HAL
open science

Integrated clinical and omics approach to rare diseases novel genes and oligogenic inheritance in holoprosencephaly

Artem Kim, Clara Savary, Christèle Dubourg, Wilfrid Carré, Charlotte Mouden, Houda Hamdi-Roze, Hélène Guyodo, Jérôme Le Douce, Laurent Pasquier, Elisabeth Flori, et al.

To cite this version:

Artem Kim, Clara Savary, Christèle Dubourg, Wilfrid Carré, Charlotte Mouden, et al.. Integrated clinical and omics approach to rare diseases novel genes and oligogenic inheritance in holoprosencephaly. Brain - A Journal of Neurology , 2019, 142 (1), pp.35-49. 10.1093/brain/awy290 . hal-01975696

HAL Id: hal-01975696

https://univ-rennes.hal.science/hal-01975696

Submitted on 16 Apr 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Integrated Clinical and Omics Approach to Rare Diseases : Novel Genes and Oligogenic Inheritance in Holoprosencephaly

Running Title : Novel Genes and Oligogenicity of Holoprosencephaly

Artem Kim[1] Ph.D., Clara Savary[1] Ph.D., Christèle Dubourg[1,2] Pharm.D., Ph.D., Wilfrid Carré[2] Ph.D., Charlotte Mouden[1] Ph.D., Houda Hamdi-Rozé[1,2] M.D, Ph.D., Hélène Guyodo[1], Jerome Le Douce[1] M.S., FREX Consortium, GoNL Consortium, Laurent Pasquier[3] M.D., Elisabeth Flori[4] M.D., Marie Gonzales[5] M.D., Claire Bénéteau[6], M.D., Odile Boute[7] M.D., Tania Attié-Bitach[8] M.D., Ph.D., Joelle Roume[9] M.D., Louise Goujon[3], Linda Akloul M.D.[3], Sylvie Odent[3] M.D., Ph.D., Erwan Watrin[1] Ph.D., Valérie Dupé[1] Ph.D., Marie de Tayrac[1,2]* Ph.D., Véronique David[1,2]* Pharm.D., Ph.D.

1 - Univ Rennes, CNRS, IGDR (Institut de génétique et développement de Rennes) - UMR 6290, F - 35000 Rennes, France

2 - Service de Génétique Moléculaire et Génomique, CHU, Rennes, France.

3 - Service de Génétique Clinique, CHU, Rennes, France.

4 - Strasbourg University Hospital

5 - Service de Génétique et Embryologie Médicales, Hôpital Armand Trousseau, Paris, France

6 - Service de Génétique, CHU, Nantes, France

7 - Service de Génétique, CHU, Lille, France

8 - Service d'Histologie-Embryologie-Cytogénétique, Hôpital Necker-Enfants-Malades, Université Paris Descartes, 149, rue de Sèvres, 75015, Paris, France

9 - Department of Clinical Genetics, Centre de Référence "AnDDI Rares", Poissy Hospital GHU PIFO, Poissy, France.

** - equally contributing authors*

Corresponding Author: Dr. Marie de Tayrac.
Email: marie.detayrac@univ-rennes1.fr
Phone: +33 (0)2 23 23 45 43

The authors have declared that no conflict of interest exists.

Abstract.

Holoprosencephaly is a pathology of forebrain development characterized by high phenotypic heterogeneity. The disease presents with various clinical manifestations at the cerebral or facial levels. Several genes have been implicated in holoprosencephaly but its genetic basis remains unclear: different transmission patterns have been described including autosomal dominant, recessive and digenic inheritance. Conventional molecular testing approaches result in a very low diagnostic yield and most cases remain unsolved. In our study, we address the possibility that genetically unsolved cases of holoprosencephaly present an oligogenic origin and result from combined inherited mutations in several genes.

26 unrelated families, for whom no genetic cause of holoprosencephaly could be identified in clinical settings (Whole Exome Sequencing and CGH-array analyses), were reanalysed under the hypothesis of oligogenic inheritance. Standard variant analysis was improved with a gene prioritization strategy based on clinical ontologies and gene co-expression networks. Clinical phenotyping and exploration of cross-species similarities were further performed on a family-by-family basis. Statistical validation was performed on 248 ancestrally similar control trios provided by the Genome of the Netherlands project and on 574 ancestrally matched controls provided by the French Exome Project.

Variants of clinical interest were identified in 180 genes significantly associated with key pathways of forebrain development including Sonic hedgehog (SHH) and primary cilia. Oligogenic events were observed in 10 families and involved both known and novel HPE genes including recurrently mutated *FAT1*, *NDST1*, *COL2A1* and *SCUBE2*. The incidence of oligogenic combinations was significantly higher in HPE patients compared to two control populations ($p < 1E-9$). We also show that depending on the affected genes, patients present with particular clinical features.

This study reports novel disease genes and supports oligogenicity as clinically relevant model in HPE. It also highlights key roles of SHH signalling and primary cilia in forebrain development. We hypothesize that distinction between different clinical manifestations of HPE lies in the degree of overall functional impact on SHH signalling. Finally, we underline that integrating clinical phenotyping in genetic studies is a powerful tool to specify the clinical relevance of certain mutations.

Keywords: exome, holoprosencephaly, oligogenic inheritance, sonic hedgehog, primary cilia

Introduction

Holoprosencephaly (HPE1, OMIM #236100) is a severe developmental defect resulting from incomplete forebrain cleavage. The disease is characterized by incomplete separation of cerebral hemispheres with several anatomical classes ranging from microforms to alobar HPE. Affected individuals present with typical craniofacial midline defects of varying severity including proboscis, cleft lip and palate, ocular hypotelorism and solitary median incisor. HPE occurs in about 1 in 10,000 to 20,000 live births worldwide (Mercier *et al.*, 2011).

The genetic basis of HPE remains unclear and different transmission patterns have been described including autosomal dominant, recessive and digenic inheritance (Dubourg *et al.*, 2018). Most mutations associated with HPE display incomplete penetrance and variable expressivity, *i.e.* close relatives carrying the same pathogenic variant can be asymptomatic or present distinct HPE-spectrum anomalies (Mercier *et al.*, 2011). *Sonic Hedgehog* (*SHH*) was the first discovered gene implicated in HPE (Roessler *et al.*, 1996) and its variants remain the most common cause of non-chromosomal HPE (Dubourg *et al.*, 2018). In 2011, molecular screening of 645 HPE probands revealed that mutations in the *SHH*, *ZIC2*, *SIX3* and *TGIF1* genes were the most frequent ones and collectively accounted for 25 % of the cases (Mercier *et al.*, 2011). The following studies reported that *GLI2* might also be considered as a major HPE gene in terms of frequency (Dubourg *et al.*, 2016), although variants in *GLI2* rarely result in classic HPE but instead cause a distinct phenotype that includes pituitary insufficiency and subtle facial features (Bear *et al.*, 2014). Pathogenic variants in *FGF8*, *FGFR1*, *DISP1*, and *DLL1* were also found in approximately 7 % of HPE cases (Dupé *et al.*, 2011; Dubourg *et al.*, 2016). The other HPE genes reported so far are *TDGF1*, *FOXH1*, *TGIF1*, *CDON*, *NODAL*, *GAS1*, *STIL* and *SUFU* whose frequency is not established due to the small

number of reported cases (Mouden *et al.*, 2015, 2016; Dubourg *et al.*, 2018; Kruszka *et al.*, 2018).

Clinical genetic testing of HPE has improved, but approximately 70 % of familial cases remain without a clear molecular diagnosis. Most of known HPE genes belong to the Sonic Hedgehog (SHH) pathway, which represents the primary pathway implicated in the disease (Mercier *et al.*, 2013; Dubourg *et al.*, 2016; Kruszka *et al.*, 2018). Therefore, defective SHH-related processes are likely to be substantially involved in HPE.

Whole-exome sequencing (WES) has been very successful for Mendelian disease-gene discovery and differential diagnosis (Bamshad *et al.*, 2011). WES analysis employs filtering approaches for candidate variant prioritization combined with comprehensive clinical evaluation. A variety of additional strategies has been developed to further improve the performance of WES in clinical settings. Collaborative platforms such as Matchmaker Exchange (Philippakis *et al.*, 2015) are used to search for recurrence in patients affected by similar phenotypes. Integrative variant-prioritization algorithms such as the Exomiser suite (Smedley *et al.*, 2015) combine WES with different phenotype-driven approaches (based on clinical data and cross-species phenotype comparisons) and analysis of protein interactome data. As useful as they are, these strategies are limited: collaborative platforms are not efficient in case of very rare genetic diseases while pipelines such as Exomiser are not designed to study non-Mendelian disorders. Studying HPE faces these two challenges: (i) HPE live-born infants are excessively rare, and, (ii) although HPE is considered as a Mendelian disorder, the wide range of severity must necessitate strong modifying factors such that a single pathogenic variant may be neither necessary nor sufficient for pathogenesis.

Recent studies have highlighted that non-Mendelian disease phenotypes could present an oligogenic etiology and result from accumulation of inherited low-penetrance variants in multiple genes (Li *et al.*, 2017). However, such events are likely overlooked in clinical genetic studies if variants are inherited from a clinically unaffected parent.

In this study, we address the additional yield that can be obtained for HPE patients who underwent medical WES evaluation in clinical settings that failed to establish a molecular diagnosis. Given the wide clinical spectrum of the disease, as well as incomplete penetrance and variable expressivity of HPE mutations, we raised the possibility that the low diagnostic yield is partly due to the complex etiology of HPE and hypothesized that a part of unsolved HPE cases results from oligogenic events, *i.e.* accumulation of several rare hypomorphic variants in distinct, functionally connected genes.

Our study involved patients for whom no disease etiology could be determined by conventional diagnostic approaches. Similarly to previous WES studies (Lee *et al.*, 2014; Stark *et al.*, 2017), we used clinically-driven prioritization approach to identify genes associated with specific clinical features as reported in gene-phenotype reference databases and mouse models. Complementarily, we developed and used a prioritization strategy based on gene co-expression networks of the developing human brain to select genes with spatio-temporal expression patterns compatible with those of known HPE genes. Finally, we used in-depth clinical phenotyping together with cross-species similarities to further strengthen the evidence of causality.

This study highlights novel HPE genes and identifies new disease-related pathways including the primary cilia pathway. Our findings also illustrate the high degree of oligogenicity of HPE and suggest that the disease requires a joint effect of multiple hypomorphic mutations.

Materials & Methods

Patient selection and preliminary genetic analyses

Study protocol was approved by the Ethics Committee of Rennes Hospital. Patients diagnosed with HPE and relatives were recruited using the clinical database of Holoprosencephaly Reference Center of Rennes Hospital. Study participation involved informed written consent, availability of clinical data, and either DNA or peripheral blood sample.

The main selection criteria for this study was the absence of clear genetic cause of HPE after conventional diagnostic procedures. As part of routine diagnosis, all patients were scanned for rare damaging mutations by targeted HPE gene-panel sequencing (Dubourg *et al.*, 2016) and for copy number variants (CNVs) using CGH-array and MLPA. Patients for whom no genetic cause of HPE (*i.e.*, a fully-penetrant causal mutation in known HPE gene or a chromosomic aberration/copy number variant explaining the pathology) could be established, underwent trio-based WES for further analysis. WES was performed using standard procedures as previously described (Mouden *et al.*, 2015, 2016). The scheme for variant classification followed the American College of Medical Genetics and Genomics–Association (ACMG) guidelines (Richards *et al.*, 2015) and included a hypothesis-free analysis of all *de novo* and homozygous variants on a family-by-family basis. Patients for whom no such variants of clinical interest had been detected were considered eligible for the hypothesis of oligogenic inheritance and included in this study.

Variant selection under oligogenic hypothesis

As discussed in previous studies, ACMG guidelines are useful in identifying variants with strong effect on phenotype but are unhelpful in case of modifier variants (Hong *et al.*, 2017). Therefore, the ACMG classification was not taken into account for variant selection dedicated to the analysis of oligogenic events. WES trio data was reanalysed using more permissive settings (filtering protocols used in this study are described in the *Supplementary Appendix)*. The exome analysis was complemented with two gene prioritization strategies based on available clinical knowledge and co-expression networks.

Clinically-driven approach.

We established two clinician-generated lists of relevant phenotypes reminiscent of HPE in human and mouse models respectively (*Table S3*). Genes associated with the phenotypes of interest were identified with publicly available clinical resources and associated ontologies. Human gene-phenotype associations were extracted from relevant databases (*Figure S1*) using R package *VarFromPDB* (*https://github.com/cran/VarfromPDB*). The Mouse Genome Informatics (MGI) (Smith *et al.*, 2018) database and a homemade workflow were used to retrieve genes associated with any of the corresponding phenotypes in mouse mutants. Human and mouse results were combined and redundancy was removed to establish a list of clinically-driven candidate genes associated with HPE-related anomalies (*Table S4*).

Identification of HPE-related genes by weighted-gene co-expression network analysis

We used Weighted Gene Co-expression Network Analysis (WGCNA) (Langfelder and Horvath, 2008) on the RNA-Seq data from the Human Development Biology Resource (HDBR) (Lindsay *et al.*, 2016) to identify genes sharing highly similar expression patterns with

4 classical genes associated to HPE – *SHH*, *SIX3*, *ZIC2* and *TGIF1* – during cerebral development. Data from samples corresponding to forebrain, cerebral cortex, diencephalon, telencephalon and temporal lobe structures taken between the 4th and 10th post-conception weeks (pcw) were selected (*Figure S9*). RNA-seq data were analyzed with the iRAP pipeline (*https://github.com/nunofonseca/irap*). We used R package WGCNA to construct co-expression networks and identify modules of co-expressed genes. The detailed protocols for WGCNA analysis are described in the *Supplementary Appendix*. The Topological Overlap Matrix (TOM) matrix was used to establish a list of transcriptome-driven candidate genes sharing highly similar expression profiles with *SHH, ZIC2, SIX3* and *TGIF1* (*Table S5*).

Integration and identification of oligogenic events

The two gene prioritization schemes were combined with the WES results to identify a restricted list of rare variations located in genes identified by either the transcriptomic or the clinical prioritization approach (*Figure 1*). Further analyses of the candidate variants were performed on a family-by-family basis. Oligogenic events were defined as combinations of candidate variants in ≥ 2 genes co-segregating with disease, *i.e.* unique to the affected individuals of each family. Variants could be either inherited from the parents — at least one each from the mother and the father - or occur *de novo* in the affected child.

To further evaluate the impact of candidate genes, we performed deep clinical phenotyping to characterize similarities between unrelated patients and/or published knockout mice. Special attention was given to genes harboring distinct rare variants in at least two affected patients with striking phenotypic overlap. Phenotypic overlaps between patients and mouse mutants deficient for the corresponding candidate genes were also examined. The most

interesting oligogenic combinations of rare deleterious variants in the affected children were finally discussed during multidisciplinary meetings.

To determine significantly enriched biological processes and pathways, functional annotation was performed by *g:profiler (http://biit.cs.ut.ee/gprofiler)* and Bonferroni adjusted p-value were considered significant below 0.05 value (KEGG, REACTOME and Gene Ontology Biological Processes).

Control Cohorts and validation

To test whether the identified oligogenic combinations were specific to the HPE cohort, we used SNV and INDELS data from 248 healthy trios (744 individuals) provided by Genome of the Netherlands (GoNL) sequencing project as a control cohort (Genome of the Netherlands Consortium, 2014). Additional control cohort consisting of 574 unrelated French individuals was provided by the French Exome Project (FREX).

We applied the same variant filtering approach and the same strategy for selection of oligogenic events. Proportion of families and/or individuals presenting oligogenic events were then compared between HPE cohort and the control cohorts. P-values were calculated using two-sided Fisher's exact test (*fisher.test* function in R, version 3.4.2).

Results

Clinical findings

We assembled a cohort of 26 families representing a total of 80 individuals including 29 affected children diagnosed with lobar (*n=3*), semilobar (*n=11*), alobar (*n=13*) or microform HPE (*n=2*) (*Table 1)*. Common HPE clinical manifestations were observed among the

probands and included cleft lip and palate (38 %), hypotelorism (34 %), microcephaly (31 %) and arhinencephaly (31 %). Ancestry analysis identified that 24 families were of European descent and two of South East Asia and African descent (*Figure S10*). 8 parents presented minor signs of midline facial anomalies and 3 parents were diagnosed with HPE microforms.

The initial targeted sequencing had identified point mutations in known HPE genes in 13 families and a full heterozygous deletion of *SIX3* gene had been detected by CGH-array in one family (*Figure 2* and *Figure S8)*. All anomalies were later confirmed by WES analysis. They were inherited from asymptomatic or mildly affected parents and were considered as insufficient to fully explain the pathogenesis of HPE, suggesting that the presence of additional risk factors was required for the disease to occur.

HPE variants overview and identification of disease-related pathways

Combined clinically- and transcriptome-driven analysis of the exome data identified a total of 232 rare candidate variants in 180 genes (Figure 1 and Table S6). All variants presented a Minor Allele Frequency (MAF) below 1 % and were predicted to be highly deleterious to protein function (*see Supplementary*). 153 variants concerned genes associated with HPE phenotypes among which 32 were located in genes reported to induce HPE-like phenotypes in mutant mice (*Table S8*). 102 variants were located in genes sharing expression profiles highly similar to those of HPE genes. Overlap between phenotype and gene co-expression network analysis contains 23 variants including 14 previously described mutations in known HPE genes (*SHH, ZIC2, SIX3, GLI2, TGIF1* and *PTCH1*).

Consistent with known disease etiology, functional profiling of the 180 genes revealed a significant enrichment for biological processes implicated in forebrain development (Table S7) including Sonic Hedgehog signalling pathway *(REAC:5358351, p-value = $2.79E^{-5}$;*

KEGG:04340, p-value = 1E^{-4}), Primary Cilia (*REAC:5617833, p-value = 1E^{-6}; GO:0060271, p-value = 2E^{-6}*) and Wnt/Planar Cell Polarity (PCP) signalling pathway (GO:0016055, p-value = 2E^{-5}). The SHH pathway is the primary pathway implicated in HPE and the primary cilium is required for the transduction of SHH signalling (Gorivodsky *et al.*, 2009; Murdoch and Copp, 2010) while components of Wnt/PCP pathway regulate both SHH signalling and primary cilia (Goetz *et al.*, 2009; Murdoch and Copp, 2010).

In-depth analyses highlighted 10 families with oligogenic events (*Figure 2*) clustered among 19 genes *(Table 2)* that functionally relate to disease-relevant pathways (*Figure 3*). These combinations of variants were unique to the affected probands. Main findings are presented below and full reports are available in the *Supplementary Appendix.*

Recurrent oligogenic events involving FAT1

4 different families, *i.e.* 15 % of the 26 families studied here, presented oligogenic events involving *FAT1* in combination with rare variants in known HPE genes (*SHH, PTCH1),* as well as in *NDST1, COL2A1* and *LRP2* genes (*Figure 2A*). FAT1 is a protocadherin and its knockdown in mouse causes severe midline defects including HPE, (Ciani *et al.*, 2003) when in Drosophilia it has been shown to regulate the PCP pathway (Rock *et al.*, 2005). *LRP2*, *NDST1* and *COL2A1* are all functionally relevant to the SHH pathway (*Figure 3*): *NDST1* and *COL2A1* mice mutants exhibit HPE phenotype and reduced SHH signalling in the forebrain (Grobe, 2005; Leung *et al.*, 2010), while *LRP2* acts as an auxiliary receptor of SHH during forebrain development and its inactivation in mouse similarly leads to HPE phenotype (Christ *et al.*, 2012).

Oligogenic events involved the following combinations: *SHH/FAT1/NDST1* (family F3), *FAT1/NDST1/COL2A1* (Family F16), *FAT1/COL2A1/PTCH1* (family F26) and *FAT1/LRP2* (Family

F23) (*Figure 2A* and *Table 2*). The details are provided in the *Supplementary appendix, Case report (1)*.

In family F3, Sanger sequencing of additional family members revealed that the *SHH/FAT1/NDST1* combination was unique to the affected individuals (*Figure 2A)*. For family F16, only the foetus carrying the *FAT1/NDST1/COL2A1* combination was affected by semilobar HPE, while the sibling carrying *NDST1/COL2A1* variants presented only a microform (*Figure 2A*). These observations are fully consistent with the oligogenic inheritance model where accumulation of multiple variants in genes associated to HPE phenotypes and/or HPE-related molecular pathways is required.

Recurrent oligogenic events involving SCUBE2/BOC implicated in SHH signalling

Similarly, 2 families presented oligogenic events implicating combined variants in *BOC* and *SCUBE2* genes (*Figure 2B and Table 2)*. BOC is an auxiliary receptor of SHH and was recently reported as HPE modifier in Human (Hong *et al.*, 2017). *SCUBE2* shares a highly similar expression pattern with *SHH* and *SIX3* and is implicated in the release of *SHH* from the secreting cell (Jakobs *et al.*, 2014). In family F4, combination of *SCUBE2/BOC* variants was associated with additional variants in *SHH*, *STK36* (see below) and *WNT4,* a member of Wnt pathway, implicated in regulation of SHH signalling (Murdoch and Copp, 2010). In family F22, the *SCUBE2* variant results in a premature stop codon at position 525 (*Figure S7*), which results in truncation of its CUB domain and is predicted to directly affect its SHH-related activity (Jakobs *et al.*, 2014). This family presented an additional candidate variant in *HIC1,* which genetically interacts with *PTCH1* (Briggs *et al.*, 2008). Mice deficient for *HIC1* exhibit craniofacial defects including HPE (Carter, 2000).

The reported variant combinations were observed exclusively in the affected probands and were absent in asymptomatic individuals. Altogether, these results reveal recurrent mutations in *SCUBE2/BOC* and further strengthen the oligogenic inheritance model of HPE.

Implication of primary cilium in HPE

Remarkably, 5 families presented candidate variants in genes related to the primary cilium - *STK36, IFT172, B9D1, MKS1, TCTN3* and *TULP3* (*Figure 2C*). Ciliary proteins are known to play essential roles in the transduction of SHH signalling downstream of PTCH1 during forebrain development (Goetz *et al.*, 2009; Murdoch and Copp, 2010).

STK36, also known as *Fused*, is a ciliary protein implicated in SHH signalling and associated to craniofacial phenotypes (Goetz *et al.*, 2009; Murdoch and Copp, 2010). *IFT172* codes for a core component of intraflagellar transport complex IFT-B required for ciliogenesis and regulation of SHH signal transduction. Moreover, *IFT172*$^{-/-}$ mice exhibit reduced expression of *Shh* in the ventral forebrain and severe craniofacial malformations including HPE (Gorivodsky *et al.*, 2009). *B9D1, MKS1* and *TCTN3* are all members the transition zone protein complex implicated in regulation of ciliogenesis (Garcia-Gonzalo *et al.*, 2011). The disruption of *B9D1* and *MKS1* in mouse models causes craniofacial defects that include HPE (Dowdle *et al.*, 2011; Wheway *et al.*, 2013). Although no mouse model is available for *TCTN3,* its expression profile is highly similar to that of *SHH* and disruption of its protein complex partners (*TCTN1*, *TCTN2, CC2D2A, MKS1, B9D1*) leads to HPE in mouse (Dowdle *et al.*, 2011; Garcia-Gonzalo *et al.*, 2011; Wheway *et al.*, 2013). Moreover, *TCTN3* was shown to be necessary for the transduction of SHH signal and *TCTN3* mutations were found in patients affected by ciliopathies (Thomas *et al.*, 2012). Finally, *TULP3* is a critical repressor of *Shh*

signalling in mouse and is associated to various craniofacial defects (Murdoch and Copp, 2010).

Additional variants observed in these families include a heterozygous deletion of *SIX3*, missense mutations in *SHH*, *SCUBE2*, *BOC* and *LRP2 (*described above*)* as well as two genes implicated in PCP pathway (*Figure 3*): *CELSR1* (2 families) and *PRICKLE1*, both associated with craniofacial defects in mouse mutants (*Figure 2C*) (Goetz *et al.*, 2009; Murdoch and Copp, 2010; Yang *et al.*, 2014). Similarly to previously described cases, the oligogenic events were present exclusively in the affected children.

Given the essential role of the primary cilium in SHH signal transduction, these observations strongly suggest that rare variants in ciliary genes contribute to the disease onset in these families.

Correspondence between affected genes and secondary clinical features

To provide additional evidence, we performed an in-depth analysis of secondary clinical features associated with HPE in our patients. Deep clinical phenotyping identified clinical similarities between unrelated patients (*Table 2*) as well as overlaps of secondary clinical features between patients and the corresponding mouse mutants.

Interestingly, the 2 patients with variants in ciliary genes (*IFT172/PRICKLE1* and *SIX3/TCTN3/TULP3*) both presented polydactyly, a clinical feature commonly associated with ciliopathies (Goetz *et al.*, 2009). Importantly, the patient with the oligogenic combination *IFT172/PRICKLE1* presented a large set of overlapping clinical features with the corresponding mouse mutants including polydactyly, cleft palate and eye defects (Gorivodsky *et al.*, 2009; Yang *et al.*, 2014).

Noteworthy, the two unrelated patients having variants in *FAT1* and *NDST1* shared a large set of specific secondary clinical features, including mandibular and ear abnormalities. Intrauterine growth restriction (IUGR) was found exclusively in the two patients with *COL2A1* variants. The most severely affected child in family F16 (*FAT1/NDST1/COL2A1*) presented a strong overlap with *NDST1-null* and *COL2A1*-null mutant mice (HPE, mandibular anomalies, absent olfactory bulb, abnormal nose morphology) (Grobe, 2005; Leung *et al.*, 2010). Similarly, proboscis and eye defects were observed in both *FAT1/NDST1/SHH* patient and *FAT1*$^{-/-}$ mice (Ciani *et al.*, 2003).

Finally, the 2 unrelated *SCUBE2/BOC* cases in families F4 and F22 presented cebocephaly, a midline facial anomaly characterized by ocular hypotelorism and a single nostril, which was absent in all other patients. Consistently, *SCUBE2* is highly expressed in the nasal septum in mouse (Xavier and Cobourne, 2011), and cebocephaly was previously associated with *CDON* – another known HPE gene sharing highly similar functions and structure with *BOC* (Zhang *et al.*, 2006).

While these clinical features are not specific to HPE, the described overlaps provide additional support for disease implication of the presented candidate variants.

Statistical validations

The identified oligogenic events were clustered among 19 genes (*Table 2* and *Figure 2*). To assess the frequency of healthy individuals presenting similar variant combinations in these genes, we applied the same family-by-family variant analysis to the 248 control trios provided by GoNL. This control cohort was chosen as 24/26 (92 %) of the HPE families included in the study were of European descent (*Figure S10*).

The approach identified a total of 3 families among controls presenting variant combinations satisfying the criteria that we established for the oligogenic events (gene, variant and parental inheritance). The 3 oligogenic events found in the control cohort were *FAT1/B9D1*, *SCUBE2/PTCH1* and *SCUBE2/LRP2/PTCH1/CELSR1 (Table S9)*. Although one *SCUBE2* variant (p.Thr285Met) was found in both the HPE and the control cohort, none of the combinations found among controls corresponded to oligogenic events identified in the HPE cohort. The incidence of oligogenic events was significantly lower in the GoNL families (3/248*, i.e.* 1.2 %) as compared to the HPE cohort (10/26, *i.e.* 38 %) with a Fisher's exact test p-value of 2.301e-09 (*Table 3*).

Three additional children of the GoNL cohort harboured combinations of rare deleterious variants in ≥ 2 candidate genes. However, in these cases, all variants were inherited from the same parent. Therefore, these combinations were not considered as oligogenic events similar to those of HPE patients. Nevertheless, even when taking into account these 3 additional cases, the proportion of children having variants in ≥ 2 candidate genes was significantly different between the HPE cohort (13/29, *i.e. 45 %)* and the GoNL cohort (6/248, *i.e.* 2,4 %) with a Fisher's exact test p-value of *1.902e-10.*

Finally, a total of 14 individuals of the GoNL cohort (parents and children combined) harboured rare deleterious variants in ≥ 2 genes. Without taking into account the relatedness between the GoNL individuals, the proportion of individuals having variants in ≥ 2 candidate genes remained significantly different between the HPE cohort (21/80, i.e. 26 %) and the GoNL control cohort (14/744*,* i.e. 1.8 %), as confirmed by Fisher's exact test (p-value = *3.237e-14*).

To further assess the frequency of control individuals presenting rare variant combinations in the identified candidate genes (*Table 2* and *Figure 2*), we analysed a second control cohort. The FREX data was chosen as it consists in 574 unrelated French individuals ancestrally matching the HPE cohort.

Screening of the FREX cohort revealed that 16/574 individuals (i.e. 2.7 %) harbored rare deleterious variants in ≥ 2 candidate genes. This proportion was statistically different from that observed in HPE cohort (21/80, i.e. 26 % vs 16/574, i.e. 2.7 %; p-value= 1.521e-11, Fisher's exact test).

Additionally, the two control cohorts (GoNL and FREX) did not present statistically significant differences in terms of proportions of individuals having rare deleterious variants in ≥ 2 candidate genes: 14/744 (i.e. 1.8 %) for the GoNL cohort vs. 16/574 (i.e. 2.7 %) for the FReX (p-value = 0.35, Fisher's exact test).

The analysis of the GoNL and FReX cohorts illustrates that the incidence of combined rare deleterious variants in the identified candidate genes is significantly higher in HPE patients as compared to a control population. All performed comparisons showed a statistically significant p-value between the cases and the controls (Table 3), thus providing evidence for oligogenicity as clinically relevant model in HPE.

Discussion

In this study, we addressed the relevance of oligogenic model for unsolved HPE cases. We provide evidence that the onset of HPE arises from the combined effects of hypomorphic variants in several genes belonging to critical biological pathways of brain development. To circumvent the limitations of classical WES analysis in complex rare disorders, we combined clinically-driven and co-expression network analyses with classical WES variant prioritization. This strategy was applied to 26 HPE families and allowed prioritization of 180 genes directly linked to the SHH signalling, cilium and Wnt/PCP pathways (*Figure 3*). The analysis of oligogenic events in patients with HPE anomalies revealed 19 genes including 15 genes previously unreported in human HPE patients (*Table 2*). All these genes are either associated with HPE phenotypes in corresponding mouse models (such as *FAT1*, *NDST1*), present highly similar expression patterns with already known HPE genes in the developing brain (such as *SCUBE2, TCTN3*) or both. We observed co-occurrence of mutations in several gene pairs such as *FAT1/NDST1* and *SCUBE2/BOC,* which provides additional arguments towards their implication in HPE. The incidence of oligogenic combinations was significantly higher in HPE patients compared to the GoNL and FREX control populations. We additionally show that in-depth evaluation of secondary clinical features in patients with HPE anomalies and comparison to published mouse knockout models may provide additional arguments for the causality of candidate genes.

The main challenge in disease-gene discovery by Whole Exome Sequencing is to identify disease-related variants among a large background of non-pathogenic polymorphisms (Bamshad *et al.*, 2011; MacArthur *et al.*, 2014). For example, the presented *FAT1* encodes a large protocadherin gene spanning over 139 kb in the human genome and presenting over

2000 missense variants with MAF below 1 % in the gnomAD database. Despite this high number of variations found in the general population, rare variants in *FAT1* were recently implicated in several genetic disorders including facioscapulohumeral dystrophy-like disease (Puppo *et al.*, 2015). Hence, correct interpretations and conclusions require extremely careful assessment of available biological and clinical knowledge.

To improve the pertinence of our study, we developed a strategy to restrict the potential candidates by targeting genes with biological and clinical arguments for their implication in the disease. Implication of a given gene in a disease is often supported by the similarity between the human pathology and the phenotype obtained in relevant animal models (MacArthur *et al.*, 2014). Accordingly, in this study, the main evidence of causality for candidate genes was that their disruption leads to clinically-defined HPE-related phenotypes in corresponding published mutant mouse models. Unlike other phenotypes such as reduced body weight (Reed *et al.*, 2008), holoprosencephaly is a rare effect of gene knockout in mice as it is associated with less than 1 % of knockout mice (as reported in the MGI database). Recent exome sequencing studies have applied similar phenotype-driven approaches to identify causal variants in monogenic disorders. Dedicated tools have been developed in that aim (Exomiser, Phive) (Smedley *et al.*, 2015) but none are designed for non-Mendelian traits involving hypomorphic variants with mild effects. We provide a method to specifically address such cases and show that further developments are necessary to improve the diagnosis of genetic disorders especially by taking into account oligogenic inheritance. Inclusion of carefully defined mouse mutant phenotypes is of powerful value as certain phenotypes like HPE are very informative due to their rarity.

Prioritization tools can also include protein–protein interaction (PPI) networks information, which improves performance in cases where candidate genes do not have an associated knockout mouse model. However, PPI-based prioritization is limited when disease investigation requires incorporation of tissue-specific data. The key process affected by HPE is the elaboration of the forebrain and its dorso-ventral patterning (Fernandes and Hébert, 2008). Deciphering the biological mechanisms involved in the early brain development is therefore necessary to provide relevant information to select disease-related genes. To incorporate tissue-specificity, we performed analysis using the RNA-Seq data of embryonic human brain at the earliest available developmental stages (from 4 to 17 pcw) as provided by the Human Development Biology Resource (Lindsay *et al.*, 2016). We defined relevant co-expression modules and selected candidate genes of which expression patterns follow those of known HPE genes. Further analysis showed that the resulting candidate genes, such as *SCUBE2* and *TCTN3*, are pertinent as they are equally implicated in the SHH pathway that is the primary HPE pathway (Thomas *et al.*, 2012; Jakobs *et al.*, 2014). Co-expression analysis provides additional insight into disease pathogenesis by establishing the first link between previously unrelated genes. A future challenge will be to generalise this approach but such a task will face the necessity to incorporate disease relevant co-expression modules that need to be pre-computed.

Patients exhibiting HPE-anomalies present enrichment of rare variants in genes related to the SHH pathway, as well as to the Wnt/PCP and primary cilia pathways, which were both shown to functionally interact with and regulate SHH pathway (Goetz *et al.*, 2009; Gorivodsky *et al.*, 2009; Murdoch and Copp, 2010; Wheway *et al.*, 2013). Accumulation of multiple rare variants in genes related to these pathways will likely disrupt the dorso-ventral gradient of the SHH morphogen (Fernandes and Hébert, 2008), leading to an incomplete

cleavage of the forebrain and, ultimately, to HPE. In this model, distinction between different manifestations of HPE lies in the degree of overall functional impact on SHH signalling (Mercier *et al.*, 2013). Moreover, depending on the affected genes and pathways, HPE patients would present different secondary clinical features.

The observed overlapping secondary clinical features further support the causality of the reported variants for HPE. As hypomorphic mutations do not have the same impact as the complete inactivation of a gene in most cases, phenotypic overlaps may be challenging to detect and require expert assessment of clinical and biological data. For example, mice deficient in *NDST1* exhibit agnathia (Grobe, 2005) (absence of the lower jaw) while unrelated patients presenting candidate variants in *NDST1* exhibit respectively prognathia and retrognathia (abnormal positioning of the lower jaw). All three phenotypes are part of the same spectrum of mandibular anomalies. From a clinical perspective, overlap of secondary clinical features between the patient and the animal models provides additional critical evidence of a causal relationship between candidate gene and disease. Key issue here remains the semantic representation of patient's phenotype and the use of a well-established phenotypic ontology during the examination processes. Explorations of secondary clinical features should be performed in future studies of genetic diseases.

Additional molecular screenings in larger populations of HPE patients are necessary to definitely assess the implication of our candidate genes in the disease. Therefore, we propose to include these novel genes into future genetic screenings of HPE patients.

In conclusion, this paper presents novel genes implicated in HPE and illustrates that HPE presents an oligogenic inheritance pattern requiring the joint effect of multiple genetic variants acting as hypomorphic mutations. The proposed inheritance pattern accounts for a

wide clinical spectrum of HPE and explains the significant part of cases in which no molecular diagnosis could be established by conventional approaches. It also explains the incomplete penetrance and variable expressivity of inherited causal mutations observed in the reported cases of HPE (Mercier *et al.*, 2011). We propose that in cases of non-Mendelian diseases with variable phenotypes, the possibility of oligogenic inheritance needs to be evaluated. Exploration of such events will improve the diagnostic yield of complex developmental disorders and will contribute to better understand the mechanisms that coordinate normal and pathological embryonic development.

Acknowledgements.

This work was supported by La Fondation Maladie Rares and the Agence de la Biomedecine. The authors acknowledge the Centre de Ressources Biologiques (CRB)-Santé (http://www.crbsante-rennes.com) of Rennes for managing patient samples. We would like to thank the families for their participation in the study, all clinicians who referred HPE cases, the eight CLAD (Centres Labellisés pour les Anomalies du Développement) within France that belong to FECLAD, French centers of prenatal diagnosis (CPDPN) and the SOFFOET for fetus cases, and the "filière AnDDI-Rares." We particularly thank all members of the Molecular Genetics Laboratory (CHU, Rennes) and of the Department of Genetics and Development (UMR6290 CNRS, Université Rennes 1) for their help and advice.

This Work was supported by France Génomique National infrastructure, funded as part of "Investissement d'avenir" program managed by Agence Nationale pour la Recherche (contrat ANR-10-INBS-09) https://www.france-genomique.org/spip/spip.php?article158

This study makes use of data generated by the Genome of the Netherlands Project. Funding for the project was provided by the Netherlands Organization for Scientific Research under

award number 184021007, dated July 9, 2009 and made available as a Rainbow Project of the Biobanking and Biomolecular Research Infrastructure Netherlands (BBMRI-NL). Samples where contributed by LifeLines (http://lifelines.nl/lifelines-research/general), The Leiden Longevity Study (http://www.healthy-ageing.nl; http://www.langleven.net), The Netherlands Twin Registry (NTR: http://www.tweelingenregister. org), The Rotterdam studies, (http://www.erasmus-epidemiology.nl/rotterdamstudy) and the Genetic Research in Isolated Populations program (http://www.epib.nl/research/geneticepi/research.html#gip). The sequencing was carried out in collaboration with the Beijing Institute for Genomics (BGI).

Funding.

This work was supported by Fondation Maladie Rares (grant PMO1201204), Agence Nationale de la Recherche (grant ANR-12-BSV1-0007-01) and the Agence de la Biomedecine (AMP2016).

Appendix. Collaborators

The FREX Consortium

Principal Investigators:
Emmanuelle Génin (chair), Inserm UMR1078, CHRU, Univ Brest, Brest, France
Dominique Campion, Inserm UMR1079, Faculté de Médecine, Rouen, France
Jean-François Dartigues, Inserm UMR1219, Univ Bordeaux, France
Jean-François Deleuze, Centre National de Génotypage, CEA, Fondation Jean Dausset-CEPH, Evry, France
Jean-Charles Lambert, Inserm UMR1167, Institut Pasteur, Lille, France
Richard Redon, Inserm UMR 1087 / CNRS UMR 6291, l'institut du thorax, Nantes, France

Collaborators:

Bioinformatics group:
Thomas Ludwig (chair), Inserm UMR1078, CHRU, Univ Brest, Brest
Benjamin Grenier-Boley, Inserm UMR1167, Institut Pasteur, Lille
Sébastien Letort, Inserm UMR1078, CHRU, Univ Brest, Brest
Pierre Lindenbaum, Inserm UMR 1087 / CNRS UMR 6291, l'institut du thorax, Nantes
Vincent Meyer, Centre National de Génotypage, CEA, Evry
Olivier Quenez, Inserm UMR1079, Faculté de Médecine, Rouen

Statistical genetics group:
Christian Dina (chair), Inserm UMR 1087/CNRS UMR 6291, l'institut du thorax, Nantes
Céline Bellenguez, Inserm UMR1167, Institut Pasteur, Lille

Camille Charbonnier-Le Clézio, Inserm UMR1079, Faculté de Médecine, Rouen
Joanna Giemza, Inserm UMR 1087 / CNRS UMR 6291, l'institut du thorax, Nantes

Data collection:
Stéphanie Chatel, Inserm UMR 1087 / CNRS UMR 6291, l'institut du thorax, Nantes
Claude Férec, Inserm UMR1078, CHRU, Univ Brest
Hervé Le Marec, Inserm UMR 1087 / CNRS UMR 6291, l'institut du thorax, Nantes
Luc Letenneur, Inserm UMR1219, Univ Bordeaux
Gaël Nicolas, Inserm UMR1079, Faculté de Médecine, Rouen
Karen Rouault, Inserm UMR1078, CHRU, Univ Brest

Sequencing:
Delphine Bacq, Centre National de Génotypage, CEA, Evry
Anne Boland, Centre National de Génotypage, CEA, Evry
Doris Lechner, Centre National de Génotypage, CEA, Evry

Genome of the Netherlands Consortium:

Steering group. Cisca Wijmenga[1,2] (principal investigator), Morris A. Swertz[1–3], P. Eline Slagboom[4], Gert-Jan B. van Ommen[5], Cornelia M. van Duijn[6], Dorret I. Boomsma[7], Paul I.W. de Bakker[8–11]

Ethical, legal, and social issues. Jasper A. Bovenberg[12]

Cohort collection and sample management. P. Eline Slagboom[4], Anton J.M. de Craen[4], Marian Beekman[4], Albert Hofman[6], Dorret I. Boomsma[7], Gonneke Willemsen[7], Bruce Wolffenbuttel[13], Mathieu Platteel[1]

Sequencing. Yuanping Du[14], Ruoyan Chen[14], Hongzhi Cao[14], Rui Cao[14], Yushen Sun[14], Jeremy Sujie Cao[14]

Analysis group. Morris A. Swertz[1–3] (Co-Chair), Freerk van Dijk[1,2], Pieter B.T. Neerincx[1,2], Patrick Deelen[1,2], Martijn Dijkstra[1,2], George Byelas[1,2], Alexandros Kanterakis[1,2], Jan Bot[15], Kai Ye[4], Eric-Wubbo Lameijer[4], Martijn Vermaat[3,5,16], Jeroen F.J. Laros[3,5,16], Johan T. den Dunnen[5,16], Peter de Knijff[5], Lennart C. Karssen[6], Elisa M. van Leeuwen[6], Najaf Amin[6], Vyacheslav Koval[17], Fernando Rivadeneira[17], Karol Estrada[17], Jayne Y. Hehir-Kwa[18], Joep de Ligt[18], Abdel Abdellaoui[7], Jouke-Jan Hottenga[7], V. Mathijs Kattenberg[3,7], David van Enckevort[3], Hailiang Mei[3], Mark Santcroos[19], Barbera D.C. van Schaik[19], Robert E. Handsaker[11,20], Steven A. McCarroll[11,20], Evan E. Eichler[21], Arthur Ko[21], Peter Sudmant[21], Laurent C. Francioli[8], Wigard P. Kloosterman[8], Isaac J. Nijman[8], Victor Guryev[22], Paul I.W. de Bakker[8–11] (Co-Chair)

1. Department of Genetics, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands

2. Genomics Coordination Center, University Medical Center Groningen and University of Groningen, Groningen, The Netherlands

3. Netherlands Bioinformatics Centre, Nijmegen, The Netherlands

4. Section Molecular Epidemiology, Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Leiden, The Netherlands

5. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

6. Department of Epidemiology, Erasmus Medical Center, Rotterdam, The Netherlands

7. Department of Biological Psychology, VU University, Amsterdam, The Netherlands

8. Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands

9. Department of Epidemiology, University Medical Center Utrecht, Utrecht, The Netherlands

10. Division of Genetics, Brigham and Women's Hospital, Harvard Medical School, Boston, Massachusetts

11. Broad Institute of Harvard and MIT, Cambridge, Massachusetts

12. Legal Pathways Institute for Health and Bio Law, Aerdenhout, The Netherlands

References

Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, et al. Exome sequencing as a tool for Mendelian disease gene discovery. Nat. Rev. Genet. 2011; 12: 745–755.

Bear KA, Solomon BD, Antonini S, Arnhold IJP, França MM, Gerkes EH, et al. Pathogenic mutations in GLI2 cause a specific phenotype that is distinct from holoprosencephaly. J. Med. Genet. 2014; 51: 413–418.

Briggs KJ, Corcoran-Schwartz IM, Zhang W, Harcke T, Devereux WL, Baylin SB, et al. Cooperation between the Hic1 and Ptch1 tumor suppressors in medulloblastoma. Genes Dev. 2008; 22: 770–785.

Carter MG. Mice deficient in the candidate tumor suppressor gene Hic1 exhibit developmental defects of structures affected in the Miller-Dieker syndrome. Hum. Mol. Genet. 2000; 9: 413–419.

Christ A, Christa A, Kur E, Lioubinski O, Bachmann S, Willnow TE, et al. LRP2 is an auxiliary SHH receptor required to condition the forebrain ventral midline for inductive signals. Dev. Cell 2012; 22: 268–278.

Ciani L, Patel A, Allen ND, ffrench-Constant C. Mice lacking the giant protocadherin mFAT1 exhibit renal slit junction abnormalities and a partially penetrant cyclopia and anophthalmia phenotype. Mol. Cell. Biol. 2003; 23: 3575–3582.

Dowdle WE, Robinson JF, Kneist A, Sirerol-Piquer MS, Frints SGM, Corbit KC, et al. Disruption of a Ciliary B9 Protein Complex Causes Meckel Syndrome. Am. J. Hum. Genet. 2011; 89: 94–110.

Dubourg C, Carré W, Hamdi-Rozé H, Mouden C, Roume J, Abdelmajid B, et al. Mutational Spectrum in Holoprosencephaly Shows That FGF is a New Major Signaling Pathway. Hum. Mutat. 2016; 37: 1329–1339.

Dubourg C, Kim A, Watrin E, de Tayrac M, Odent S, David V, et al. Recent advances in understanding inheritance of holoprosencephaly. Am. J. Med. Genet. C Semin. Med. Genet. 2018

Dupé V, Rochard L, Mercier S, Le Pétillon Y, Gicquel I, Bendavid C, et al. NOTCH, a new signaling pathway implicated in holoprosencephaly. Hum. Mol. Genet. 2011; 20: 1122–1131.

Fernandes M, Hébert JM. The ups and downs of holoprosencephaly: dorsal versus ventral patterning forces. Clin. Genet. 2008; 73: 413–423.

Garcia-Gonzalo FR, Corbit KC, Sirerol-Piquer MS, Ramaswami G, Otto EA, Noriega TR, et al. A transition zone complex regulates mammalian ciliogenesis and ciliary membrane composition. Nat. Genet. 2011; 43: 776–784.

Genome of the Netherlands Consortium. Whole-genome sequence variation, population

structure and demographic history of the Dutch population. Nat. Genet. 2014; 46: 818–825.

Goetz SC, Ocbina PJR, Anderson KV. The Primary Cilium as a Hedgehog Signal Transduction Machine. Methods Cell Biol. 2009; 94: 199–222.

Gorivodsky M, Mukhopadhyay M, Wilsch-Braeuninger M, Phillips M, Teufel A, Kim C, et al. Intraflagellar Transport Protein 172 is essential for primary cilia formation and plays a vital role in patterning the mammalian brain. Dev. Biol. 2009; 325: 24–32.

Grobe K. Cerebral hypoplasia and craniofacial defects in mice lacking heparan sulfate Ndst1 gene function. Development 2005; 132: 3777–3786.

Hong M, Srivastava K, Kim S, Allen BL, Leahy DJ, Hu P, et al. BOC is a modifier gene in holoprosencephaly. Hum. Mutat. 2017; 38: 1464–1470.

Jakobs P, Exner S, Schürmann S, Pickhinke U, Bandari S, Ortmann C, et al. Scube2 enhances proteolytic Shh processing from the surface of Shh-producing cells. J. Cell Sci. 2014; 127: 1726–1737.

Kruszka P, Martinez AF, Muenke M. Molecular testing in holoprosencephaly. Am. J. Med. Genet. C Semin. Med. Genet. 2018

Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008; 9: 559.

Lee H, Deignan JL, Dorrani N, Strom SP, Kantarci S, Quintero-Rivera F, et al. Clinical exome sequencing for genetic identification of rare Mendelian disorders. JAMA 2014; 312: 1880–1887.

Leung AWL, Wong SYY, Chan D, Tam PPL, Cheah KSE. Loss of procollagen IIA from the anterior mesendoderm disrupts the development of mouse embryonic forebrain. Dev. Dyn. 2010; 239: 2319–2329.

Li L, Bainbridge MN, Tan Y, Willerson JT, Marian AJ. A Potential Oligogenic Etiology of Hypertrophic Cardiomyopathy: A Classic Single-Gene Disorder. Circ. Res. 2017; 120: 1084–1090.

Lindsay SJ, Xu Y, Lisgo SN, Harkin LF, Copp AJ, Gerrelli D, et al. HDBR Expression: A Unique Resource for Global and Individual Gene Expression Studies during Early Human Brain Development [Internet]. Front. Neuroanat. 2016; 10[cited 2017 May 2] Available from: http://journal.frontiersin.org/article/10.3389/fnana.2016.00086/full

MacArthur DG, Manolio TA, Dimmock DP, Rehm HL, Shendure J, Abecasis GR, et al. Guidelines for investigating causality of sequence variants in human disease. Nature 2014; 508: nature13127.

Mercier S, David V, Ratié L, Gicquel I, Odent S, Dupé V. NODAL and SHH dose-dependent double inhibition promotes an HPE-like phenotype in chick embryos. Dis. Model. Mech. 2013; 6: 537–543.

Mercier S, Dubourg C, Garcelon N, Campillo-Gimenez B, Gicquel I, Belleguic M, et al. New findings for phenotype-genotype correlations in a large European series of holoprosencephaly cases. J. Med. Genet. 2011; 48: 752–760.

Mouden C, Dubourg C, Carré W, Rose S, Quelin C, Akloul L, et al. Complex mode of inheritance in holoprosencephaly revealed by whole exome sequencing. Clin. Genet. 2016; 89: 659–668.

Mouden C, Tayrac M de, Dubourg C, Rose S, Carré W, Hamdi-Rozé H, et al. Homozygous STIL Mutation Causes Holoprosencephaly and Microcephaly in Two Siblings. PLOS ONE 2015; 10: e0117418.

Murdoch JN, Copp AJ. The relationship between Sonic hedgehog signalling, cilia and neural tube defects. Birt. Defects Res. A. Clin. Mol. Teratol. 2010; 88: 633–652.

Philippakis AA, Azzariti DR, Beltran S, Brookes AJ, Brownstein CA, Brudno M, et al. The Matchmaker Exchange: a platform for rare disease gene discovery. Hum. Mutat. 2015; 36: 915–921.

Puppo F, Dionnet E, Gaillard M-C, Gaildrat P, Castro C, Vovan C, et al. Identification of Variants in the 4q35 Gene FAT1 in Patients with a Facioscapulohumeral Dystrophy-Like Phenotype. Hum. Mutat. 2015; 36: 443–453.

Reed DR, Lawler MP, Tordoff MG. Reduced body weight is a common effect of gene knockout in mice. BMC Genet. 2008; 9: 4.

Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. Off. J. Am. Coll. Med. Genet. 2015; 17: 405–424.

Rock R, Schrauth S, Gessler M. Expression of mouse dchs1, fjx1, and fat-j suggests conservation of the planar cell polarity pathway identified in drosophila. Dev. Dyn. 2005; 234: 747–755.

Roessler E, Belloni E, Gaudenz K, Jay P, Berta P, Scherer SW, et al. Mutations in the human Sonic Hedgehog gene cause holoprosencephaly. Nat. Genet. 1996; 14: 357–360.

Smedley D, Jacobsen JOB, Jäger M, Köhler S, Holtgrewe M, Schubach M, et al. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nat. Protoc. 2015; 10: 2004–2015.

Smith CL, Blake JA, Kadin JA, Richardson JE, Bult CJ, Mouse Genome Database Group. Mouse Genome Database (MGD)-2018: knowledgebase for the laboratory mouse. Nucleic Acids Res. 2018; 46: D836–D842.

Solomon BD, Gropman A, Muenke M. Holoprosencephaly Overview [Internet]. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, et al., editor(s). GeneReviews®. Seattle (WA): University of Washington, Seattle; 1993. [cited 2018 Mar 11] Available from:

http://www.ncbi.nlm.nih.gov/books/NBK1530/

Stark Z, Dashnow H, Lunke S, Tan TY, Yeung A, Sadedin S, et al. A clinically driven variant prioritization framework outperforms purely computational approaches for the diagnostic analysis of singleton WES data. Eur. J. Hum. Genet. EJHG 2017; 25: 1268–1272.

Thomas S, Legendre M, Saunier S, Bessières B, Alby C, Bonnière M, et al. TCTN3 mutations cause Mohr-Majewski syndrome. Am. J. Hum. Genet. 2012; 91: 372–378.

Wheway G, Abdelhamed Z, Natarajan S, Toomes C, Inglehearn C, Johnson CA. Aberrant Wnt signalling and cellular over-proliferation in a novel mouse model of Meckel–Gruber syndrome. Dev. Biol. 2013; 377: 55–66.

Xavier GM, Cobourne MT. Scube2 expression extends beyond the central nervous system during mouse development. J. Mol. Histol. 2011; 42: 383–391.

Yang T, Jia Z, Bryant-Pike W, Chandrasekhar A, Murray JC, Fritzsch B, et al. Analysis of PRICKLE1 in human cleft palate and mouse development demonstrates rare and common variants involved in human malformations. Mol. Genet. Genomic Med. 2014; 2: 138–151.

Zhang W, Kang J-S, Cole F, Yi M-J, Krauss RS. Cdo functions at multiple points in the Sonic Hedgehog pathway, and Cdo-deficient mice accurately model human holoprosencephaly. Dev. Cell 2006; 10: 657–665.

Figures. ~~Legends~~

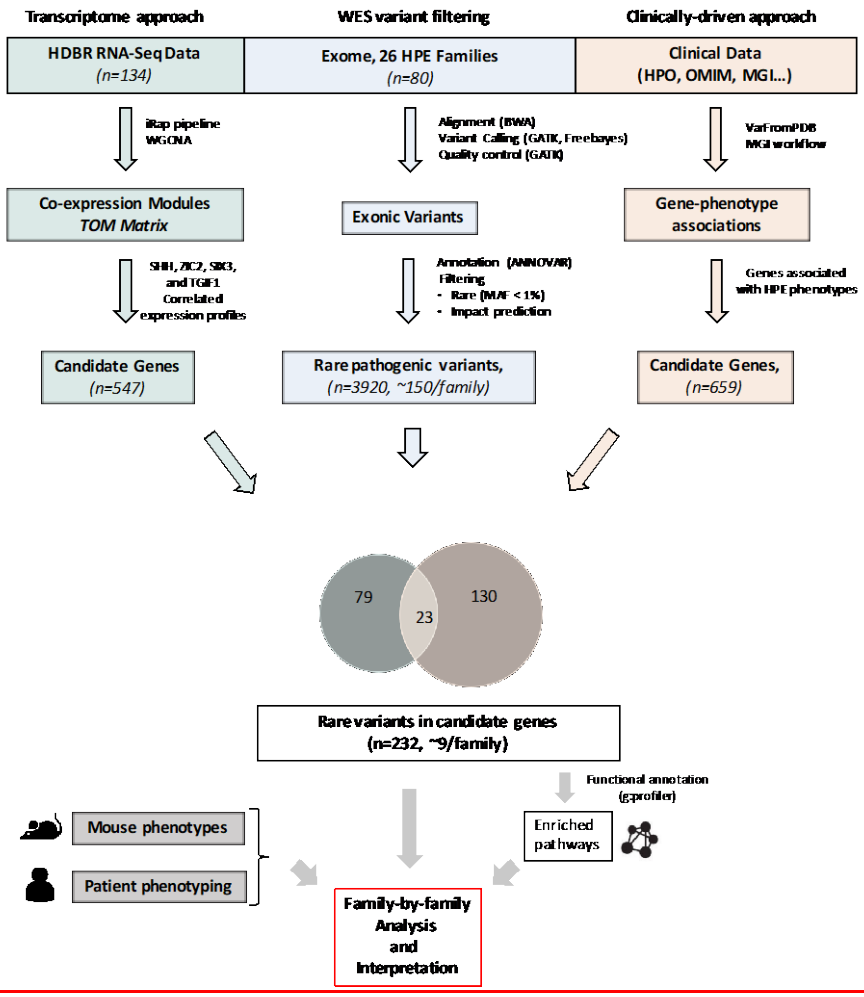

Figure 1. Flow chart illustrating the prioritization strategy.

Classical WES analysis was performed (blue) and combined with two prioritization approaches: (1) based on gene co-expression networks (green) and (2) based on clinical knowledge (salmon). Details of the pipeline are also provided in the Supplementary Appendix.

Variant overlaps were selected and further analysed by functional annotation analysis and on a family-by-family basis, by integrating a comprehensive clinical phenotyping of patients and exploration of cross-species similarities.

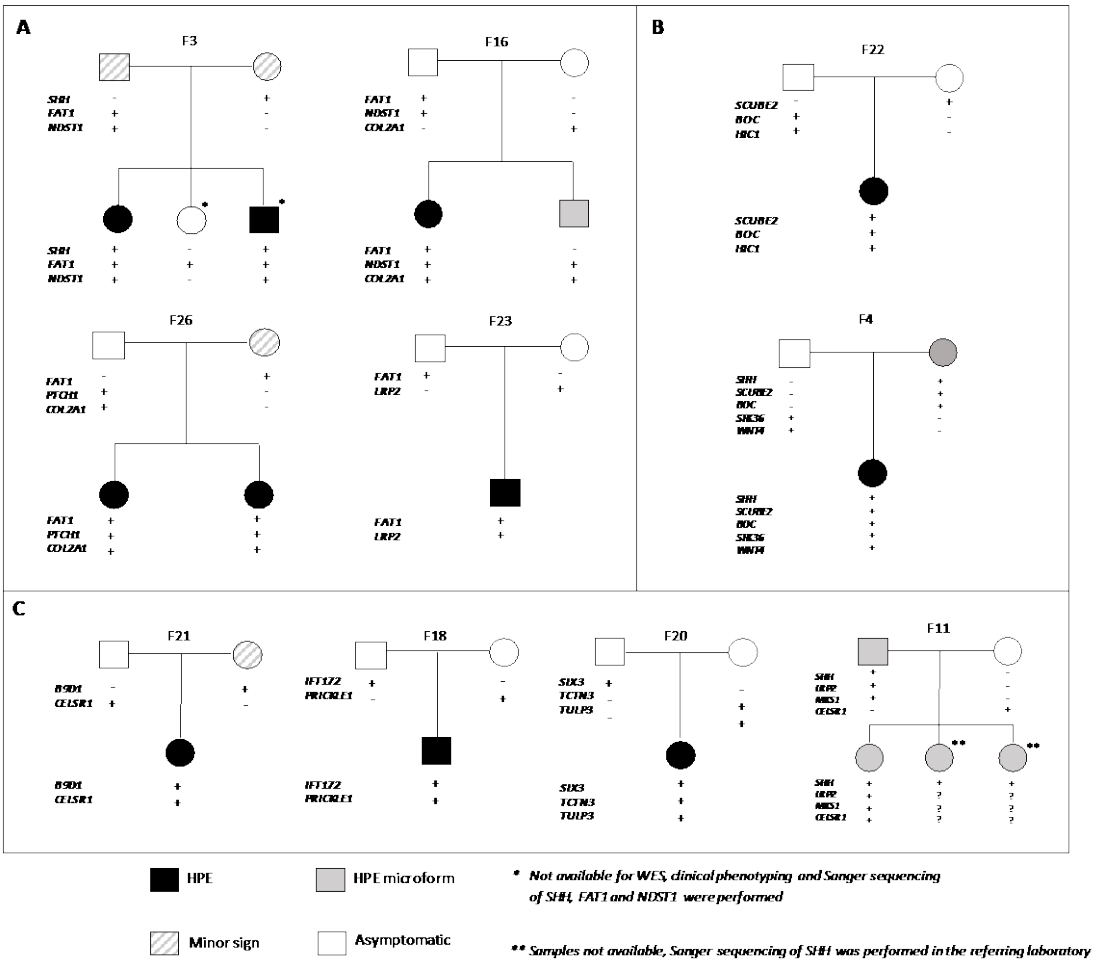

Figure 2. Oligogenic events reported in this study Candidate genes are listed for each family. Individuals carrying/not carrying the variants are identified respectively by the +/-. Variants information is available in Table 2 and Table S6 of the Supplementary Appendix.

(**A)** Oligogenic events involving *FAT1*. **(B)** Oligogenic events involving variants in *SCUBE2* and *BOC.*(***C)*** Oligogenic events involving mutations in genes related to the primary cilium.

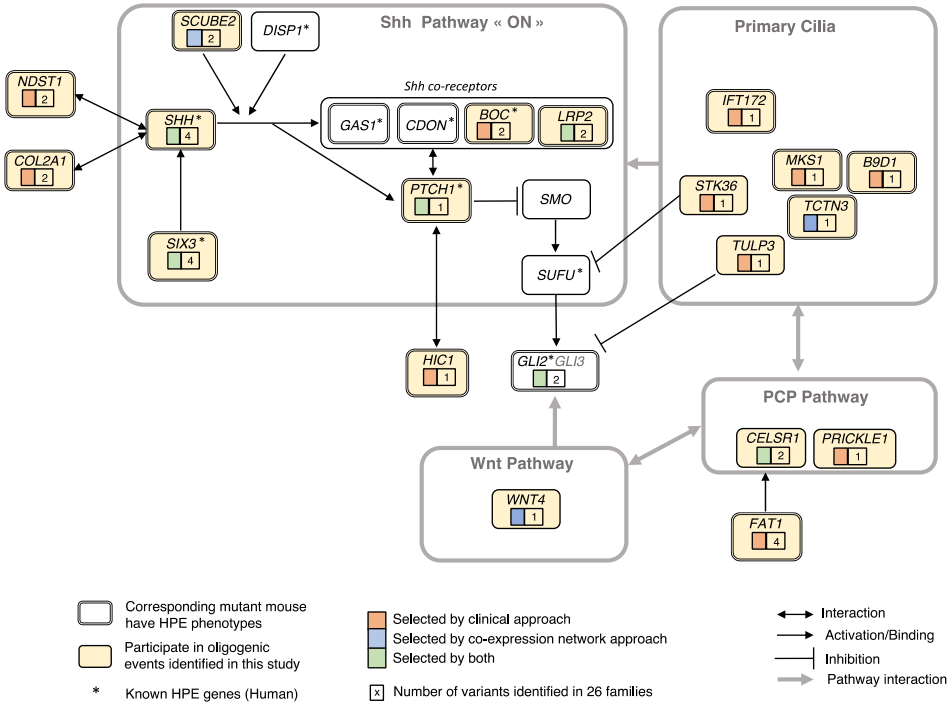

Figure 3 Implication of the candidate genes in the signaling pathways involved in HPE. Key affected pathways and genes are presented. Under each gene name, the selection methods (clinical or co-expression networks approach or both) is shown in the left panel and the number of variants for each gene is shown in the right panel. Genes known in HPE are marked with an asterisk, and genes for which corresponding mutant mouse have HPE phenotypes are surrounded by a double line. The genes implicated in an oligogenic events in the study are presented in yellow background.

Category	Feature	N	%
Proband Sex	Male	6	21
	Female	20	69
	Unkown	3	10
	TOTAL	**29**	**100**
Clinical Phenotype of the parents	Unaffected	40	78
	Minor sign	8	16
	hypotelorism	4	8
	incomplete iris	1	2
	epicanthus	1	2
	narrow palate	1	2
	nasal anomaly	1	2
	HPE microform	3	6
	TOTAL	**51**	**100**
Clinical Characteristics of the probands	HPE	29	100
	lobar	3	10
	semilobar	11	38
	alobar	13	45
	microform	2	7
	Cleft lip/palate	11	38
	Hypotelorism	10	34
	Microcephaly	9	31
	Arhinencephaly	9	31
	Agenesis of corpus callosum	7	24
	Flat head (plagiocephaly)	6	21
	Thalami Fusion	6	21
	Ventricles Fusion	6	21
	Premaxilliary agenesia	5	17
	Fusion frontal lobes	4	14
	Flat nose	4	14
	Proboscis	3	10
	Cyclopia	2	7
	TOTAL	**29**	**100**
Families with mutations in HPE genes	SHH	4	15,4
	ZIC2	1	3,8
	SIX3	5*	19,2
	TGIF1	2	7,7
	PTCH1	1	3,3
	ZIC2/GLI2	1	3,8
	No mutation	12	46,2
	TOTAL	**26**	**100,0**
Families Ethnicity	European	21	81
	African	1	4
	South Asian	1	4
	Admix	3	12
	TOTAL	26	100,0

Table 1. Clinical description of 26 HPE families.

* For SIX3, point mutations were found in four families (targeted sequencing) and a heterozygous deletion was detected by CGH-array in one family.

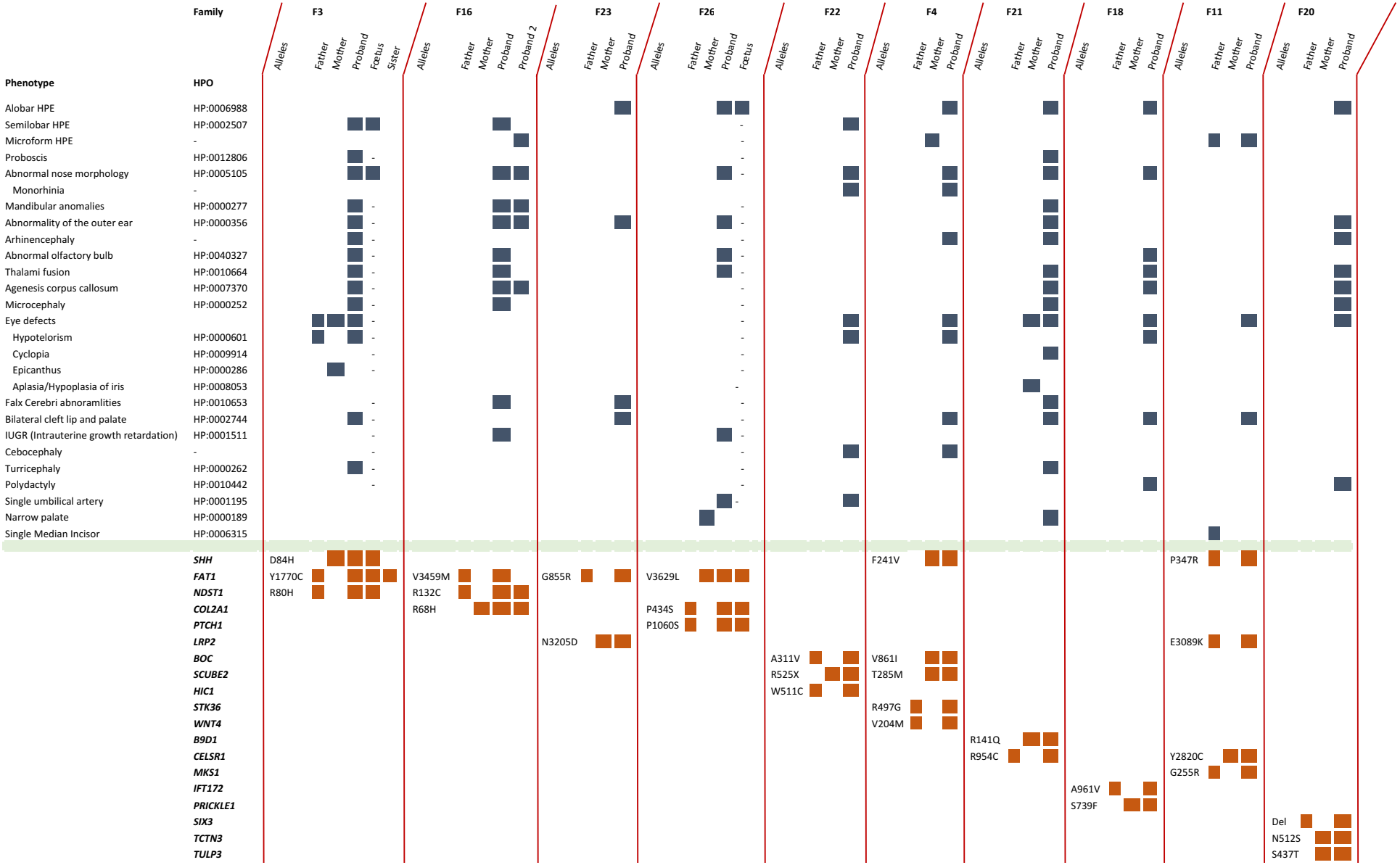
Family
F3
F16
F23
F26
F22
F4
F21
F18
F11
F20
Alleles
Father
Mother
Proband
Fœtus
Sister
Proband 2
Phenotype
HPO
Alobar HPE HP:0006988
Semilobar HPE HP:0002507
Microform HPE -
Proboscis HP:0012806
Abnormal nose morphology HP:0005105
Monorhinia -
Mandibular anomalies HP:0000277
Abnormality of the outer ear HP:0000356
Arhinencephaly -
Abnormal olfactory bulb HP:0040327
Thalami fusion HP:0010664
Agenesis corpus callosum HP:0007370
Microcephaly HP:0000252
Eye defects
Hypotelorism HP:0000601
Cyclopia HP:0009914
Epicanthus HP:0000286
Aplasia/Hypoplasia of iris HP:0008053
Falx Cerebri abnoramlities HP:0010653
Bilateral cleft lip and palate HP:0002744
IUGR (Intrauterine growth retardation) HP:0001511
Cebocephaly -
Turricephaly HP:0000262
Polydactyly HP:0010442
Single umbilical artery HP:0001195
Narrow palate HP:0000189
Single Median Incisor HP:0006315
SHH
FAT1
NDST1
COL2A1
PTCH1
LRP2
BOC
SCUBE2
HIC1
STK36
WNT4
B9D1
CELSR1
MKS1
IFT172
PRICKLE1
SIX3
TCTN3
TULP3
D84H
Y1770C
R80H
V3459M
R132C
R68H
G855R
N3205D
V3629L
P434S
P1060S
A311V
R525X
W511C
F241V
V861I
T285M
R497G
V204M
R141Q
R954C
A961V
S739F
P347R
E3089K
Y2820C
G255R
Del
N512S
S437T

Table 2. Comparison of clinical features in the studied families. Occurrences of phenotypes are marked with blue squares for each individual. Unrelated families are separated by red lines. Hyphen is used when no observation was possible on foetuses. Heterozygous variants in the different genes are marked with orange squares. Human symbols indicate that the genes are known HPE disease genes. Mouse symbols indicate the existence of mouse mutant exhibiting HPE for the corresponding gene.

Comparison	HPE	GoNL	FREX	P-value HPE vs. GoNL	P-value HPE vs. FREX	P-Value GoNL vs. FREX
Families with oligogenic events	10/26 (38 %)	3/248 (1.2 %)	NA	2.301e-09	NA	NA
Children harbouring rare deleterious variants in >= 2 candidate genes	13/29 (45 %)	6/248 (2.4 %)	NA	1.902e-10	NA	NA
All individuals harbouring rare deleterious variants in >= 2 candidate genes	21/80 (26 %)	14/744 (1.8 %)	16/574 (2.7 %)	3.237e-14	1.521e-11	0.35

Table 3. Statistical validations. Fisher's exact test analysis for oligogenic events. Oligogenic inheritance is defined as presence of combined rare deleterious variants in ≥ 2 genes described in *Table 2/Figure 2*. The proportion of individuals harbouring combined rare deleterious variants in the identified genes is significantly higher in HPE cohort as compared to two control populations GoNL and FREX (Fisher's exact test).