<td>KEGG:04724</td>
<td>glutamatergic synapse</td>
<td>2.67E-05</td>
<td>44</td>
</tr>
<tr>
<td>KEGG:04935</td>
<td>growth hormone synthesis, secretion and action</td>
<td>8.32E-05</td>
<td>44</td>
</tr>
<tr>
<td>KEGG:05032</td>
<td>morphine addiction</td>
<td>8.98E-05</td>
<td>36</td>
</tr>
<tr>
<td>KEGG:04211</td>
<td>longevity regulating pathway</td>
<td>2.05E-04</td>
<td>35</td>
</tr>
<tr>
<td>KEGG:04213</td>
<td>longevity regulating pathway - multiple species</td>
<td>2.81E-04</td>
<td>27</td>
</tr>
<tr>
<td>KEGG:04810</td>
<td>regulation of actin cytoskeleton</td>
<td>3.47E-04</td>
<td>66</td>
</tr>
<tr>
<td>KEGG:04976</td>
<td>bile secretion</td>
<td>3.78E-04</td>
<td>35</td>
</tr>
<tr>
<td>KEGG:04012</td>
<td>ErbB signaling pathway</td>
<td>5.57E-04</td>
<td>33</td>
</tr>
<tr>
<td>KEGG:04512</td>
<td>ECM-receptor interaction</td>
<td>6.19E-04</td>
<td>34</td>
</tr>
<tr>
<td>KEGG:04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>9.34E-04</td>
<td>97</td>
</tr>
<tr>
<td>KEGG:04022</td>
<td>cGMP-PKG signaling pathway</td>
<td>1.23E-03</td>
<td>53</td>
</tr>
<tr>
<td>KEGG:04728</td>
<td>dopaminergic synapse</td>
<td>2.00E-03</td>
<td>44</td>
</tr>
<tr>
<td>KEGG:01522</td>
<td>endocrine resistance</td>
<td>2.03E-03</td>
<td>35</td>
</tr>
<tr>
<td>KEGG:04611</td>
<td>platelet activation</td>
<td>2.46E-03</td>
<td>42</td>
</tr>
<tr>
<td>KEGG:04725</td>
<td>cholinergic synapse</td>
<td>2.98E-03</td>
<td>39</td>
</tr>
<tr>
<td>KEGG:04928</td>
<td>parathyroid hormone synthesis, secretion and action</td>
<td>3.61E-03</td>
<td>37</td>
</tr>
<tr>
<td>KEGG:04540</td>
<td>gap junction</td>
<td>4.93E-03</td>
<td>32</td>
</tr>
<tr>
<td>KEGG:04730</td>
<td>long-term depression</td>
<td>5.70E-03</td>
<td>24</td>
</tr>
<tr>
<td>KEGG:04713</td>
<td>circadian entrainment</td>
<td>6.84E-03</td>
<td>34</td>
</tr>
<tr>
<td>KEGG:04020</td>
<td>calcium signaling pathway</td>
<td>6.91E-03</td>
<td>59</td>
</tr>
<tr>
<td>KEGG:04911</td>
<td>insulin secretion</td>
<td>7.87E-03</td>
<td>31</td>
</tr>
<tr>
<td>Term ID</td>
<td>Term name</td>
<td>Adjusted p-value</td>
<td># genes</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>KEGG:04926</td>
<td>relaxin signaling pathway</td>
<td>1.31E-02</td>
<td>41</td>
</tr>
<tr>
<td>KEGG:05100</td>
<td>bacterial invasion of epithelial cells</td>
<td>1.40E-02</td>
<td>27</td>
</tr>
<tr>
<td>KEGG:05412</td>
<td>arrhythmogenic right ventricular cardiomyopathy</td>
<td>1.48E-02</td>
<td>28</td>
</tr>
<tr>
<td>KEGG:04915</td>
<td>estrogen signaling pathway</td>
<td>1.52E-02</td>
<td>43</td>
</tr>
<tr>
<td>KEGG:04720</td>
<td>long-term potentiation</td>
<td>2.12E-02</td>
<td>25</td>
</tr>
<tr>
<td>KEGG:04520</td>
<td>adherens junction</td>
<td>2.25E-02</td>
<td>26</td>
</tr>
<tr>
<td>KEGG:04723</td>
<td>retrograde endocannabinoid signaling</td>
<td>2.43E-02</td>
<td>45</td>
</tr>
<tr>
<td>KEGG:04971</td>
<td>gastric acid secretion</td>
<td>3.04E-02</td>
<td>27</td>
</tr>
<tr>
<td>KEGG:04062</td>
<td>chemokine signaling pathway</td>
<td>4.03E-02</td>
<td>54</td>
</tr>
<tr>
<td>KEGG:04070</td>
<td>phosphatidylinositol signaling system</td>
<td>4.05E-02</td>
<td>32</td>
</tr>
<tr>
<td>KEGG:04912</td>
<td>GnRH signaling pathway</td>
<td>4.07E-02</td>
<td>31</td>
</tr>
<tr>
<td>KEGG:04925</td>
<td>aldosterone synthesis and secretion</td>
<td>4.99E-02</td>
<td>32</td>
</tr>
</tbody>
</table>
Table 2: Candidate germline de novo variants in RAS pathway genes

<table>
<thead>
<tr>
<th>Gene</th>
<th>Patient ID</th>
<th>Patient phenotype</th>
<th>Variant (hg38)</th>
<th>Variant consequence</th>
<th>gnomAD MAF</th>
<th>Predicted effect on splicing?</th>
<th>Damaging functional DNA scores?</th>
<th>TFBS?</th>
<th>Open chromatin in HUVECs?</th>
<th>OMIM phenotype</th>
<th>Homozygous knockout mouse model</th>
<th>CNV(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DLC1</td>
<td>Patie1nt 1</td>
<td>Infantile hemangioma; Cerebellar hypoplasia; Migration anomaly of cortex; Middle cerebral artery hypoplasia; Chorioretinal coloboma</td>
<td>8:13089422: C:T</td>
<td>intronic SNV (intron 15)</td>
<td>1.9e-4</td>
<td>—</td>
<td>—</td>
<td>transcriptional elongation</td>
<td>colorectal cancer, somatic</td>
<td>neural tube, heart, brain, and vascular placenta defects; embryonic lethality&lt;sup&gt;1&lt;/sup&gt;</td>
<td>I duplication in DECIPHER [ID: 395935]</td>
<td></td>
</tr>
<tr>
<td>DLC1</td>
<td>Patient 17</td>
<td>Infantile hemangioma; Coarctation of aorta; Sternal pit</td>
<td>8:13104433: G:A</td>
<td>intronic SNV (intron 7)</td>
<td>novel</td>
<td>—</td>
<td>EIGEN=0.25</td>
<td>transcriptional elongation</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>DLC1</td>
<td>Patient 3</td>
<td>Infantile hemangioma; Internal carotid artery hypoplasia; Internal carotid artery tortuosity; Absent vertebral artery; Tortuous aorta with pseudo-coarctation</td>
<td>8:13253055: G:C</td>
<td>intronic SNV (intron 5)</td>
<td>novel</td>
<td>HSF: activates cryptic acceptor site</td>
<td>—</td>
<td>—</td>
<td>strong enhancer</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>EPHA3</td>
<td>Patient 9</td>
<td>Infantile hemangioma; Sternal cleft; Dental enamel defects</td>
<td>3:89246570: G:A</td>
<td>intronic SNV (intron 3)</td>
<td>novel</td>
<td>HSF: activates cryptic acceptor site, changes exonic splicing silencer / enhancer</td>
<td>EIGEN=0.58</td>
<td>CTCF, TCF12, RAD21</td>
<td>—</td>
<td>—</td>
<td>perinatal death due to cardiac failure&lt;sup&gt;2&lt;/sup&gt;</td>
<td>I duplication in PHACE</td>
</tr>
<tr>
<td>PIK3CA</td>
<td>Patient 8</td>
<td>Infantile hemangioma; Subglottic hemangioma; Supra-umbilical raphe</td>
<td>3:179200365: T:G</td>
<td>intronic SNV (intron 3)</td>
<td>novel</td>
<td>HSF: changes exonic splicing silencer / enhancer</td>
<td>EIGEN=0.21</td>
<td>—</td>
<td>weakly transcribed</td>
<td>CLAPO syndrome, somatic; CLOVE syndrome, somatic; Cowden syndrome; keratosis, seborrhoeic, somatic; macrodactyly, somatic; megalencephaly-capillary malformation-polymergyria syndrome, somatic; nevus, epidermal, somatic; variety of cancers, somatic</td>
<td>vascular defects; hemorrhage; growth retardation; embryonic lethality&lt;sup&gt;3&lt;/sup&gt;</td>
<td>I duplication in DECIPHER [ID: 283584]</td>
</tr>
</tbody>
</table>

<sup>1</sup> (Durkin et al. 2005)<br><sup>2</sup> (Vaidya et al. 2003)<br><sup>3</sup> (Bi et al. 1999; Graupera et al. 2008; Lelievre et al. 2005)
<table>
<thead>
<tr>
<th>Gene</th>
<th>Patient ID</th>
<th>Patient phenotype</th>
<th>Variant (hg38)</th>
<th>gnomAD MAF</th>
<th>Predicted effect on splicing?</th>
<th>Damaging functional DNA scores?</th>
<th>TFBS?</th>
<th>Open chromatin in HUVECs?</th>
<th>OMIM phenotype</th>
<th>Homozygous knockout mouse model</th>
<th>CNV(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RASA3</td>
<td>Patient 4</td>
<td>Hypoplasia of cerebellum; Infantile hemangioma; Internal carotid artery hypoplasia; Internal carotid artery tortuosity; Optic nerve dysplasia</td>
<td>13:114041023: C:T</td>
<td>p.Val85Met</td>
<td>1.2e-5</td>
<td>HSF: changes exonic splicing silencer / enhancer</td>
<td>CADD=23.6; FATHMM=0.95; GWAVA=0.44</td>
<td>—</td>
<td>transcriptional elongation</td>
<td>hemorrhage; embryonic lethality&lt;sup&gt;4&lt;/sup&gt;</td>
<td>1 deletion in DECIPHER [ID: 393047]</td>
</tr>
<tr>
<td>STAMBP</td>
<td>Patient 16</td>
<td>Infantile hemangioma; Abnormal cerebral arterial morphology; Ventricular septal defect (VSD); Dental enamel hypoplasia</td>
<td>2:73850410: T:A</td>
<td>p.Ile301Asn</td>
<td>novel</td>
<td>—</td>
<td>CADD=27.7; FATHMM=0.99</td>
<td>weakly transcribed</td>
<td>microcephaly-capillary malformation syndrome</td>
<td>postnatal growth retardation; limb-clasping; hypocellular cerebral cortex; loss of hippocampal neurons; blepharoptosis; death of starvation at weaning&lt;sup&gt;6&lt;/sup&gt;</td>
<td>—</td>
</tr>
<tr>
<td>THBS2</td>
<td>Patient 14</td>
<td>Infantile hemangioma; Absent left internal carotid artery</td>
<td>6:169225343: C:T</td>
<td>p.Glu120Gly</td>
<td>novel</td>
<td>—</td>
<td>CADD=28.4; FATHMM=0.96</td>
<td>—</td>
<td>—</td>
<td>blood vessel abnormalities; premature death&lt;sup&gt;5&lt;/sup&gt;</td>
<td>5 deletions in DECIPHER [IDs: 3923007, 392970&lt;sup&gt;8&lt;/sup&gt;, 392995&lt;sup&gt;9&lt;/sup&gt;, 396392&lt;sup&gt;10&lt;/sup&gt;, 402744]</td>
</tr>
<tr>
<td></td>
<td>Patient 12</td>
<td>Infantile hemangioma; Aberrant right subclavian artery; Fetal origin of cerebral artery; Internal carotid artery hypoplasia; Internal carotid artery tortuosity; Vertebral artery hypoplasia; Apical muscular VSD</td>
<td>6:169243142: C:CCACATCT</td>
<td>p.Asp589Asn</td>
<td>novel</td>
<td>—</td>
<td>—</td>
<td>IKZF1</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

<sup>4</sup> (Iwashita et al. 2007)<br> <sup>5</sup> (Ishii et al. 2001)<br> <sup>6</sup> (Kyriakides et al. 1998)<br> <sup>7</sup> (Peeden et al. 1983)<br> <sup>8</sup> (Römke et al. 1987)<br> <sup>9</sup> (Petit et al. 1980)<br> <sup>10</sup> (Hopkin et al. 1997)
Table 3: Patient demographics

<table>
<thead>
<tr>
<th>Gender</th>
<th>Number of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>79</td>
</tr>
<tr>
<td>Male</td>
<td>19</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>85</td>
</tr>
<tr>
<td>Black or African American</td>
<td>4</td>
</tr>
<tr>
<td>Asian</td>
<td>2</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>2</td>
</tr>
<tr>
<td>Other or Unknown</td>
<td>5</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Not Hispanic or Latino</td>
<td>88</td>
</tr>
<tr>
<td>Hispanic or Latino</td>
<td>10</td>
</tr>
<tr>
<td>Facial hemangioma</td>
<td>90</td>
</tr>
<tr>
<td>Age (average)</td>
<td>5.8 years</td>
</tr>
</tbody>
</table>

All rights reserved. No reuse allowed without permission.

(which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

The copyright holder for this preprint
this version posted August 9, 2021. doi: medRxiv preprint
FIGURE LEGENDS

Figure 1: Location and consequences of THBS2–p.Asp859Asn variant
A) The location and interactions of p.Asp859 in the THBS2 protein structure 1YO8 in PDB. The p.Asp859 amino acid is highlighted
in green. The surrounding amino acids that participate in hydrogen binding with p.Asp859 and/or coordination with the calcium ions
are labeled. B-C) The wild-type (B) and mutant (C) structure predictions from DynaMut are given. The affected amino acid is
highlighted in teal. Red dashes indicate hydrogen bonds, and orange dashes indicate weak hydrogen bonds. The p.Asp859Asn variant
disrupts many of the interactions of p.Asp859 with its neighboring amino acids.

Figure 2: Expression of candidate genes in the developing heart at single cell level
A-F) Clustering of single cells by Uniform Manifold Approximation and Projection (UMAP) based on the differentially accessibility
of open chromatin regions (scATAC-seq). Gene expression levels of RASA3 (A), DLC1 (B), AFF2 (C), EPHA3 (D), PIK3CA (E), and
THBS2 (F) imputed from gene accessibility scores (scATAC-seq) and RNA expression (scRNA-seq). Annotation shows the
expression level of each gene in the corresponding cell type cluster.
Figure 3: Diagram of proposed RAS signaling pathway

RAS is activated downstream of growth factor binding to cellular receptors, which signal through GRB2/SOS/STAMBP, and deactivated by RAS-GAP proteins (RASA1, RASA3), which interact with EPHA3 and DLC1. Activated RAS-GTP signals through the MEK/ERK and PI3K/AKT/mTOR pathways to upregulate cell growth, cell migration, and angiogenesis. PI3K is also activated through extracellular matrix (ECM) signaling: THBS2 binds to the ECM and to integrins, which link the cell to the ECM, and THBS2 signals through focal adhesion kinase (FAK) to PI3K. Protein names denoted with an asterisk (DLC1, EPHA3, PIK3CA, RASA3, STAMBP, THBS2) were found to have de novo variants in our cohort. Image created using BioRender.