TITLE

Models for infantile hypertrophic pyloric stenosis development in patients with esophageal atresia

Authors and affiliations

Chantal A. ten Kate[1], Rutger W.W. Brouwer[2], Yolande van Bever[3], Vera K. Martens[3], Tom Brands[3], Nicole W.G. van Beelen[1], Alice S. Brooks[3], Daphne Huigh[3], Bert J.F.M.M. Eussen[3], Wilfred F.J. van IJcken[2], Hanneke IJsselstijn[1], Dick Tibboel[1], Rene M.H. Wijnen[1], Annelies de Klein[3], Robert M.W. Hofstra[3], Erwin Brosens[3,#]

[1] Department of Pediatric Surgery and Intensive Care Children, Erasmus MC University Medical Center - Sophia Children's Hospital, Rotterdam, The Netherlands

[2] Center for Biomics, Erasmus MC University Medical Center, Rotterdam, The Netherlands

[3] Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands

Corresponding Author:

E. Brosens, PhD

Erasmus Medical Centre – Sophia Children's Hospital

Room Ee-987e

P.O. Box 2040

3000 CA Rotterdam

Tel +31 10 70 37643

Fax: +31 10 70 44736

e.brosens@erasmusmc.nl

ORCID ID: https://orcid.org/0000-0001-8235-4010

Running title: EA and IHPS: proposal for multifactorial hypothesis

Keywords: esophageal atresia, tracheoesophageal fistula, pyloric stenosis, VACTERL

Summary statement:

Instead of one affected gene, the higher incidence of IHPS in EA patients is more likely the result of multiple (epi)genetic and environmental factors together shifting the balance to disease development.

ABSTRACT

Patients born with esophageal atresia (EA) have a 30 times higher prevalence of infantile hypertrophic pyloric stenosis (IHPS). This makes sense from a developmental perspective as both the esophagus and the pyloric sphincter are foregut derived structures. EA and IHPS are variable features in several (monogenetic) syndromes. This, and twin and familial studies, indicates a genetic component for both conditions as single entities. We hypothesized that genetic defects, disturbing foregut morphogenesis, are responsible for this combination of malformations. Non-genetic factors could also contribute, as mice exposed to Adriamycin develop EA and *in utero* diethylstilbestrol exposure is associated with EA.

We investigated the copy number profiles and protein coding variants of 15 patients with both EA and IHPS. As all parents were unaffected, we first considered dominant (*de novo)* or recessive inheritance models but could not identify putatively deleterious mutations or recessive variants. We did identify inherited variants in genes either known to be involved in EA or IHPS or important in foregut morphogenesis in all patients. Unfortunately, variant burden analysis did not show a significant difference with unaffected controls. However, the IHPS associated risk SNP rs1933683 had a significantly higher incidence (OR 3.29, p=0.009).

Although the genetic variation in likely candidate genes as well as the predisposing locus near *BARX1* (rs1933683) suggest a genetic component, it does not fully explain the abnormalities seen in these patients. Therefore, we hypothesize that a combination of high impact genetic, mechanical and environmental factors together can shift the balance to abnormal development.

INTRODUCTION

Esophageal atresia (EA) is a rare congenital malformation caused by a faulty development of the foregut which leads to a discontinuity of the esophagus. It occurs in about 2.5 cases per 10,000 births within Europe (Pedersen et al., 2012, Oddsberg et al., 2012) and over three-quarters of patients present with a tracheoesophageal fistula (TEF) (Pedersen et al., 2012, Macchini et al., 2017). EA is considered – etiologically as well as phenotypically – a highly heterogeneous condition (Brosens et al., 2014). It can present either as an isolated defect but is often seen in combination with other malformations. Frequently, these malformations are part of the VACTERL (Vertebral, Anorectal, Cardiac, Tracheoesophageal, Renal or urinary tract of Limb malformations) association. VACTERL association is a diagnosis of exclusion in which three or more features of the VACTERL spectrum are present and a known genetic syndrome is not identified (Solomon et al., 2012). However, clustering of one or more of these features with additional specific associated malformations could also be the results of a shared genetic etiology.

One of the more prevalent, but less well-known, associated malformations is Infantile Hypertrophic Pyloric Stenosis (IHPS) (Rollins et al., 1989). In contrast to EA, IHPS is often considered an acquired disorder. The pyloric muscle hypertrophies in the first weeks of life, causing a narrowing of the pyloric channel (Panteli, 2009). Healthy-born infants present at week 3 to 6 of life with projectile postprandial vomiting. They need surgery where the upper layer of the circular smooth muscle of the pylorus will be incised, to release the passage from the stomach to the intestine again.

Previously, we described a 30 times higher prevalence (7.5%) of IHPS in patients with EA compared to the normal population (0.25%) (van Beelen et al., 2014). This increased prevalence has been reported in other retrospective studies (3.3-13%) as well (Palacios M.E.C. et al., 2014, Deurloo et al., 2002). The clinical presentation of IHPS seen in patients with EA/IHPS is not different from patients with isolated IHPS. However, the diagnosis of IHPS is more difficult and often delayed in patients with EA. Relatively common complications after EA repair, such as stenosis of the anastomosis, can protect against reflux and lead to just regurgitation. By the time these patients start vomiting, there must be massive gastroesophageal reflux.

The presentation of both EA and IHPS makes sense from a developmental perspective as the esophagus and the pyloric sphincter are both foregut derived structures. Organ specification during embryonic development is under tight spatiotemporal control of specific growth factors, transcription factors and signaling cascades (Li et al., 2009, Jacobs and Que, 2013). Disturbances in these pathways could impact proper development. In mice, the esophagus is specified from the foregut tube between embryonic day E9.5 and E11.5. In humans, the esophagus, as well as the

stomach, starts developing from the fourth week after conception onwards. The stomach turns around its anterior-posterior axis during embryonic development (Cetin et al., 2006). The developing pylorus can be visualized with immunostaining at week six after gestation and differentiates during fetal life (Koyuncu et al., 2009).

Environmental (Zwink et al., 2016, Felix et al., 2008, Feng et al., 2016, Markel et al., 2015, Krogh et al., 2012, Sorensen et al., 2002) and genetic contributions (Peeters et al., 2012, Brosens et al., 2014, Solomon et al., 2012) have been described for both EA and IHPS as single entities or in combination with other anatomical malformations. It has been suggested that *in utero* exposure to diethylstilbestrol (DES) is associated with the development of EA (Felix et al., 2007a). Moreover, both malformations are variable features in specific and often phenotypically overlapping genetic syndromes (Table 1). The presence of both conditions as variable features in the phenotypical spectrum of known genetic syndromes is indicative of a genetic background for EA and IHPS. More evidence for a genetic contribution can be deduced from twin studies and animal models (de Jong et al., 2010). The concordance rates in monozygotic twins compared to dizygotic twins is higher for EA (Veenma et al., 2012) and IHPS (Krogh et al., 2010) as single entities. Also, the recurrence risk is elevated for siblings and offspring of affected individuals with EA in combination with other associated anomalies (Robert et al., 1993, Van Staey et al., 1984, Warren et al., 1979, McMullen et al., 1996). In contrast, the recurrence risk for isolated EA is low (Schulz et al., 2012) and moderate for IHPS (Krogh et al., 2010, Elinoff et al., 2005). In contrast to EA, there has been reported a male predominance for IHPS (4:1) (MacMahon, 2006). There have been risk loci associated to IHPS (Everett and Chung, 2013, Feenstra et al., 2012, Feenstra et al., 2013, Svenningsson et al., 2012, Fadista et al., 2019). To date, no risk loci have been described for EA.

Considering the increased prevalence of IHPS in patients with EA, their common developmental origin and previous evidence in genetic studies, we hypothesized that Copy Number Variants (CNVs) or other protein coding alterations affecting one specific gene, or genetic disturbances in more genes all important for foregut morphogenesis are responsible for the higher incidence of IHPS in patients with EA.

RESULTS

Patient cohort

In total, 27 out of 664 patients (4.1%) born with EA between 1970-2017, developed IHPS. Twenty patients have been described previously (van Beelen et al., 2014). Parental informed consent for

whole exome sequencing (WES) was obtained for 15 patients. Several phenotypical characteristics stood out in this EA/IHPS cohort: a sacral dimple was present in seven patients (25.9%), anomalies of the vertebrae or ribs in eight patients (29.7%) and genitourinary anomalies in six patients (22.2%) of which two patients (7.4%) had hypospadias. Four patients (14.8%) had three or more anomalies within the VACTERL spectrum (Solomon, 2011). A full phenotypical description of the 27 EA/IHPS patients is given in Table 2.

Copy Number analysis

Our previous study described rare CNVs and their inheritance pattern in patients with EA (Brosens et al., 2016b), seventeen EA/IHPS patients were included in this previous study. None of the six large CNVs identified were *de novo*, all were inherited from one of the unaffected parents. Patient SKZ_400 had a paternal inherited rare gain of chromosomal region 11q15. Patient SKZ_0887 had maternal inherited putative deleterious gains on Xq26.1 and Xp22.33. Patient SKZ_1003 had a maternal inherited loss of chromosomal region 17q11 and patient SKZ_1248 maternal inherited rare gains in chromosomal regions 4q35 and 5p15.1. Additional exon-level CN-profiling using the normalized coverage profiles (Amarasinghe et al., 2013) of the exome sequencing data confirmed the presence of the CNV seen with SNP-array. All CN profiles of main EA and IHPS disease genes (Brosens et al., 2014, Peeters et al., 2012) were normal. There were no overlapping rare CNVs in this patient cohort. All rare CNVs, classified as VUS or (likely) deleterious are described in Table S1.

Exome sequence analysis

Sequencing resulted in at least 5 Giga-bases of raw sequence data with an average coverage of 70X and 90% of target bases covered over 20X. Quality of the sequence data is listed in Table S2. As none of the parents of the 15 investigated patients were affected we first considered dominant *de novo* and recessive modes of inheritance.

We could not identify *de novo* pathogenic variation in main EA and IHPS disease genes (Brosens et al., 2014, Peeters et al., 2012). Subsequently, we searched for possible *de novo* mutations exome wide. For this, we focused on putative deleterious ultra-rare protein coding or splice site variants (n=100) (Bennett et al., 2017). Variations were considered ultra-rare when they were absent in the gnomAD dataset (123,136 whole exomes and 15,496 whole genomes) (http://gnomad.broadinstitute.org/) (Lek et al., 2016). Twenty-five variants proved to be sequencing artifacts. Furthermore, we could not confirm the segregation of 15 mutations due to lack of parental DNA. We determined the segregation of all ultra-rare variants predicted to be of unknown significance (VUS, n=37) or (likely) deleterious (n=23). All putative deleterious variants tested proved to be inherited from one of the unaffected parents.

Considering a recessive mode of inheritance, we searched for genes with homozygous or compound heterozygous variants. Six variants in three genes (*FLNC*, *ATP6V0A1* and *FAM46A*) fitted a putative compound heterozygous model, two genes (*KCNN3* and *VDAC3*) had homozygous variants and two genes (*MID2* and *SH3KBP1*) had variants on chromosome X in a male patient. All variants were predicted to be likely deleterious or VUS and intolerant to missense variants (Z-score ≥3) or loss of function variants (PLI or PLIrec ≥0.9). With segregation analysis, we could confirm the compound heterozygous mode of inheritance of the variant in the *FAM46A* gene in patient SKZ_2023 and the maternally inherited X-linked variant in the *SK3KBP1* gene in patient SKZ_1260. The other recessive candidate genes could not be validated due to technical difficulties (and are likely sequencing artifacts) or due to lack of parental DNA. None of the recessive candidate genes were affected twice or more in this cohort. All predicted deleterious variants were submitted to the ClinVar database https://www.ncbi.nlm.nih.gov/clinvar/ (Landrum et al., 2014).

We inspected the CN profiles from WES-CN and SNP-array for partial overlap with genes affected by heterozygous variant predicted to be deleterious in (recessive) loss of function intolerant or missense intolerant genes (n=48) and could not detect unmasking of a recessive mutation by a CNV. Ultra-rare variants (n=78), X-linked or recessive variants are depicted in Table S4 and uploaded to the ClinVar database (https://www.ncbi.nlm.nih.gov/clinvar/)

Pathway enrichment analysis of genes affected by rare variants

When looking at the selected protein altering variants (Z-score ≥3, n=44) or loss of function intolerant (PLI ≥0.9, n=4), two relevant pathways were significantly enriched (p-value $<1x10^{-5}$): proliferation and differentiation of smooth muscle cells (*INSR*, *ITGB1*, *NOTCH1*, *TCF4*, *PDE4D*, *TERT*, *ANKRD17*, *DICER1*) and self-renewal of satellite cells (*ITGB1*, *NOTCH1*).

Variant prioritization using different in silico tools

We prioritized all rare variants with three in silico tools (see Methods section). Fifty-four variants in 34 genes had an overlap between VAAST (Yandell et al., 2011, Hu et al., 2013, Kennedy et al., 2014)), which prioritizes based on variant deleteriousness and Phevor and PhenIX which prioritize more on phenotype (Singleton et al., 2014, Zemojtel et al., 2014)). Top ranking variants can be found in Table S3.

Additionally, we found variants in the same gene in multiple patients (Fig. 1). Of these 116 genes (VUS=87, likely deleterious=30), 36 genes were found in ≥3 patients of which six genes were present in more than five patients (*CNTN2*, *DSPP*, *NOTCH4*, *PRRC2A*, *SEC16B*, *ZNF717*). Four (*AMBRA1*, *ATP2A3*, *DSCAM*, *NOTCH1*) out of 116 genes were predicted to be intolerant for missense variants (Z-score ≥3). See also Table S4.

Gene burden analysis

An exome wide gene burden analysis showed ten genes which were enriched for rare putatively deleterious variation compared to the 1000 Genomes project phase 3 samples. The results are shown in Table 3. There were no genes with more than two distinct variants. Each variant was observed only once. A second burden test – only evaluating genes from developmental important pathways and known disease genes – showed no significant difference between our 15 patients and a control group of 44 healthy individuals, who were sequenced in a previous study (Table 4). Also, the number of putative deleterious variants between these two groups was not significantly different (Table 5). Unfortunately, a burden test comparing the variant profiles of these genes between the patients and their parents was not possible since no WES data of the parents was available.

Expression of main candidate gene during development

With public micro-array transcriptome data we evaluated which genes were upregulated at a specific time-point in the foregut, esophagus or pyloric sphincter and used the output as an indicator of gene expression (see Methods section and Table S5). Of the genes classified as VUS or likely deleterious in our exome sequencing results, 28 genes were upregulated in both the foregut or esophagus as well as the pyloric sphincter: *ADAMTSL4*, *AGRN*, *ANKRD29*, *ARHGAP29*, *CAMTA1*, *CDHR5*, *CNTN2*, *COL11A1*, *DNAJC11*, *HIVEP3*, *HMCN1*, *HMGCS2*, *HSPG2*, *ITGB3BP*, *LDB3*, *MYOF*, *NKX2-3*, *NUP133*, *PCSK9*, *PKN2*, *PRDM16*, *PUM1*, *RET*, *SEC16B*, *SERINC2*, *TMEM82*, *VPS13D* and *ZBTB7B*.

Unfortunately, none of the genes enriched in our burden analysis were differentially expressed in mice foregut between E8.5 and E16.5. Seven out of 116 genes with putative deleterious variants in more than one patient were differentially expressed in mice foregut: *Adamtsl4* at E8.5, E14.5 and E16.5; *Ankrd26* at E14.5; *Cntn2* at E8.5, E15.5 and E18.5; *Hspg2* at E8.25, E8.5, E14.5 and E18.5; *Kcnn3* at E8.5 and E15.5; *Ldb3* at E8.5, E14.5 and E15.5; *Sec16b* at E8.5, E14.5 and E16.5. Of the top candidate genes in the manual burden analysis (see Table 4 and Table S6) only *Ret* was differentially expressed in mice at E8.25, E8.5, E11.5, E14.5, E15.5, E16.5 and E18.5.

Detection of common SNPs associated with IHPS

Determination of the risk allele frequency of four loci highly associated with IHPS (rs11712066, rs573872, rs29784 and rs1933683 near genes *MBNL1*, *NKX2-5* and *BARX1*, respectively) revealed a significantly higher incidence of rs1933683 in our EA/IHPS cohort compared to the population frequency (OR 3.29 (95% CI 1.27-8.56), p=0.009, see Table S7). The risk allele frequency of the other risk loci was not significantly different from the normal population. We did not detect rare putatively deleterious variants in *MBNL1*, *NKX2-5* and *BARX1* in the patient exome sequencing data.

DISCUSSION

We hypothesized that the increased prevalence of IHPS in patients with EA compared to the prevalence of IHPS in the normal population was due to shared CNVs or protein coding alterations in a specific gene, or due to genetic disturbances in genes of shared biological networks during development. As mentioned earlier, both EA and IHPS are variable features in specific genetic syndromes (Table 1). Therefore, to find genetic aberrations that contribute to EA/IHPS we initially searched for pathogenic alterations in known EA or IHPS associated genes (Table S6).

There are no pathogenic changes in known disease genes

As all parents were unaffected, we started this study by focusing on *de novo*, recessive or X-linked changes affecting these known disease genes. However, we could not identify deleterious protein coding alterations, exonic gains or losses or larger CNVs affecting these genes. This is in line with previous studies in which limited causal changes could be detected in patients with EA and associated anomalies (Zhang et al., 2017, Hilger et al., 2015, Brosens et al., 2016b)

Next, we extended our analysis to all genes covered in the exome capture. Given the small sample size (n=15), the low prevalence of the disorder and the high impact on development, we concentrated on genes intolerant to variation (Lek et al., 2016, Ruderfer et al., 2016) harboring rare putative deleterious single nucleotide changes or large CNVs. Moreover, we determined the segregation of alterations in these candidate genes, putative recessive (X-linked, compound heterozygous and homozygous recessive) and all ultra-rare protein altering changes absent from the gnomAD database (Lek et al., 2016) as the later have a high chance of being *de novo* (Bennett et al., 2017). Unfortunately, we did not identify any *de novo* mutations or *de novo* CNVs. None of the identified inherited rare CNVs overlapped in these patients. We could confirm the presence of a compound heterozygous variant in *FAM46A* in one patient and an X-linked variant in *SH3KBP1* in another patient. However, *FAM46A* and *SH3KBP1* are not known to be associated with the gastrointestinal or respiratory tract and were not differentially expressed at the time points important for foregut morphogenesis. These findings made us conclude that neither a dominant nor a recessive model can explain the combination of EA and IHPS in these patients.

The coding sequences of genes crucial in esophageal and pyloric sphincter formation are affected

Subsequently, we focused on genes involved in foregut development. Literature research together with data of previous expression studies provided an overview of genes important for foregut development (Fig. 2).

The development of the foregut is most studied in mouse models. In mice, early foregut formation starts with *Foxa2* stimulation of the anterior endoderm at E8.0 (Heath, 2010). The endodermal sheet folds and forms a tube at E8.75 (Sherwood et al., 2009). Next, signals from the notochord start dorsal-ventral patterning around E9.0, with high *Nkx2.1*/absent *Sox2* in the ventral future trachea and absent *Nkx2.1*/high *Sox2* in the dorsal future esophagus and stomach (Que et al., 2007). These dorsal-ventral patterns lead to compartmentalization of the foregut. Between E9.5 and E11.5 the foregut separates in the primordial esophagus and stomach, and in the primordial trachea. Primordial lung buds become apparent at E9.5 (Sherwood et al., 2009). The separation site is marked by mesenchymal expression of *Barx1* (Woo et al., 2011). The esophagus is completely separated from the trachea at E11.5.

Pyloric sphincter formation is mostly studied in chick and mouse models. This formation starts with the thickening of the circular smooth muscle layer between the antrum and the duodenum around E14.5 and the primordial pyloric sphincter is complete around E18.5 (Smith et al., 2000a, Self et al., 2009). In addition to its functioning in foregut separation, the *Barx1* homeobox gene is also vital for stomach differentiation and stomach smooth muscle development. It inhibits *Wnt* signaling (Woo et al., 2011) and modulates the expression of *Bapx1*, another important factor required for pyloric sphincter morphogenesis (Jayewickreme and Shivdasani, 2015, Stringer et al., 2008, Verzi et al., 2009).

Given the described importance of these genes in normal development, we hypothesized that variations in multiple genes important for foregut morphogenesis might explain the higher incidence of IHPS in patients with EA. We compared a selection of genes – known to be important for foregut morphogenesis or syndromatically associated with EA or IHPS – between the patients and the healthy controls (Table 4, Table S6). Interestingly, in *TNXB* (NM_019105.6:c.4444G>A, p.Val1482Met), *WDR11* (NM_018117.11:c.1138G>T, p.Val380Phe), *PEX3* (NM_003630.2:c.1012A>G, p.Ser338Gly), *TBX3* (NM_016569.3:c.506G>A, p.Arg169Gln), and *GDF6* (NM_001001557.2:c.281C>G, p.Pro94Arg) rare variants were present that were absent in the control group. These variants might not be sufficient to result in disease but are predicted to impact the protein and might contribute together with other unknown factors to disease development.

There is no high frequency burden of rare variants

Given the limited number of samples, we will only detect a gene burden if it is large and has a high impact. We compared the total rare and ultra-rare variant burden of putative deleterious variants in all genes. The number of ultra-rare variants was slightly higher in the patient group compared to the control group but did not differ significantly (Table 5). A second burden analysis identified ten genes

with more variants compared to those seen in the 1000 Genome cohort, two variants were predicted to be deleterious (Table 3). Unfortunately, they did not show any overlap with the results of the expression analysis or candidate genes selected from the literature. Therefore, these variants are not likely to explain the increased incidence of IHPS in EA patients or EA/IHPS development. A rare variant burden might exist but we could not detect it due to limited sample size and/or focus on known candidate genes.

Of all the protein coding changes classified as VUS or higher (Table S4; Table S8), 116 genes were affected with a variant in more than one patient (Fig. 1). Seven of these genes (*ADAMTSL4*, *ANKRD26, CNTN2*, *HSPG2*, *KCNN3*, *LDB3*, *SEC16B*) were differentially expressed in the developing foregut, esophagus or pyloric sphincter in mice between E8.25 and E16.5. However, none of these genes could explain the combination of EA and IHPS within a patient based on their function; none of these genes is known to be associated with the gastrointestinal or respiratory tract. Furthermore, most variants had a population frequency above the prevalence of EA/TEF. If these variants are highly penetrant, they would not be the likely cause. Increasing sample sizes (drastically) would allow an analysis going beyond known intolerant genes, allow us to consider reduced penetrance and potentially identify a shared genetic etiology.

Known common variants associated with IHPS development could have an impact in some patients

Since certain SNPs have been identified with GWAS to be highly associated with IHPS, we wondered if these known common haplotypes could also play a role in the higher incidence of IHPS in patients with EA. In our cohort, we found a significantly higher incidence of the risk loci rs1933683 compared to the population frequency (Table S7). Three patients were homozygous for the risk allele and have a substantially increased risk for IHPS development. The common risk haplotype might therefore impact IHPS development in some of the IHPS patients. However, further research is needed to confirm the impact of this haplotype in a larger EA and EA/IHPS population.

Possible contribution of non-genetic factors

All the data presented so far made us conclude that dominant *de novo* variations in possible disease causing genes do not play a role in our cohort. Recessive inheritance cannot totally be excluded, although our results are not suggestive for this mode of inheritance. We did identify in all patients putative disease-causing variants. Nevertheless, as all parents from whom these variants were inherited were not affected, these variants could contribute but not cause the disease. Previous studies suggested the contribution of non-genetic factors as an explanation for the combined occurrence of EA and IHPS.

Could IHPS be an acquired condition related to surgery or treatment of EA?

The overrepresentation of IHPS in EA patients made us wonder if IHSP could also be the result of the atresia itself, potentially as a result of the surgical procedure to correct the atresia or the result of treatment. Previous studies also mentioned vagal nerve lesions, a gastrostomy and transpyloric feeding tubes as possible causes for an increased incidence of IHPS after correction of EA (Ilhan et al., 2018). IHPS has been suggested to be a neuromuscular disorder with the involvement of smooth muscle cells, interstitial cells of Cajal and the enteric nervous system. The hypertrophy is suggested to be the result of discoordinated movements of the pyloric sphincter and the contractions of the stomach (Hayes and Goldenberg, 1957), perhaps as the result of absent nitric oxide synthase activity (Vanderwinden et al., 1992). Impaired gastric contractility and esophageal relaxation were observed in Adriamycin and doxorubicin induced EA in mice (Tugay et al., 2003, Tugay et al., 2001). Mechanistically, this association between EA and IHPS seems plausible. However, it does not explain why IHPS is not fully penetrant in patients with EA. The most common thought is that mechanical and environmental factors disturb the developmental field. To which extent these factors influence the development of the child, depends on the specific risk factors and their timing. Further research on the cause and other specific clinical risk factors for patients with EA should be considered, e.g. the late start of oral feeding or the long-term feeding through a tube instead of drinking themselves.

Models for EA/IHPS disease etiology

Since we hypothesized that genetic defects, disturbing foregut morphogenesis, would be responsible for the combination of EA/IHPS, we started with the thought of a (monogenetic) syndromic model. However, we have not been able to find a central gene impacted in most patients which can explain the increased prevalence of IHPS in patients with EA. An as yet unknown syndrome is unlikely since we have not found *de novo* (as all parents are unaffected) or shared high impact variants in the same gene multiple patients. However, we cannot exclude *de novo* mutations which have been seen in the GnomAD exomes or genomes, nor did we look beyond the coding part of our genome.

Furthermore, we have detected inherited rare variants in candidate genes and genes affected more than once by variants with a low (unaffected) population frequency. Therefore, we cannot exclude a genetic component. Another option is that, although the combination of EA and IHPS could not be explained by one gene or locus, the underlying cause of EA and the other associated anomalies is the result of hits in multiple genes. About 10% of the patients with EA have an underlying genetic syndrome (Brosens et al., 2014) and more can be expected. The phenotypical spectrum of this cohort is very heterogeneous and could be the result of impacts on multiple genes. IHPS could then independently be caused by mechanical factors such as the surgical procedure.

Furthermore, environmental risk factors have been suggested for EA and IHPS, like pesticides, smoking, herbicides and periconceptional alcohol or multivitamin use (Zwink et al., 2016, Felix et al., 2008, Feng et al., 2016, Markel et al., 2015, Krogh et al., 2012, Sorensen et al., 2002). Considering the absence of highly penetrant recurring genetic variations, we now hypothesize different multifactorial models for disease development. In all these models the combination of CNVs, deleterious protein alterations (Felix et al., 2007b, Brosens et al., 2014), severe changes in the developmental field during the organogenesis (Martinez-Frias, 1994, Martinez-Frias and Frias, 1997) and/or environmental inducing epigenetic changes (Sorensen et al., 2002) together can modulate the phenotypical spectrum seen in these patients. Examples of possible mechanical and environmental factors disturbing the developmental field are mice models with Adriamycin induction or dorsal-ventral patterning signals from the notochord.

Our first hypothesis is based on the earlier published theory of Brosens et al. about disturbed ENS development (Brosens et al., 2016a). It includes a seesaw model in which risk factors are in balance with protective mechanisms. In this model, the fulcrum can be shifted by a genetic variation in a central gene, which automatically disrupts the balance. When applying this theory on EA/IHPS patients, this would lead to more or less affected organ systems within the VACTERL syndrome. However, in this study we have not been able to detect one central gene. This makes a seesaw model with a variable fulcrum less likely as an explanation for the increased prevalence of IHPS in the EA population.

A second hypothesis is a burden model (Fig. 3A). Similar, (epi)genetic, environmental and mechanical factors form a burden of risk factors, which balances with protective mechanisms. In this model, the point of balance is not shifted by a mutation in a central gene and every person has contributions of certain risk factors. But in most cases this does not lead to affected organ systems. There is an intermediate range between normal and affected in which individuals can have the genetic burden but lacks an abnormal phenotype (reduced penetrance) or their symptoms differ in severity (variable expressivity). The latter would fit the results in this study; maybe we did have detected variants but have we failed to interpret them correctly as parents were seemingly unaffected and/or the variant frequency can be higher in unaffected controls to be of relevance in patients. Mechanical or environmental factors could have made the difference in shifting the balance. All together the burden model is a plausible explanation for the disease development.

Last, we hypothesize a slippery slope model (Fig. 3B). In this model, the burden of low impact genetic variants and environmental disturbances alone does not impact the balance seen in the seesaw

model unless it crosses a certain threshold. Moreover, we hypothesize that the protective mechanisms (e.g. compensatory mechanisms) during development are very strong, making it really difficult to shift the balance. Most fetuses do not develop any malformations despite the combined genetic and environmental burden or do not survive. But once the threshold is reached, the balance is immediately greatly disrupted and often multiple organ systems are affected. This model fits with the phenotypical results in this study since four patients (14.8%) had three or more anomalies within the VACTERL spectrum. In this model there is a high tolerance for low impact genetic variation and only high impact variation (aneuploidies, exposure to toxic substances, pathogenic changes in developmental crucial genes) shifts the balance. When the balance is disturbed, it shifts drastically. We did not detect high impact changes responsible for the EA/IHPS combination. As parents are unaffected it is (in this model) unlikely that inherited variants impact disease development, nor would variants which are seen in the (unaffected) population controls.

Limitations

Not finding any positive correlation between DNA variations in specific genes or developmental pathways is partly due to the small data set we have (15 patients). The small sample size is no problem for the *de novo* and our recessive model strategy in known disease genes, but it is so for the heterozygous variant burden analysis. Another limitation is the lack of data on the expression of genes involved in normal foregut development in human embryos. Our gene selection was based on mouse transcriptome data. Little human data is available since human embryos of 4 to 6 weeks old are generally not preserved. However, although it is unclear how precisely the foregut development in mice corresponds with humans, it is unlikely that this is very different in its early phases. Finally, one could argue that variations in the non-coding part of the genome are major contributors. Although we did investigate known IHPS risk loci and determined genome wide CNV profiles, we did not determine genome wide variation.

Conclusions

To conclude, *de novo* mutations (a dominant model) and homozygous or compound heterozygous mutations (a recessive model) in the protein coding part of the genome are not a likely cause for the combination of EA and IHPS. Although the presence of genetic variation in likely candidate genes suggests a genetic component, there does not seem to be an enrichment of genetic variants in good candidate genes in patients. There are putative deleterious variants in foregut or disease genes which might contribute to disease development and although there is no difference in burden, some variants might contribute more than others and this is not taken into account in a burden test. We

might have misinterpreted the impact of some of the inherited variants. Furthermore, in some patients the IHPS predisposing locus rs1933683 is present.

We hypothesized several multifactorial models in which the combination of multiple high impact genetic, mechanical and environmental factors together can shift the balance from normal to abnormal development. A burden model with reduced penetrance or variable expressivity is most likely if genetic factors contribute. Future research should investigate the incidence of IHPS in bigger EA patients cohorts to further explore this theory. To exclude the role of treatment or surgery, clinical factors related to the surgical correction of EA – for example vagal nerve lesions after surgery, the late start of oral feeding or transpyloric feeding tubes – should be systematically registered.

MATERIAL AND METHODS

Patient cohort

This study was approved by the Medical Ethical Review Board of Erasmus MC - Sophia Children's Hospital (MEC 193.948/2000/159). We searched the Erasmus University MC-Sophia EA-cohort and the database of the standardized prospective longitudinal follow up program in our hospital for children with congenital anatomical anomalies (Gischler et al., 2009) for patients born between 1970-2017 with a combination of both EA and IHPS in history. Patients were included and analyzed after parental informed consent

SNP-array analysis

Micro-array analysis was performed using the single-nucleotide polymorphism (SNP) CytoSNP-850Kv0 BeadChip (Illumina Inc., San Diego) using standard protocols and the GenomeStudio genotyping module (v1.9.4, www.illumnia.com). Visualization of Copy Number Variations (CNVs), Runs of Homozygosity (ROH) and comparisons to in-house control cohorts as well as published cohorts of affected and control individuals was done using Biodiscovery Nexus CN7.5. (Biodiscovery Inc., Hawthorne, CA, USA) and described previously (Brosens et al., 2016b).

Variant pre-filtering and prioritization

The initial variant filtering method has been described previously (Halim et al., 2017). In brief, we included all variants with an allele frequency below 1% in 1000 Genomes phase 3 version 5, Exome Variant Server 6500 v0.0.30, Genome of the Netherlands (Genome of the Netherlands, 2014), ExAC 0.3 and our in-house cohort (n=906), consisting of individuals captured with the SureSelect Human All Exon 50 Mb Targeted exome enrichment kit v4 (n=279), SureSelect Clinical Research Exome v1

(n=387) and Haloplex Exome target enrichment system (n=240), Agilent Technologies, Inc., Santa Clara, California).

All nonsense variants, variants predicted to affect splicing and all variants with a Combined Annotation-Dependent Depletion (CADD) score (Kircher et al., 2014) above 20 were selected for individual patient analysis in downstream tools. Different downstream tools were used to prioritize the variants. Prioritized variants were further classified according to the criteria in Table S8. Determination of variant segregation and confirmation of *de novo* of inherited status of variants was done with Sanger sequencing unless otherwise indicated.

Variant burden test and prioritization using Opal

We used the Variant Annotation, Analysis & Search Tool (VAAST) (Yandell et al., 2011, Hu et al., 2013, Kennedy et al., 2014) cohort analysis embedded in Opal 4.29.5 (Fabric Genomics, Oakland, CA, USA) to rank the variants in the individual patients. Secondly, we performed a burden test on the full exomes using Exome Variant Server 6500 v0.0.30 and 1000 Genomes phase 3 version 5 as a control cohort. We used a 1% allele frequency cut-off for recessive (hemizygous and homozygous) variants and 0.1% cut-off for heterozygous variants. Compound heterozygosity was not considered in this analysis as we did not know the phase of the haplotypes. Only putative protein changing (nonsense, missense, initiator codon variants, in-frame indels, splice sites and splice regions) variants were taken into account. Since we were only interested in putative deleterious variants we used an Omicia score of 0.79 as a threshold as this cut-off has a false positive rate of 5%. Omicia is an algorithm included in the Opal software that combines SIFT (Ng and Henikoff, 2001), PolyPhen (Adzhubei et al., 2010), MutationTaster (Schwarz et al., 2010) and PhyloP (Siepel et al., 2005) to predict deleteriousness of variants.

For the VAAST burden test we used a minimum significance of 0.05 and a gene had to have at least two distinct variants in the case set. These genes were used as a gene panel in the individual patient analysis. Individual variants were prioritized before individual inspection as follows. First, all recessive (X-linked and putative homozygous and compound heterozygous), putative rare (MAF ≤0.001%) and damaging *de novo* variants were selected. Secondly, the top 10 of variants ranked by the VAAST 1.1 prioritization algorithm and subsequently the top 10 variants re-ranked by the Phevor algorithm (Singleton et al., 2014) were included. We used the Human Phenotype Ontology (HPO) (Singleton et al., 2014) terms esophageal atresia and pyloric stenosis as phenotype terms in the algorithm. Finally, variants passing the pre-filtering criteria in genes from the burden test were included.

Variant prioritization using bioinformatic genotype-phenotype correlation tools

Three modules were used: PhenIX (Zemojtel et al., 2014) (http://compbio.charite.de/PhenIX/), the Exomiser (Robinson et al., 2014) (http://www.sanger.ac.uk/resources/software/exomiser/submit/) and the HPO prioritization incorporated within the Cartagenia software. Settings were as followed.

Using PhenIX the full patient phenotype in HPO terms was used, the exome target region filter is on and allele frequency filter of 0.1%, pathogenicity filter was on and mode of inheritance unknown. Genes were prioritized using PhenIX which compares patient phenotypes against human phenotypes only. As a cut-of we used a gene relevance score of 0.8 in combination with a variant score of 0.8, or a total score of 0.9.

When using the Exomiser tool we used similar settings: full patient phenotype in HPO terms, exome target region filter is off, allele frequency filter 0.1%, pathogenicity filter on. We did not remove dbSNP variants nor used an inheritance model. Genes are now prioritized using hiPhive, which compares phenotypes against all species. As a cut-of we used a phenotype score of 0.8 in combination with a variant score of 0.8, or an Exomiser score of 0.9.

Pathway enrichment analysis of genes affected by rare variants

To investigate if specific pathways are enriched with ultra-rare variants, Gene IDs with variants in canonical splice sites (n=16), nonsense variants (n=21), protein altering inframe InDels (n=28) and missense variants (n=557) were uploaded to Ingenuity pathway Analysis (Qiagen, Venlo, The Netherlands). Additionally, a more stringent set was uploaded with loss of function variants, predicted to be loss of function intolerant (PLI ≥0.9, n=4) and protein altering variants with a Z-score ≥3 (n=44).

Expression of candidate genes

Candidate gene expression was determined at relevant developmental time points in human and mouse. Gene expression of top-ranking genes derived from the burden analysis and individual patient sample prioritizations were determined using datasets (GSE13040, GSE19873, GSE34278, GSE15872, GSE43381) downloaded from the Gene Expression Omnibus (GEO) (Edgar et al., 2002). We used public data on mice on the endoderm, mesoderm and ectoderm at E8.25, foregut at E8.5 and esophagus, stomach, pyloric sphincter and intestine at E11.5-E18.5 (https://www.ncbi.nlm.nih.gov/geo/) (Stephens et al., 2013, Li et al., 2009, Sherwood et al., 2009, Millien et al., 2008, Chen et al., 2012). These datasets were imported into BRB-ArrayTools Version: 4.5.0 - Beta_2. (http://linus.nci.nih.gov/BRB-ArrayTools.html), annotated by Bioconductor (www.bioconductor.org), R version 3.2.2 Patched (2015-09-12 r69372) and normalized. We

535 determined differential expression between tissue types and classified upregulated genes being
536 expressed in the tissue under investigation.
537
538 *Detection of common SNP associated with IHPS*
539 Genome-wide association studies (GWAS) revealed five loci highly associated with IHPS (rs11712066,
540 rs573872, rs29784, rs1933683 and rs6736913), pointing towards *MBNL1*, *NKX2-5*, *BARX1* and *EML4*
541 as candidate genes (Feenstra et al., 2012, Everett and Chung, 2013, Fadista et al., 2019). Since
542 rs6736913 is a low frequency missense variant, we did not further analyze this SNP in our patients.
543 We used Sanger sequencing to determine the risk allele frequency of the other four SNPs. With a chi
544 square test we compared the allele frequency in our patients with the gnomAD dataset
545 (http://gnomad.broadinstitute.org/).
546

ACKNOWLEDGEMENTS

We are grateful for the help of families, patients and the cooperation of the patient cooperation "Vereniging voor Ouderen en Kinderen met een Slokdarmafsluiting". We would like to thank Tom de Vries Lentsch for preparing the figures.

COMPETING INTERESTS

Authors do not have any potential conflicts (financial, professional, or personal) relevant to the manuscript to disclose.

FUNDING

Erwin Brosens was funded by the Sophia Foundations for Scientific Research, projects SSWO-493 and SWOO13-09.

DATA AVAILABILITY

Variants included in Table S4 are submitted to the ClinVar database. (http://www.ncbi.nlm.nih.gov/clinvar/)

AUTHOR CONTRIBUTIONS STATEMENT

Conceptualization: D.T., R.W., R.H., A.K., E.B.; Methodology: R.B., H.E., W.IJ., A.K., E.B.; Software: R.B., T.B.; Validation: C.K., V.M., D.H.; Formal analysis: C.K.; Investigation: C.K., Resources: Y.B., N.B., A.B., H.E., H.IJ.; Data curation: E.B.; Writing – original draft: C.K., R.H., E.B.; Writing – review & editing: C.K., R.B., Y.B., A.B., H.E., W.IJ., H.IJ. D.T., R.W., R.H., A.K., E.B.; Visualization: C.K., Supervision: E.B.; Project administration: D.T., R.W., R.H., A.K., E.B.; Funding acquisation: D.T., R.W., R.H., A.K.

REFERENCES

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S. & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nat Methods,* **7,** 248-9.

Amarasinghe, K. C., Li, J. & Halgamuge, S. K. (2013). CoNVEX: copy number variation estimation in exome sequencing data using HMM. *BMC Bioinformatics,* **14 Suppl 2,** S2.

Anderson, R. B., Newgreen, D. F. & Young, H. M. (2006). Neural crest and the development of the enteric nervous system. *Adv Exp Med Biol,* **589,** 181-96.

Ayme, S., Julian, C., Gambarelli, D., Mariotti, B., Luciani, A., Sudan, N., Maurin, N., Philip, N., Serville, F., Carles, D., et al. (1989). Fryns syndrome: report on 8 new cases. *Clin Genet,* **35,** 191-201.

Barberia Leache, E., Saavedra Ontiveros, D. & Maroto Edo, M. (2003). Etiopathogenic analysis of the caries on three patients with Noonan Syndrome. *Med Oral,* **8,** 136-42.

Battaglia, A., Chines, C. & Carey, J. C. (2006). The FG syndrome: report of a large Italian series. *Am J Med Genet A,* **140,** 2075-9.

Bennett, C. A., Petrovski, S., Oliver, K. L. & Berkovic, S. F. (2017). ExACtly zero or once: A clinically helpful guide to assessing genetic variants in mild epilepsies. *Neurol Genet,* **3,** e163.

Blank, C. E. (1960). Apert's syndrome (a type of acrocephalosyndactyly)-observations on a British series of thirty-nine cases. *Ann Hum Genet,* **24,** 151-64.

Brosens, E., Burns, A. J., Brooks, A. S., Matera, I., Borrego, S., Ceccherini, I., Tam, P. K., Garcia-Barcelo, M. M., Thapar, N., Benninga, M. A., et al. (2016a). Genetics of enteric neuropathies. *Dev Biol,* **417,** 198-208.

Brosens, E., Marsch, F., De Jong, E. M., Zaveri, H. P., Hilger, A. C., Choinitzki, V. G., Holscher, A., Hoffmann, P., Herms, S., Boemers, T. M., et al. (2016b). Copy number variations in 375 patients with oesophageal atresia and/or tracheoesophageal fistula. *Eur J Hum Genet,* **24,** 1715-1723.

Brosens, E., Ploeg, M., Van Bever, Y., Koopmans, A. E., H, I. J., Rottier, R. J., Wijnen, R., Tibboel, D. & De Klein, A. (2014). Clinical and etiological heterogeneity in patients with tracheo-esophageal malformations and associated anomalies. *Eur J Med Genet*.

Cates, M., Billmire, D. F., Bull, M. J. & Grosfeld, J. L. (1989). Gastroesophageal dysfunction in Cornelia de Lange syndrome. *J Pediatr Surg,* **24,** 248-50.

Cetin, E., Malas, M. A., Albay, S. & Cankara, N. (2006). The development of stomach during the fetal period. *Surg Radiol Anat,* **28,** 438-46.

Chen, H., Li, J., Li, H., Hu, Y., Tevebaugh, W., Yamamoto, M., Que, J. & Chen, X. (2012). Transcript profiling identifies dynamic gene expression patterns and an important role for Nrf2/Keap1 pathway in the developing mouse esophagus. *PLoS One,* **7,** e36504.

Christiano, A. M., Mcgrath, J. A., Tan, K. C. & Uitto, J. (1996). Glycine substitutions in the triple-helical region of type VII collagen result in a spectrum of dystrophic epidermolysis bullosa phenotypes and patterns of inheritance. *Am J Hum Genet,* **58,** 671-81.

Christiano, A. M., Suga, Y., Greenspan, D. S., Ogawa, H. & Uitto, J. (1995). Premature termination codons on both alleles of the type VII collagen gene (COL7A1) in three brothers with recessive dystrophic epidermolysis bullosa. *J Clin Invest,* **95,** 1328-34.

De Jong, E. M., Felix, J. F., De Klein, A. & Tibboel, D. (2010). Etiology of esophageal atresia and tracheoesophageal fistula: "mind the gap". *Curr Gastroenterol Rep,* **12,** 215-22.

Deurloo, J. A., Ekkelkamp, S., Schoorl, M., Heij, H. A. & Aronson, D. C. (2002). Esophageal atresia: historical evolution of management and results in 371 patients. *Annals of Thoracic Surgery,* **73,** 267-72.

Edgar, R., Domrachev, M. & Lash, A. E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res,* **30,** 207-10.

Elinoff, J. M., Liu, D., Guandalini, S. & Waggoner, D. J. (2005). Familial pyloric stenosis associated with developmental delays. *J Pediatr Gastroenterol Nutr,* **41,** 129-32.

Everett, K. V. & Chung, E. M. (2013). Confirmation of two novel loci for infantile hypertrophic pyloric stenosis on chromosomes 3 and 5. *J Hum Genet,* **58,** 236-7.

Fadista, J., Skotte, L., Geller, F., Bybjerg-Grauholm, J., Gortz, S., Romitti, P. A., Caggana, M., Kay, D. M., Matsson, H., Boyd, H. A., et al. (2019). Genome-wide meta-analysis identifies BARX1 and EML4-MTA3 as new loci associated with infantile hypertrophic pyloric stenosis. *Hum Mol Genet,* **28,** 332-340.

Fausett, S. R. & Klingensmith, J. (2012). Compartmentalization of the foregut tube: developmental origins of the trachea and esophagus. *Wiley Interdiscip Rev Dev Biol,* **1,** 184-202.

Feenstra, B., Geller, F., Carstensen, L., Romitti, P. A., Korberg, I. B., Bedell, B., Krogh, C., Fan, R., Svenningsson, A., Caggana, M., et al. (2013). Plasma lipids, genetic variants near APOA1, and the risk of infantile hypertrophic pyloric stenosis. *Jama,* **310,** 714-21.

Feenstra, B., Geller, F., Krogh, C., Hollegaard, M. V., Gortz, S., Boyd, H. A., Murray, J. C., Hougaard, D. M. & Melbye, M. (2012). Common variants near MBNL1 and NKX2-5 are associated with infantile hypertrophic pyloric stenosis. *Nat Genet,* **44,** 334-7.

Felix, J. F., Steegers-Theunissen, R. P., De Walle, H. E., De Klein, A., Torfs, C. P. & Tibboel, D. (2007a). Esophageal atresia and tracheoesophageal fistula in children of women exposed to diethylstilbestrol in utero. *Am J Obstet Gynecol,* **197,** 38.e1-5.

639 **Felix, J. F., Tibboel, D. & De Klein, A.** (2007b). Chromosomal anomalies in the aetiology of
640 oesophageal atresia and tracheo-oesophageal fistula. *Eur J Med Genet,* **50,** 163-75.
641 **Felix, J. F., Van Dooren, M. F., Klaassens, M., Hop, W. C., Torfs, C. P. & Tibboel, D.** (2008).
642 Environmental factors in the etiology of esophageal atresia and congenital diaphragmatic
643 hernia: results of a case-control study. *Birth Defects Res A Clin Mol Teratol,* **82,** 98-105.
644 **Feng, Y., Chen, R., Li, X. & Mo, X.** (2016). Environmental factors in the etiology of isolated and
645 nonisolated esophageal atresia in a Chinese population: A case-control study. *Birth Defects*
646 *Res A Clin Mol Teratol,* **106,** 840-846.
647 **Franceschini, P., Guala, A., Licata, D., Franceschini, D., Signorile, F. & Di Cara, G.** (1997). Esophageal
648 atresia with distal tracheoesophageal fistula in a patient with fronto-metaphyseal dysplasia.
649 *Am J Med Genet,* **73,** 10-4.
650 **Freeman, S. B., Torfs, C. P., Romitti, P. A., Royle, M. H., Druschel, C., Hobbs, C. A. & Sherman, S. L.**
651 (2009). Congenital gastrointestinal defects in Down syndrome: a report from the Atlanta and
652 National Down Syndrome Projects. *Clin Genet,* **75,** 180-4.
653 **Gargiulo, A., Auricchio, R., Barone, M. V., Cotugno, G., Reardon, W., Milla, P. J., Ballabio, A.,**
654 **Ciccodicola, A. & Auricchio, A.** (2007). Filamin A is mutated in X-linked chronic idiopathic
655 intestinal pseudo-obstruction with central nervous system involvement. *Am J Hum Genet,* **80,**
656 751-8.
657 **Genome of the Netherlands, C.** (2014). Whole-genome sequence variation, population structure and
658 demographic history of the Dutch population. *Nat Genet,* **46,** 818-25.
659 **Gillis, L. A., Mccallum, J., Kaur, M., Descipio, C., Yaeger, D., Mariani, A., Kline, A. D., Li, H. H.,**
660 **Devoto, M., Jackson, L. G., et al.** (2004). NIPBL mutational analysis in 120 individuals with
661 Cornelia de Lange syndrome and evaluation of genotype-phenotype correlations. *Am J Hum*
662 *Genet,* **75,** 610-23.
663 **Gischler, S. J., Mazer, P., Duivenvoorden, H. J., Van Dijk, M., Bax, N. M., Hazebroek, F. W. &**
664 **Tibboel, D.** (2009). Interdisciplinary structural follow-up of surgical newborns: a prospective
665 evaluation. *J Pediatr Surg,* **44,** 1382-9.
666 **Gripp, K. W. & Lin, A. E. (**1993). Costello Syndrome. *In:* PAGON, R. A., ADAM, M. P., ARDINGER, H. H.,
667 BIRD, T. D., DOLAN, C. R., FONG, C. T., SMITH, R. J. H. & STEPHENS, K. (eds.) *GeneReviews(R).*
668 Seattle (WA): University of Washington, Seattle
669 University of Washington, Seattle. All rights reserved.
670 **Gross, R. E.** (1947). Atresia of the esophagus. *Am J Dis Child,* **74,** 369.
671 **Halim, D., Brosens, E., Muller, F., Wangler, M. F., Beaudet, A. L., Lupski, J. R., Akdemir, Z. H. C.,**
672 **Doukas, M., Stoop, H. J., De Graaf, B. M., et al.** (2017). Loss-of-Function Variants in MYLK

673 Cause Recessive Megacystis Microcolon Intestinal Hypoperistalsis Syndrome. *Am J Hum*
674 *Genet,* **101,** 123-129.
675 **Hayes, M. A. & Goldenberg, I. S.** (1957). The problems of infantile pyloric stenosis. *Surg Gynecol*
676 *Obstet,* **104,** 105-38.
677 **Heath, J. K. (**2010). Chapter Four - Transcriptional Networks and Signaling Pathways that Govern
678 Vertebrate Intestinal Development. *In:* PETER, K. (ed.) *Current Topics in Developmental*
679 *Biology*. Academic Press.
680 **Hilger, A. C., Halbritter, J., Pennimpede, T., Van Der Ven, A., Sarma, G., Braun, D. A., Porath, J. D.,**
681 **Kohl, S., Hwang, D. Y., Dworschak, G. C., et al.** (2015). Targeted Resequencing of 29
682 Candidate Genes and Mouse Expression Studies Implicate ZIC3 and FOXF1 in Human
683 VATER/VACTERL Association. *Hum Mutat,* **36,** 1150-4.
684 **Hovnanian, A., Hilal, L., Blanchet-Bardon, C., De Prost, Y., Christiano, A. M., Uitto, J. & Goossens,**
685 **M.** (1994). Recurrent nonsense mutations within the type VII collagen gene in patients with
686 severe recessive dystrophic epidermolysis bullosa. *Am J Hum Genet,* **55,** 289-96.
687 **Hu, H., Huff, C. D., Moore, B., Flygare, S., Reese, M. G. & Yandell, M.** (2013). VAAST 2.0: improved
688 variant classification and disease-gene identification using a conservation-controlled amino
689 acid substitution matrix. *Genet Epidemiol,* **37,** 622-34.
690 **Ilhan, O., Bor, M., Gunendi, T. & Dorterler, M. E.** (2018). Hypertrophic pyloric stenosis following
691 repair of oesophageal atresia and tracheo-oesophageal fistula in a neonate. *BMJ Case Rep,*
692 **2018**.
693 **Jacobs, I. J. & Que, J.** (2013). Genetic and cellular mechanisms of the formation of esophageal atresia
694 and tracheoesophageal fistula. *Dis Esophagus,* **26,** 356-8.
695 **Jayewickreme, C. D. & Shivdasani, R. A.** (2015). Control of stomach smooth muscle development and
696 intestinal rotation by transcription factor BARX1. *Dev Biol,* **405,** 21-32.
697 **Jongmans, M. C., Van Ravenswaaij-Arts, C. M., Pitteloud, N., Ogata, T., Sato, N., Claahsen-Van Der**
698 **Grinten, H. L., Van Der Donk, K., Seminara, S., Bergman, J. E., Brunner, H. G., et al.** (2009).
699 CHD7 mutations in patients initially diagnosed with Kallmann syndrome--the clinical overlap
700 with CHARGE syndrome. *Clin Genet,* **75,** 65-71.
701 **Kennedy, B., Kronenberg, Z., Hu, H., Moore, B., Flygare, S., Reese, M. G., Jorde, L. B., Yandell, M. &**
702 **Huff, C.** (2014). Using VAAST to Identify Disease-Associated Variants in Next-Generation
703 Sequencing Data. *Curr Protoc Hum Genet,* **81,** 6 14 1-25.
704 **Kim, H. G., Kurth, I., Lan, F., Meliciani, I., Wenzel, W., Eom, S. H., Kang, G. B., Rosenberger, G.,**
705 **Tekin, M., Ozata, M., et al.** (2008). Mutations in CHD7, encoding a chromatin-remodeling
706 protein, cause idiopathic hypogonadotropic hypogonadism and Kallmann syndrome. *Am J*
707 *Hum Genet,* **83,** 511-9.

Kircher, M., Witten, D. M., Jain, P., O'roak, B. J., Cooper, G. M. & Shendure, J. (2014). A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet,* **46,** 310-5.

Koyuncu, E., Malas, M. A., Albay, S., Cankara, N. & Karahan, N. (2009). The development of fetal pylorus during the fetal period. *Surg Radiol Anat,* **31,** 335-41.

Kroes, H. Y., Pals, G. & Van Essen, A. J. (2003). Ehlers-Danlos syndrome type IV: unusual congenital anomalies in a mother and son with a COL3A1 mutation and a normal collagen III protein profile. *Clin Genet,* **63,** 224-7.

Krogh, C., Fischer, T. K., Skotte, L., Biggar, R. J., Oyen, N., Skytthe, A., Goertz, S., Christensen, K., Wohlfahrt, J. & Melbye, M. (2010). Familial aggregation and heritability of pyloric stenosis. *JAMA,* **303,** 2393-9.

Krogh, C., Gortz, S., Wohlfahrt, J., Biggar, R. J., Melbye, M. & Fischer, T. K. (2012). Pre- and perinatal risk factors for pyloric stenosis and their influence on the male predominance. *Am J Epidemiol,* **176,** 24-31.

Kuivaniemi, H., Kontusaari, S., Tromp, G., Zhao, M. J., Sabol, C. & Prockop, D. J. (1990). Identical G+1 to A mutations in three different introns of the type III procollagen gene (COL3A1) produce different patterns of RNA splicing in three variants of Ehlers-Danlos syndrome. IV. An explanation for exon skipping some mutations and not others. *J Biol Chem,* **265,** 12067-74.

Landrum, M. J., Lee, J. M., Riley, G. R., Jang, W., Rubinstein, W. S., Church, D. M. & Maglott, D. R. (2014). ClinVar: public archive of relationships among sequence variation and human phenotype. *Nucleic Acids Res,* **42,** D980-5.

Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., O'donnell-Luria, A. H., Ware, J. S., Hill, A. J., Cummings, B. B., et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature,* **536,** 285-91.

Li, X., Udager, A. M., Hu, C., Qiao, X. T., Richards, N. & Gumucio, D. L. (2009). Dynamic patterning at the pylorus: formation of an epithelial intestine-stomach boundary in late fetal life. *Dev Dyn,* **238,** 3205-17.

Lodha, A. K., Satodia, P. & Whyte, H. (2005). Fetal alcohol syndrome and pyloric stenosis: alcohol induced or an association? *J Perinat Med,* **33,** 262-3.

Macchini, F., Parente, G., Morandi, A., Farris, G., Gentilino, V. & Leva, E. (2017). Classification of Esophageal Strictures following Esophageal Atresia Repair. *Eur J Pediatr Surg*.

Macmahon, B. (2006). The continuing enigma of pyloric stenosis of infancy: a review. *Epidemiology,* **17,** 195-201.

Mangyanda, M. K., Mbuila, C., Geniez, L., Personne, A., Boize, P., Gasmi, E. H., Saf, E. & Hayat, P. (1998). [Fetal alcohol syndrome and hypertrophic pyloric stenosis in two brothers]. *Arch Pediatr,* **5,** 695-6.

Markel, T. A., Proctor, C., Ying, J. & Winchester, P. D. (2015). Environmental pesticides increase the risk of developing hypertrophic pyloric stenosis. *J Pediatr Surg,* **50,** 1283-8.

Martinez-Frias, M. L. (1994). Developmental field defects and associations: epidemiological evidence of their relationship. *Am J Med Genet,* **49,** 45-51.

Martinez-Frias, M. L. & Frias, J. L. (1997). Primary developmental field. III: Clinical and epidemiological study of blastogenetic anomalies and their relationship to different MCA patterns. *Am J Med Genet,* **70,** 11-5.

Mcmullen, K. P., Karnes, P. S., Moir, C. R. & Michels, V. V. (1996). Familial recurrence of tracheoesophageal fistula and associated malformations. *Am J Med Genet,* **63,** 525-8.

Millien, G., Beane, J., Lenburg, M., Tsao, P. N., Lu, J., Spira, A. & Ramirez, M. I. (2008). Characterization of the mid-foregut transcriptome identifies genes regulated during lung bud induction. *Gene Expr Patterns,* **8,** 124-39.

Mori, M., Yamagata, T., Mori, Y., Nokubi, M., Saito, K., Fukushima, Y. & Momoi, M. Y. (1996). Elastic fiber degeneration in Costello syndrome. *Am J Med Genet,* **61,** 304-9.

Nezelof, C., Jaubert, F. & Lyon, G. (1976). [Familial syndrome combining short small intestine, intestinal malrotation, pyloric hypertrophy and brain malformation. 3 anatomoclinical case reports]. *Ann Anat Pathol (Paris),* **21,** 401-12.

Ng, P. C. & Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Res,* **11,** 863-74.

Oddsberg, J., Lu, Y. & Lagergren, J. (2012). Aspects of esophageal atresia in a population-based setting: incidence, mortality, and cancer risk. *Pediatr Surg Int,* **28,** 249-57.

Palacios M.E.C., Sanz, J. C., Tàrranga A.B.D., San Roman C.G. & J.J.V., C. (2014). Esophageal atresia and hypertrofic pyloric stenosis: An association to consider. *Case Rep Intern Med,* **1,** 187-90.

Panteli, C. (2009). New insights into the pathogenesis of infantile pyloric stenosis. *Pediatr Surg Int,* **25,** 1043-52.

Pedersen, R. N., Calzolari, E., Husby, S., Garne, E. & Group, E. W. (2012). Oesophageal atresia: prevalence, prenatal diagnosis and associated anomalies in 23 European regions. *Arch Dis Child,* **97,** 227-32.

Peeters, B., Benninga, M. A. & Hennekam, R. C. (2012). Infantile hypertrophic pyloric stenosis--genetics and syndromes. *Nat Rev Gastroenterol Hepatol,* **9,** 646-60.

Pelz, L., Unger, K. & Radke, M. (1994). Esophageal stenosis in acrocephalosyndactyly type I. *Am J Med Genet,* **53,** 91.

Perin, S., Mccann, C. J., Borrelli, O., De Coppi, P. & Thapar, N. (2017). Update on Foregut Molecular Embryology and Role of Regenerative Medicine Therapies. *Front Pediatr,* **5,** 91.

Que, J., Okubo, T., Goldenring, J. R., Nam, K. T., Kurotani, R., Morrisey, E. E., Taranova, O., Pevny, L. H. & Hogan, B. L. (2007). Multiple dose-dependent roles for Sox2 in the patterning and differentiation of anterior foregut endoderm. *Development,* **134,** 2521-31.

Rajab, A., Heathcote, K., Joshi, S., Jeffery, S. & Patton, M. (2002). Heterogeneity for congenital generalized lipodystrophy in seventeen patients from Oman. *Am J Med Genet,* **110,** 219-25.

Rajab, A., Straub, V., Mccann, L. J., Seelow, D., Varon, R., Barresi, R., Schulze, A., Lucke, B., Lutzkendorf, S., Karbasiyan, M., et al. (2010). Fatal cardiac arrhythmia and long-QT syndrome in a new form of congenital generalized lipodystrophy with muscle rippling (CGL4) due to PTRF-CAVIN mutations. *PLoS Genet,* **6,** e1000874.

Robert, E., Mutchinick, O., Mastroiacovo, P., Knudsen, L. B., Daltveit, A. K., Castilla, E. E., Lancaster, P., Kallen, B. & Cocchi, G. (1993). An international collaborative study of the epidemiology of esophageal atresia or stenosis. *Reprod Toxicol,* **7,** 405-21.

Robinson, P. N., Kohler, S., Oellrich, A., Sanger Mouse Genetics, P., Wang, K., Mungall, C. J., Lewis, S. E., Washington, N., Bauer, S., Seelow, D., et al. (2014). Improved exome prioritization of disease genes through cross-species phenotype comparison. *Genome Res,* **24,** 340-8.

Rollins, M. D., Shields, M. D., Quinn, R. J. & Wooldridge, M. A. (1989). Pyloric stenosis: congenital or acquired? *Archives of Disease in Childhood,* **64,** 138-139.

Ruderfer, D. M., Hamamsy, T., Lek, M., Karczewski, K. J., Kavanagh, D., Samocha, K. E., Daly, M. J., Macarthur, D. G., Fromer, M. & Purcell, S. M. (2016). Patterns of genic intolerance of rare copy number variation in 59,898 human exomes. *Nat Genet,* **48,** 1107-11.

Ruzzi, L., Gagnoux-Palacios, L., Pinola, M., Belli, S., Meneguzzi, G., D'alessio, M. & Zambruno, G. (1997). A homozygous mutation in the integrin alpha6 gene in junctional epidermolysis bullosa with pyloric atresia. *J Clin Invest,* **99,** 2826-31.

Schuffler, M. D., Bird, T. D., Sumi, S. M. & Cook, A. (1978). A familial neuronal disease presenting as intestinal pseudoobstruction. *Gastroenterology,* **75,** 889-98.

Schulz, A. C., Bartels, E., Stressig, R., Ritgen, J., Schmiedeke, E., Mattheisen, M., Draaken, M., Ludwig, M., Bagci, S., Muller, A., et al. (2012). Nine new twin pairs with esophageal atresia: a review of the literature and performance of a twin study of the disorder. *Birth Defects Res A Clin Mol Teratol,* **94,** 182-6.

Schwarz, J. M., Rodelsperger, C., Schuelke, M. & Seelow, D. (2010). MutationTaster evaluates disease-causing potential of sequence alterations. *Nat Methods,* **7,** 575-6.

Self, M., Geng, X. & Oliver, G. (2009). Six2 activity is required for the formation of the mammalian pyloric sphincter. *Dev Biol,* **334,** 409-17.

812 **Shah, N., Rodriguez, M., Louis, D. S., Lindley, K. & Milla, P. J.** (1999). Feeding difficulties and foregut
813 dysmotility in Noonan's syndrome. *Arch Dis Child,* **81,** 28-31.
814 **Sherwood, R. I., Chen, T. Y. & Melton, D. A.** (2009). Transcriptional dynamics of endodermal organ
815 formation. *Dev Dyn,* **238,** 29-42.
816 **Siepel, A., Bejerano, G., Pedersen, J. S., Hinrichs, A. S., Hou, M., Rosenbloom, K., Clawson, H.,**
817 **Spieth, J., Hillier, L. W., Richards, S., et al.** (2005). Evolutionarily conserved elements in
818 vertebrate, insect, worm, and yeast genomes. *Genome Res,* **15,** 1034-50.
819 **Singleton, M. V., Guthery, S. L., Voelkerding, K. V., Chen, K., Kennedy, B., Margraf, R. L., Durtschi, J.,**
820 **Eilbeck, K., Reese, M. G., Jorde, L. B., et al.** (2014). Phevor combines multiple biomedical
821 ontologies for accurate identification of disease-causing alleles in single individuals and small
822 nuclear families. *Am J Hum Genet,* **94,** 599-610.
823 **Smith, D. M., Grasty, R. C., Theodosiou, N. A., Tabin, C. J. & Nascone-Yoder, N. M.** (2000a).
824 Evolutionary relationships between the amphibian, avian, and mammalian stomachs. *Evol*
825 *Dev,* **2,** 348-59.
826 **Smith, R. L., Edwards, M. J., Notaras, E. & O'loughlin, E. V.** (2000b). Esophageal dysmotility in
827 brothers with an FG-like syndrome. *Am J Med Genet,* **91,** 185-9.
828 **Solomon, B. D.** (2011). VACTERL/VATER Association. *Orphanet J Rare Dis,* **6,** 56.
829 **Solomon, B. D., Bear, K. A., Kimonis, V., De Klein, A., Scott, D. A., Shaw-Smith, C., Tibboel, D.,**
830 **Reutter, H. & Giampietro, P. F.** (2012). Clinical geneticists' views of VACTERL/VATER
831 association. *Am J Med Genet A,* **158a,** 3087-100.
832 **Sorensen, H. T., Norgard, B., Pedersen, L., Larsen, H. & Johnsen, S. P.** (2002). Maternal smoking and
833 risk of hypertrophic infantile pyloric stenosis: 10 year population based cohort study. *BMJ,*
834 **325,** 1011-2.
835 **Stephens, D. N., Klein, R. H., Salmans, M. L., Gordon, W., Ho, H. & Andersen, B.** (2013). The Ets
836 transcription factor EHF as a regulator of cornea epithelial cell identity. *J Biol Chem,* **288,**
837 34304-24.
838 **Stringer, E. J., Pritchard, C. A. & Beck, F.** (2008). Cdx2 initiates histodifferentiation of the midgut
839 endoderm. *FEBS Lett,* **582,** 2555-60.
840 **Svenningsson, A., Soderhall, C., Persson, S., Lundberg, F., Luthman, H., Chung, E., Gardiner, M.,**
841 **Kockum, I. & Nordenskjold, A.** (2012). Genome-wide linkage analysis in families with
842 infantile hypertrophic pyloric stenosis indicates novel susceptibility loci. *J Hum Genet,* **57,**
843 115-21.
844 **Tanner, M. S., Smith, B. & Lloyd, J. K.** (1976). Functional intestinal obstruction due to deficiency of
845 argyrophil neurones in the myenteric plexus. Familial syndrome presenting with short small
846 bowel, malrotation, and pyloric hypertrophy. *Arch Dis Child,* **51,** 837-41.

Taylor, A. I. (1968). Autosomal trisomy syndromes: a detailed study of 27 cases of Edwards' syndrome and 27 cases of Patau's syndrome. *J Med Genet,* **5,** 227-52.

Tugay, M., Yildiz, F., Utkan, T., Sarioglu, Y. & Gacar, N. (2003). Gastric smooth muscle contractility changes in the esophageal atresia rat model: an in vitro study. *J Pediatr Surg,* **38,** 1366-70.

Tugay, M., Yildiz, F., Utkan, T., Ulak, G., Gacar, N. & Erden, F. (2001). Impaired esophageal reactivity in adriamycin-induced rat esophageal atresia: an in vitro study. *J Pediatr Surg,* **36,** 1569-73.

Unger, S., Mainberger, A., Spitz, C., Bahr, A., Zeschnigk, C., Zabel, B., Superti-Furga, A. & Morris-Rosendahl, D. J. (2007). Filamin A mutation is one cause of FG syndrome. *Am J Med Genet A,* **143a,** 1876-9.

Van Beelen, N. W., Mous, D. S., Brosens, E., De Klein, A., Van De Ven, C. P., Vlot, J., Ijsselstijn, H. & Wijnen, R. (2014). Increased incidence of hypertrophic pyloric stenosis in esophageal atresia patients. *Eur J Pediatr Surg,* **24,** 20-4.

Van Staey, M., De Bie, S., Matton, M. T. & De Roose, J. (1984). Familial congenital esophageal atresia. Personal case report and review of the literature. *Hum Genet,* **66,** 260-6.

Vanderwinden, J. M., Mailleux, P., Schiffmann, S. N., Vanderhaeghen, J. J. & De Laet, M. H. (1992). Nitric oxide synthase activity in infantile hypertrophic pyloric stenosis. *N Engl J Med,* **327,** 511-5.

Varki, R., Sadowski, S., Pfendner, E. & Uitto, J. (2006). Epidermolysis bullosa. I. Molecular genetics of the junctional and hemidesmosomal variants. *J Med Genet,* **43,** 641-52.

Veenma, D., Brosens, E., De Jong, E., Van De Ven, C., Meeussen, C., Cohen-Overbeek, T., Boter, M., Eussen, H., Douben, H., Tibboel, D., et al. (2012). Copy number detection in discordant monozygotic twins of Congenital Diaphragmatic Hernia (CDH) and Esophageal Atresia (EA) cohorts. *Eur J Hum Genet,* **20,** 298-304.

Verzi, M. P., Stanfel, M. N., Moses, K. A., Kim, B. M., Zhang, Y., Schwartz, R. J., Shivdasani, R. A. & Zimmer, W. E. (2009). Role of the homeodomain transcription factor Bapx1 in mouse distal stomach development. *Gastroenterology,* **136,** 1701-10.

Vivona, G., Frontali, M., Di Nunzio, M. L. & Vendemiati, A. (1987). Aplasia cutis congenita and/or epidermolysis bullosa. *Am J Med Genet,* **26,** 497-502.

Warren, J., Evans, K. & Carter, C. O. (1979). Offspring of patients with tracheo-oesophageal fistula. *J Med Genet,* **16,** 338-40.

Woo, J., Miletich, I., Kim, B. M., Sharpe, P. T. & Shivdasani, R. A. (2011). Barx1-mediated inhibition of Wnt signaling in the mouse thoracic foregut controls tracheo-esophageal septation and epithelial differentiation. *PLoS One,* **6,** e22493.

Yandell, M., Huff, C., Hu, H., Singleton, M., Moore, B., Xing, J., Jorde, L. B. & Reese, M. G. (2011). A probabilistic disease-gene finder for personal genomes. *Genome Res,* **21,** 1529-42.

882 **Zemojtel, T., Kohler, S., Mackenroth, L., Jager, M., Hecht, J., Krawitz, P., Graul-Neumann, L.,**
883 **Doelken, S., Ehmke, N., Spielmann, M., et al.** (2014). Effective diagnosis of genetic disease
884 by computational phenotype analysis of the disease-associated genome. *Sci Transl Med,* **6,**
885 252ra123.
886 **Zhang, R., Marsch, F., Kause, F., Degenhardt, F., Schmiedeke, E., Marzheuser, S., Hoppe, B.,**
887 **Bachour, H., Boemers, T. M., Schafer, M., et al.** (2017). Array-based molecular karyotyping in
888 115 VATER/VACTERL and VATER/VACTERL-like patients identifies disease-causing copy
889 number variations. *Birth Defects Res,* **109,** 1063-1069.
890 **Zwink, N., Choinitzki, V., Baudisch, F., Holscher, A., Boemers, T. M., Turial, S., Kurz, R., Heydweiller,**
891 **A., Keppler, K., Muller, A., et al.** (2016). Comparison of environmental risk factors for
892 esophageal atresia, anorectal malformations, and the combined phenotype in 263 German
893 families. *Dis Esophagus,* **29,** 1032-1042.
894

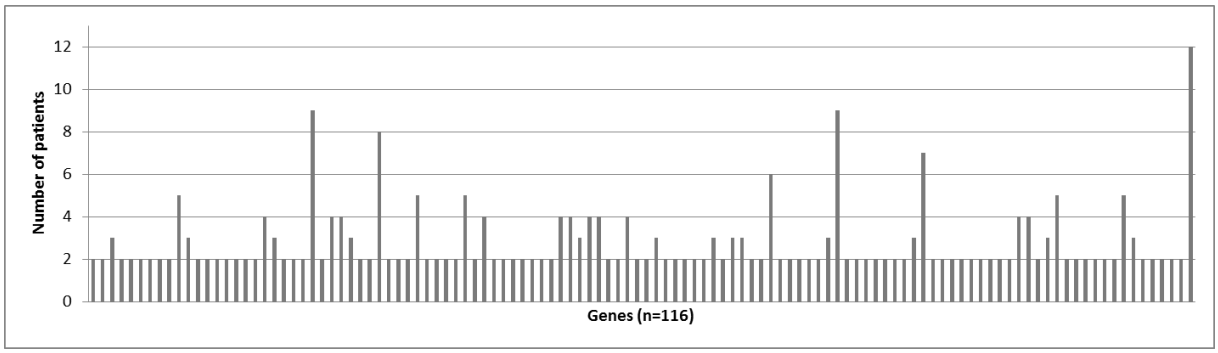

Figure 1. Number of patients with variants per gene

36 genes were found in ≥3 patients of which six genes were present in more than five patients (*CNTN2*, *DSPP*, *NOTCH4*, *PRRC2A*, *SEC16B*, *ZNF717*). See also Table S6.

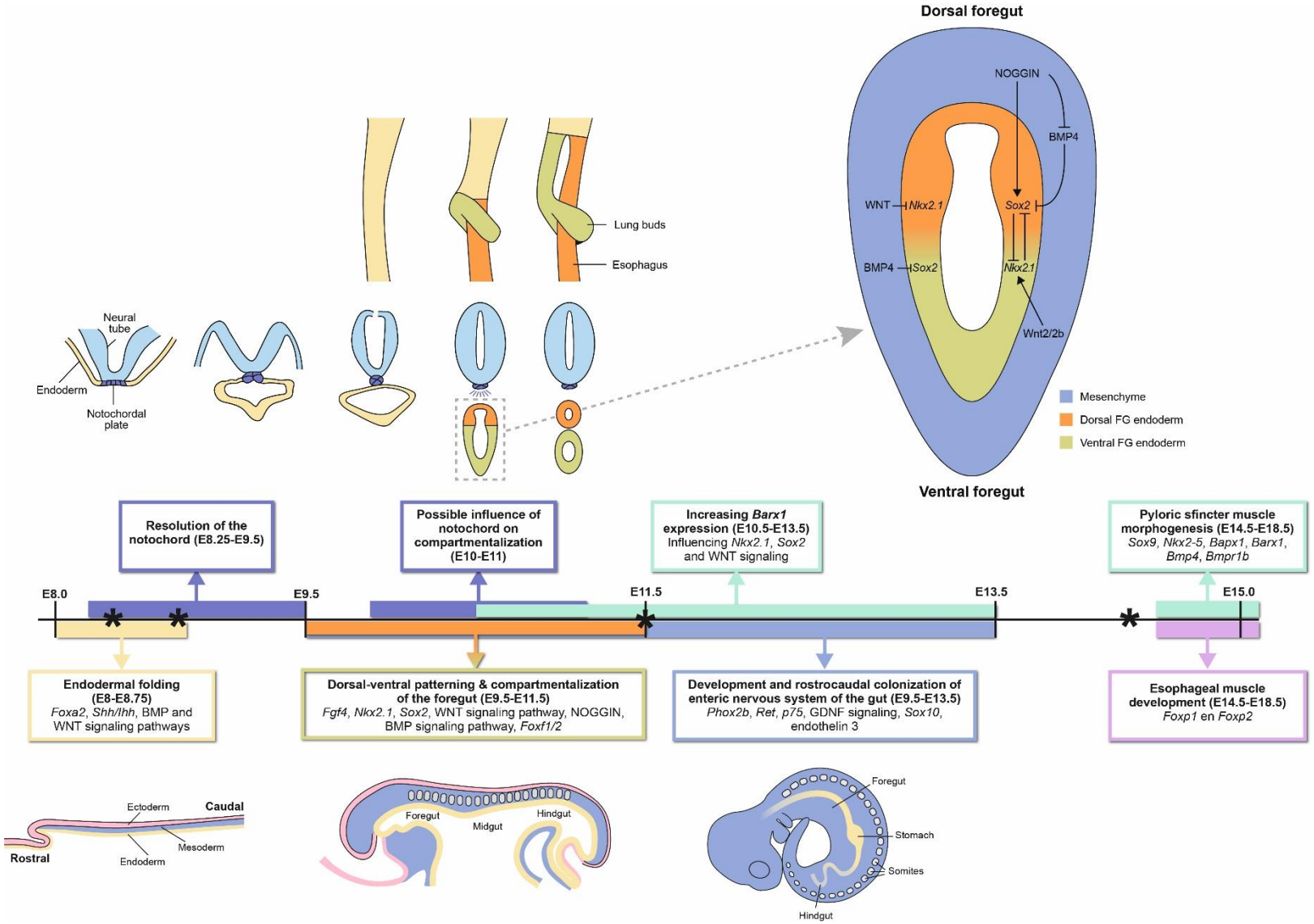

Figure 2. Timeline of models and genes known to be important for foregut development in mice

(Heath, 2010, Fausett and Klingensmith, 2012, Perin et al., 2017, Anderson et al., 2006), * = time points used in expression analysis (see Table S5)

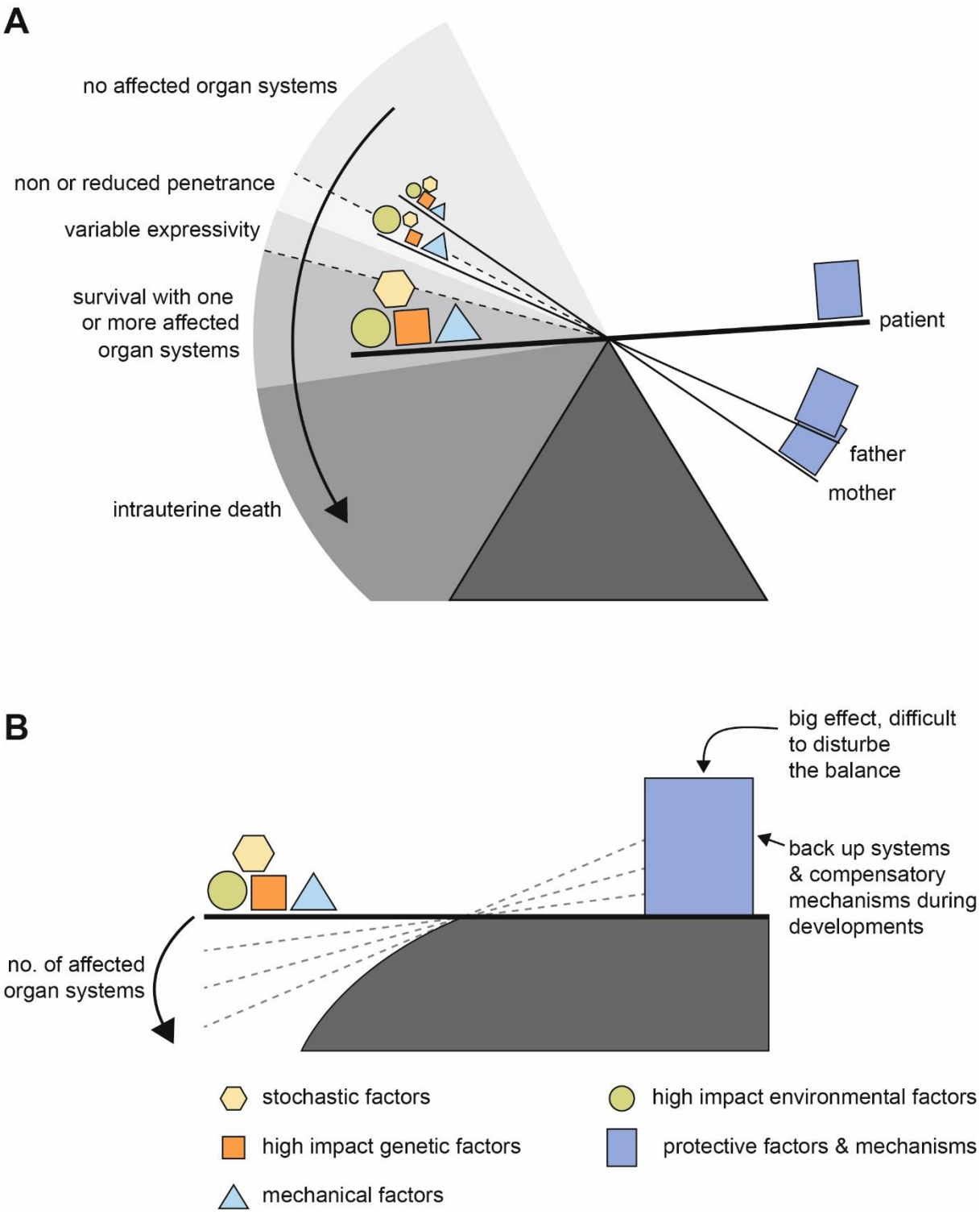

Figure 3. Two models for EA/IHPS etiology

A = burden model, B = slippery slope model. The combination of multiple high impact factors (genetic, environmental, mechanical and/or stochastic) together can modulate the phenotypical spectrum. These risk factors are in balance with protective factors like backup systems and compensatory mechanisms.

908 **Table 1. Genetic syndromes and mutated genes with tracheoesophageal and pyloric anomalies as variable features**
909

Syndrome	Esophageal or pyloric anomaly	Inheritance	Loci	Gene(s)	OMIM	Ref
Esophageal atresia or stenosis						
Epidermolysis bullosa, junctional, with pyloric stenosis or atresia [C]	Esophageal and pyloric atresia or stenosis	AR	2q31.1 17q25.1	*ITGA4* *ITGA6*	226730 226730	(Varki et al., 2006, Vivona et al., 1987) (Ruzzi et al., 1997)
Ehlers-Danlos syndrome [C]	EA and IHPS	AD	2q32.2	*COL3A1*	130050	(Kroes et al., 2003, Kuivaniemi et al., 1990)
Trisomy 13	EA/TEF and IHPS	AD	13	multiple	NA	(Brosens et al., 2014, Taylor, 1968)
Trisomy 18	EA/TEF and IHPS	AD	18	multiple	NA	(Brosens et al., 2014, Taylor, 1968)
Trisomy 21	EA/TEF and IHPS	AD	21	multiple	190685	(Brosens et al., 2014, Freeman et al., 2009)
Fryns syndrome	EA/TEF and IHPS	U	unknown	unknown	229850	(Ayme et al., 1989)
Fetal alcohol syndrome	EA/TEF and IHPS	NA	NA	NA	NA	(Brosens et al., 2014, Lodha et al., 2005, Mangyanda et al., 1998)
Motility anomalies of the esophagus						
Epidermolysis bullosa dystrophia [C]	Esophageal strictures and stenosis	AR, AD AR	3p21.31 11q22.2	*COL7A1* *MMP1*	131750 226600	(Christiano et al., 1995, Hovnanian et al., 1994, Christiano et al., 1996)
Cornelia de Lange syndrome [B,C]	Esophageal stenosis and dysmotility and IHPS	AD	5p13.2	*NIPBL*	122470	(Cates et al., 1989, Gillis et al., 2004)
Apert syndrome	Esophageal stenosis and IHPS	AD	10q26.13	*FGFR2*	101200	(Pelz et al., 1994, Blank, 1960)
Congenital generalized lipodystrophy	Esophageal dysmotility and IHPS	AR	17q21.2	*PTRF*	613327	(Rajab et al., 2010, Rajab et al., 2002)
Opitz-Kaveggia syndrome	Nutcracker esophagus and IHPS	XL	Xq13	*MED12*	305450	(Smith et al., 2000b, Battaglia et al., 2006)
Noonan syndrome [C]	Esophageal dysmotility and IHPS	AD	12q24.13	*PTPN11*	163950	(Shah et al., 1999, Barberia Leache et al., 2003)
Visceral neuropathy	Dilated non-peristaltic esophagus and IHPS	U	unknown	unknown	243180	(Schuffler et al., 1978, Tanner et al., 1976)
Costello syndrome	Loss of elastic fibers in esophagus, IHPS	AD	11p15.5	*HRAS*	218040	(Mori et al., 1996, Gripp and Lin, 1993)
Other associations						
Chronic idiopathic intestinal pseudo obstruction [B, C]	Gastro-intestinal dysmotility and IHPS	XL	Xq28	*FLNA*	300048	(Tanner et al., 1976, Gargiulo et al., 2007)
Fronto-metaphyseal dysplasia [B]	EA/TEF	XL	Xq28	*FLNA*	305620	(Franceschini et al., 1997)
X-linked periventricular heterotopia [B]	IHPS	XL	Xq28	*FLNA*	300049	(Nezelof et al., 1976)
FG syndrome [B, C]	Esophageal dysmotility and IHPS	XL	Xq28	*FLNA*	300321	(Unger et al., 2007, Peeters et al., 2012)
CHARGE syndrome [B, C]	EA/TEF	AD	8q12.1-q12.2	*CHD7*	214800	(Brosens et al., 2014)
Hypogonadotropic hypogonadism with or without anosmia [B, C]	IHPS [A]	AD	8q12.1-q12.2	*CHD7*	612370	(Jongmans et al., 2009, Kim et al., 2008)

910 This table is modified from two reviews on esophageal atresia (Brosens et al., 2014) and infantile hypertrophic pyloric stenosis (Peeters et al., 2012). AD; autosomal dominant, AR;
911 autosomal recessive, U; unknown, NA; not applicable, XL; X-linked; EA; esophageal atresia, TEF; trachea-esophageal fistula, IHPS; infantile hypertrophic pyloric stenosis. [A] In literature
912 IHPS is associated with other genes responsible for this syndrome. [B] No overlap in EA and IHPS phenotype for this syndrome, the gene mutated in this syndrome can be responsible
913 for different syndromes in which either EA or IHPS are variable features. [C] More genes associated to possible several subtypes of this syndrome.

Table 2. Phenotype description

Individual	Gender	EA-type	Phenotype	Remarks
SKZ_0027	female	C	EA/TEF, IHPS, thin ear helix, seizures	-
SKZ_0096	male	C	EA/TEF, IHPS, syndactyly second-third finger, radial dysplasia, abnormal fibula	VACTERL association
SKZ_0244	male	C	EA/TEF, IHPS, anal atresia, intestinal malrotation, sacral dimple, abnormal os coccygis, abnormal vertebrae L1, thenar hypoplasia, both sides hypoplastic "floating" thumbs, both sides dysplastic radii	VACTERL association, mother is a DES daughter
SKZ_0321	male	C	EA/TEF, IHPS, mild left sided expansion of the pyelocaliceal system , breath holding spells	-
SKZ_0353	female	C	EA/TEF, IHPS, sacral dimple, thin/slender build, diminished hearing, palpebral fissures slant up, hemolytic anemia, short phalanges	Glucose-6-phosphate dehydrogenase deficiency
SKZ_0399	male	C	EA/TEF, IHPS, anal atresia, sacral dimple, 2 umbilical vessels, posteriorly rotated ears, small ears/microtia, flat face, bifid scrotum, small penis/micropenis, small palmar crease, thick fingers, broad thumbs, proximal placement of thumbs, microstomia, thick broad neck, wide nasal bridge, patent ductus arteriosis, 4th toe abnormally placed	VACTERL association
SKZ_0400	male	C	EA/TEF, IHPS, extra ribs, fusion of vertebrae, macrocephaly, bulbar dermoid cyst, auricular tags, short thick/broad neck	Klippel-Feil syndrome
SKZ_0683	male	C	EA/TEF, IHPS, sacral dimple	-
SKZ_0760	male	C	EA/TEF, IHPS, hemivertebrae, bitemporal narrowing of the head, prominent forehead, hyper mobile/ extensible fingers, narrow thorax/funnel chest, thin lower and upper lip, spasticity, cerebral palsy	-
SKZ_0788	male	C	EA/TEF, IHPS, inguinal hernia, jaundice, deafness	-
SKZ_0790	female	C	EA/TEF, IHPS	-
SKZ_0796	male	C	EA/TEF, IHPS	Vanishing twin
SKZ_0848	male	C	EA/TEF, IHPS, sacral dimple, hypospadias, patent ductus arteriosus	-
SKZ_0887	male	C	EA/TEF, IHPS, abnormal sacrum, fusion of vertebrae, posteriorly rotated ears, small mandible/micrognathia, rocker-bottom feet, sandal gap of toes, open mouth appearance, short neck, jaundice	-
SKZ_1003	male	C	EA/TEF, IHPS, abnormal sacrum, cleft jaw, cleft palate, cleft upper lip, depressed/flat nasal bridge, fused ribs	Methyldopa (aldomet) for hypertension during pregnancy
SKZ_1248	female	C	EA/TEF, IHPS, small large fontanel, deafness, small ears, auricular tags, single palmar crease, small/hypoplastic deep set ears	-
SKZ_1260	male	C	EA/TEF, IHPS, syndactyly of second-third toe, bifid/fused ribs	-
SKZ_1353	male	C	EA/TEF, IHPS, cleft uvula, epicanthic folds, abnormal dermatoglypic patterns, hyperconvex/clubbed nails, hypoplastic scrotum, hypospadias, bifid scrotum, hydrocele of testis	-
SKZ_1407	female	A	EA, IHPS	-
SKZ_1472	male	C	EA/TEF, IHPS, eczema of hands with hyperhidrosis, blisters and erythema, Xerosis Cutis	Antibiotics for respiratory infection during pregnancy
SKZ_1961	male	C	EA/TEF, IHPS, sacral dimple, mild dysmorphic features, small mouth, pointy ears, long fingers	Maternal hypertension
SKZ_2013	male	A	EA, IHPS, persistent superior vena cava, scoliosis, Horner's syndrome	-
SKZ_2023	male	C	EA/TEF, IHPS, small chin, sacral dimple	-
SKZ_2050	male	C	EA/TEF, IHPS, atrial septum defect	
SKZ_2082	male	C	EA/TEF, IHPS, persistent tracheolaryngeal cleft, anal atresia, atrial septum defect, tracheal-laryngeal anomaly, prostate fistula	VACTERL association
SKZ_2149	male	C	EA/TEF, IHPS	-
SKZ_2171	female	C	EA/TEF, IHPS, spina bifida Th10/11, synostoses vertebrae, hydronephrosis, kyphoscoliosis	Unknown medication for headaches and nerves during pregnancy

EA; esophageal atresia, TEF; trachea-esophageal fistula, IHPS; infantile pyloric stenosis, DES; di-ethylstilbestrol. EA-type classification according to Gross classification (Gross, 1947)

Table 3. Results for gene burden analysis

Gene	Chr.	Start position	rsID	HGVS c	VAAST p-value
ALMS1	chr2	73828337	rs201777220	c.11885T>C	1,00E-06
ALMS1	chr2	73746907		c.9543delA	1,00E-06
SLC28A3	chr9	86895771	rs762396296	c.1673dupA	0,002853
SLC28A3	chr9	86912927	rs775359011	c.677G>A	0,002853
SP2	chr17	46002296	rs778698435	c.1384C>A	0,002853
SP2	chr17	45993628	rs754352270	c.191C>T	0,002853
EPB41	chr1	29435944		c.2410A>G	0,019802
EPB41	chr1	29391549	rs768609152	c.2063T>G	0,019802
AMBRA1	chr11	46563580	rs755884183	c.1717C>G	0,019802
AMBRA1	chr11	46567274		c.431C>T	0,019802
VWA8	chr13	42293773	rs371770462	c.3070G>A	0,019934
VWA8	chr13	42439870	rs201163045	c.1425+2T>C	0,019934
CLGN	chr4	141320158	rs201306926	c.731A>G	0,039604
CLGN	chr4	141315036	rs200652126	c.1309G>T	0,039604
SDK2	chr17	71357964	rs147983543	c.5326G>A	0,049505
SDK2	chr17	71415318		c.2173G>A	0,049505
PDLIM7	chr5	176911101	rs200609502	c.1141A>G	0,049505
PDLIM7	chr5	176916502	rs764108486	c.761C>T	0,049505
GUCY2F	chrX	108652306	rs7883913	c.1883G>A	0,019802
GUCY2F	chrX	108638614	rs35726803	c.2380G>A	0,019802

N/A = not applicable

922 **Table 4. Summary of overlapping top candidate genes**

	EA/IHPS patients (n=15)				Healthy controls (n=44)			
	Rare (MAF ≤0.001%)		Ultra-rare (MAF 0%)		Rare (MAF ≤0.001%)		Ultra-rare (MAF 0%)	
	LB	VUS/LD	LB	VUS/LD	LB	VUS/LD	LB	VUS/LD
Important in normal foregut development (Fig. 2)	1	1	0	0	3	5	1	2
Genes associated with genetic syndromes involving both EA and IHPS as variable features (Table 1)	1	0	0	0	2	4	1	1
Genes associated with IHPS (Peeters et al., 2012)	0	0	0	0	0	2	0	0
Genes involved in neuromuscular and connective tissue syndromes associated with IHPS (Peeters et al., 2012)	1	1	1	0	1	2	1	0
Genes involved in syndromes and signaling disturbances associated with IHPS (Peeters et al., 2012)	0	3	0	0	1	6	1	1
Genes involved in ciliopathies and disturbances of gene regulation associated with IHPS (Peeters et al., 2012)	0	1	0	0	4	2	1	0
Genes involved in lymphatic abnormalities and syndromes of environmental and unknown origin associated with IHPS (Peeters et al., 2012)	0	0	0	0	1	0	0	0
Genes involved in genetic syndromes and abnormalities	1	1	1	1	6	6	0	1

923 Variant allele count per gene. LB = likely benign, VUS = variant of unknown significance, LD = likely deleterious.
924 See Table S6 for the complete results, adapted from Peeters et al. en Brosens et al. (Peeters et al., 2012, Brosens et
925 al., 2014).
926
927
928 **Table 5. Comparison with control cohort: number of variants**
929

	Ultra-rare variants (MAF 0%)				Rare variants (MAF ≤0.001%)			
	EA/IHPS patients (n=15)		Healthy controls (n=44)		EA/IHPS patients (n=15)		Healthy controls (n=44)	
		#		#		#		#
Putative deleterious	296	*19.73*	725	*16.48*	912	*60.8*	2667	*60.6*
LOF intolerant	28	*1.87*	85	*1.93*	81	*5.4*	272	*6.2*
De novo variants	0	-	Unknown	-	0	-	Unknown	-
Recessive	291	*19.4*	715	*16.25*	898	*59.9*	2631	*59.80*
Compound heterozygous	3	*0.2*	3	*0.06*	8	*0.53*	6	*0.14*
X-linked	11	*0.73*	23	*0.52*	28	*1.87*	95	*2.16*

930 *#* = number of variants divided by the number of patients. A Chi^2 test showed no significant
931 differences between the patients and the healthy controls (all p>0.05)
932