Title: Pulmonary transcriptome analysis in the rabbit model of surgically-induced diaphragmatic hernia treated with fetal tracheal occlusion.

Authors: Alexander C. Engels\textsuperscript{1,3}, Paul D. Brady\textsuperscript{2}, Molka Kammoun\textsuperscript{2}, Julio Finalet Ferreiro\textsuperscript{2}, Philip DeKoninck\textsuperscript{1,3}, Masayuki Endo\textsuperscript{1}, Jaan Toelen\textsuperscript{1,4}, Joris R. Vermeesch\textsuperscript{2}, Jan Deprest\textsuperscript{1,3}

\textsuperscript{1} Department of Development and Regeneration, Organ System Cluster, Faculty of Medicine, KU Leuven, Leuven, Belgium

\textsuperscript{2} Centre for Human Genetics, University Hospitals KU Leuven, Leuven, Belgium

\textsuperscript{3} Clinical Department of Obstetrics and Gynaecology, Division Woman and Child, University Hospitals KU Leuven, Leuven, Belgium

\textsuperscript{4} Clinical Department of Pediatrics, Division Woman and Child, University Hospitals KU Leuven, Leuven, Belgium

Keywords: RNA-Seq, congenital diaphragmatic hernia, CDH, pulmonary hypoplasia, tracheal occlusion, fetoscopy

*Correspondence:
Jan Deprest, MD PhD FRCOG
Department of Obstetrics and Gynaecology
University Hospital Gasthuisberg
Herenstraat 49, B-3000 Leuven, Belgium
Tel : +32 16 34 42 15
Fax : + 32 16 34 42 05
E-mail: jan.deprest@uzleuven.be
**Summary statement:**

Rabbit fetuses with induced diaphragmatic hernia and treated with prenatal tracheal occlusion do have a similar pulmonary transcriptome as unaffected controls and provide a valuable database of gene expressions.

**Abstract:**

Congenital diaphragmatic hernia (CDH) is a malformation leading to pulmonary hypoplasia which can be treated *in utero* by fetal tracheal occlusion (TO). However the changes of gene expression induced by TO remain largely unknown but could be used to further improve the clinically used prenatal treatment of this devastating malformation. Therefore we aimed to investigate the pulmonary transcriptome changes due to surgical induction of diaphragmatic hernia (DH) and additional tracheal occlusion in the fetal rabbit model.

Induction of DH was associated with 378 up-regulated genes compared to controls when allowing a false discovery rate (FDR) of 0.1 and a Fold Change (FC) of 2. Those genes were again down-regulated by consecutive TO. But DH+TO was associated with an up-regulation of 157 genes compared to DH and controls. When being compared to control lungs, 106 genes were down-regulated in the DH group and were not changed by TO. Therefore, the overall pattern of gene expression in DH+TO is more similar to the control group then to the DH group. In this study we further provide a database of gene expression changes induced by surgical creation of DH and consecutive TO in the rabbit model. Future treatment strategies could be developed using this dataset. We also discuss the most relevant genes that are involved in CDH.
Introduction

The prevalence of congenital diaphragmatic hernia (CDH) ranges between 1-4/10,000 births, which translates to 542 to 2,168 children in EU-27 every year (2008) (Kotecha et al. 2012). It occurs as an isolated condition, associated with other anomalies or part of a genetic syndrome. Lung development is disturbed already from the embryonic period, but progresses because of herniation of the viscera through the defect into the chest, competing for space with the developing lungs. As a consequence, lungs of babies with CDH have fewer and less mature airway branches, a smaller cross-sectional area of pulmonary vessels, remodelled vascular architecture and an altered vasoreactivity (Kinsella et al. 2005). At birth, this leads to respiratory insufficiency and persistent pulmonary hypertension (PPHT). This is lethal in up to 30% of cases, despite prenatal referral to high volume centers that offer standardized neonatal care (Grushka et al., 2009; Hayakawa et al., 2013; van den Hout et al., 2011).

In countries with prenatal ultrasound (US) screening practices, >60% of cases are diagnosed at the latest by the second trimester. In isolated cases an individualized prognosis can be made based on the degree of liver herniation and parenchymal lung size predicting mortality and early neonatal morbidity (Cannie et al., 2008; Claus et al., 2011; Cruz-Martinez et al., 2011; Jani et al., 2007; Jani et al., 2009; Mayer et al., 2011; Ruano et al., 2012). The ability to prenatally identify a future non-survivor enables clinicians to propose a prenatal intervention that avoids this outcome. Today the clinical method to stimulate lung development is fetal tracheal occlusion (TO). TO prevents egress of lung liquid, causing increased pulmonary stretch, which accelerates lung growth (DiFiore et al., 1994). This operation is currently evaluated within the framework of a randomized clinical trial, comparing Fetoscopic Endoluminal Trachea Occlusion (FETO) to expectant management during pregnancy (Deprest et al., 2009) (Deprest et al., 2014).

Because TO still fails to save half of the fetuses with severe CDH and carries a risk for preterm delivery, there is a need for alternative prenatal strategies. Preclinical evaluation of these therapies involves the use of animal models i.e. rodents, rabbits, and eventually lambs. Though in all mammalians, lung development progresses through five stages (embryonic, pseudoglandular, canalicular, saccular, alveolar), this is not at the same pace. (Pringle, 1986). This needs to be kept in mind when choosing an animal model in studies on lung development. For instance, rodents are “early” models as CDH is induced in the embryologic phase by the administration of a teratogen (Clugston et al., 2006; Keijzer et al., 2000). In those, nitrofen (NF) is prenatally administered at E11.5 (mice) or E13.5 (rats), and effects are
studied at term, when lungs are in the canalicular to saccular phase (Beurskens et al., 2007; Correia-Pinto et al., 2010). All NF-exposed pups have lung hypoplasia, half of them also the diaphragmatic defect (Keijzer et al., 2000; Rottier and Tibboel, 2005; Schittny et al., 2000). These effects are believed to be caused by interference of NF with the retinoic acid (RA) signaling pathway, which is key in lung development (Malpel et al., 2000). This pathway is also disturbed in CDH in humans (Beurskens et al., 2009).

In larger animals hypoplasia is induced by surgical creation of a diaphragmatic defect, typically in the pseudoglandular phase. Rabbits are widely available, have low housing needs, a timed gestation and large litter size. Rabbits alveolize in utero as in man (Roubliova et al., 2010). Pups with diaphragmatic hernia (DH) display both histological and functional changes such as reduced airway and vascular development, and pathologic compliance, airway resistance, tissue damping and elastance – mimicking the clinical phenotype (Flemmer et al., 2007; Roubliova et al., 2004; Wu et al., 2000). Gene expression of a number of critical signaling molecules relevant to alveolarization, angiogenesis and regulation of vascular tone, but not to surfactant production, have been shown to be disrupted just as in man (Vuckovic et al., 2013; Vuckovic et al., 2012). However a broader study on gene expression levels in this model has not been done so far. The use of RNA-sequencing for transcriptome analysis has become increasingly widespread with the advent of massively parallel sequencing technologies, in part due to reductions in costs and increased throughput, and improved knowledge of non-model organism reference genomes. Therefore we wanted to investigate the pulmonary transcriptome after surgical induced DH creation and subsequent TO in the rabbit model. The provided gene expression database can be used to develop further treatment strategies for CDH.
Results

At harvest there were seven surviving DH+TO fetuses (mean LBWR: 0.017; SD: 0.002; CI 95%: 0.013 – 0.022) and seven DH fetuses (mean LBWR: 0.011; SD: 0.003; CI 95%: 0.003 – 0.018). We also took one random control for every third litter (n=4) (mean LBWR: 0.016; SD: 0.001; CI 95%: 0.013 – 0.018). RNA integrity (RIN) values were determined which led to the exclusion of one DH+TO fetus. All six remaining fetuses from the DH+TO group were analyzed. We selected four (DH) fetuses displaying the best RIN values (data not shown). Principal component analysis (PCA) and unsupervised hierarchical clustering revealed 3 DH+TO samples clustered together as a single homogenous group, and 2 DH+TO samples clustered with the DH group. Those 2 samples correlate with a clinical sub-group of non-responders to the TO treatment which analysis was not the aim of our study. The remaining DH+TO sample was considered as a statistical outlier based upon QC correlation analysis and calculation of MAD scores using the Array Studio software, hence it was excluded from further analysis. This distinct gene expression pattern of DH+TO fetuses correlated with two patterns of lung growth, as evidenced by the lung-to-body-weight ratio, which is a pathologic measure of the degree of lung development (Table 1).

Statistical inference analysis was performed between the groups defined above [DH (n=4), DH+TO (n=3), and control (n=4)] using the general linear model function in Array Studio (OmicSoft, Cary, US) to calculate fold change (FC) values and false discovery rate (FDR) values (applying the Benjamini–Hochberg procedure).

PCA and unsupervised hierarchical clustering of those dysregulated genes in any group comparison (described above) demonstrated clustering of homogenous groups for DH, DH+TO and control as expected, and revealed that DH+TO and control groups were more comparable than the DH group (Supplementary Figure 1, 2).

The heat map generated by the unsupervised hierarchical clustering for the total of 641 unique genes found to be dysregulated (FC +2; FDR <0.1) in any group comparison is shown in Figure 1. This reveals three large gene clusters whose expression values define the changes due to DH and after treatment with TO. The first gene cluster containing of 157 genes (Figure 1, top) demonstrates low expression in both DH and control, and high expression in DH+TO only. This cluster represents genes whose expression is induced by TO, but which are unchanged in DH. The second, and largest, gene cluster containing of 378 genes (Figure 1,
middle) demonstrates high expression in DH only and low expression in control and DH+TO. This cluster represents genes whose expression is induced by DH, and which are reversed to lower levels by subsequent TO. The third gene cluster containing of 106 genes (Figure 1, bottom) demonstrates high expression in controls, and low expression in both DH and DH+TO. This cluster represents genes whose expression is reduced following DH, and which remain expressed at low levels even after TO.

To summarize, from the hierarchical cluster heat map shown in Figure 1 it is apparent that the major change in gene expression (Figure 1, gene cluster 2, middle) is up regulation of genes in DH, which are down regulated by TO, towards, or beyond, levels seen in the controls (normal state) at the same developmental time point. One subset of genes is unchanged by TO and remains expressed at low levels, and one subset is highly upregulated by TO in comparison to both control and DH without TO.

The database of all dysregulated genes found in our study can be found in Supplementary Table 1.

Pathway Analysis

Analysis of the dysregulated gene sets for each comparison detailed above and for each of the gene clusters defined above was performed using the IPA web application (Ingenuity Pathway Analysis, Ingenuity Systems, Redwood City, CA, USA). Of the total 641 unique genes, 560 genes were mapped within IPA to known functional genes with human, mouse or rat homologues. This gene set was used to create a protein-protein interaction (PPI) network where experimental evidence supports direct interactions (i.e. predictions and indirect interactions were not allowed) and orphan molecules were subsequently removed. The resulting network is displayed in Supplementary Figure 3, where the molecules are coloured according to the gene clusters identified in Figure 1 (blue = genes with significantly increased expression in DH compared to DH+TO and/or control; yellow = genes with significantly increased expression in DH+TO compared to DH and/or control; and green = genes with significantly decreased expression in DH and/or DH+TO in comparison to control).

The Cytoscape plugin iRegulon (Janky et al., 2014) was used to identify potential transcription factors (TFs) of interest, using the 641 genes dysregulated in any comparison as the input. iRegulon searches for enrichment of cis-regulatory TF motifs in the target gene set,
and predicted FOXJ1 as a key regulatory TF in our study. The input list contains 139/641 (22%) genes, which are ‘target’ genes of FOXJ1.

In addition, we used the STEM application (Ernst and Bar-Joseph, 2006) to identify clusters of genes with similar changes in expression over time, or in our case the changes observed from control to disease (DH) to treatment (DH+TO). The input was the FC values for the 641 genes significantly dysregulated in any comparison, where timepoint 1 is control data (i.e. FC=1), timepoint 2 is the CDH vs control data, and timepoint 3 is the TO vs control data. This identified 4 significant clusters containing genes with similar changes in expression profile, of which Profile 9 contained the largest number of genes and was of highest significance (Figure 2, and Supplementary Table 2).

**Real Time Quantitative PCR (qPCR)**

PCR analysis of ten of the 12 analyzed genes yielded results that were in agreement with RNA sequencing data (Supplementary Figure 4). For eight genes, the differential expression between groups was concordant with the RNA-seq data (p<0.05, u Mann Whitney test). However, qPCR did not show a significant increase of RFX3 and LRRQI1 in DH group compared to control (Supplementary Figure 5). For BMPR2 and PDE5A, qPCR failed to show any difference between DH and wild type. This distinct gene expression pattern of DH+T fetuses correlated with two patterns of lung growth, as evidenced by the lung-to-body-weight ratio, which is a pathologic measure of the degree lung development. although PDE5A was significantly down regulated in the TO group.
Discussion

In this study we describe for the first time the pulmonary transcriptome analysis of specimens obtained in a rabbit model for pulmonary hypoplasia. The latter was induced by creating a diaphragmatic defect during the pseudoglandular phase. Conversely, forced lung growth was induced by fetal TO at the transition of the canalicular to saccular phase. We found that the largest group of genes significantly dysregulated were 378 genes that were both upregulated by DH creation and downregulated by TO to a level similar to that of controls. Further this study gives a database of genes that are significantly influenced by DH creation and consecutive TO (Supplementary Table 1). This database can be used for further understanding of the disease process and development of treatment modalities for CDH. Below we discuss some of the most relevant genes that we found dysregulated.

Relation of findings to previous gene expression analytical experiments in other models of CDH and/or TO

Many studies have documented expression changes for numerous genes in association with CDH, and to a lesser extent also the effects of TO, all of this in various animal models of CDH. This is typically done by using PCR for selected genes, or using broader arrays, at least for experiments done in (nitrofen exposed) rodents, a species where molecular tools are abundantly available. Using a more modern technique such as RNA-Seq one can now also document and screen for changes in gene expression in relevant animal models for pulmonary hypoplasia and induced lung growth, even if the genome has not been completely been identified. We herein used this technology in rabbits, and studied dysregulations associated to pulmonary hypoplasia and tracheal occlusion. The latter is done to force lung growth which has been abundantly demonstrated in several animal models of CDH using other tools (Khan et al., 2007). In this experiment we confirmed the effects of TO on cell proliferation. MKI67 is a key cell proliferation marker and was significantly upregulated by TO. The effect of TO was in proportion to the pre-existing lung size (FC 3.6056, FDR 0.0874; TO vs control), a phenomenon already clinically demonstrated (Peralta et al., 2008). Other genes related to the mechanisms of TO in previous studies were identified. To name only one, AQP4 was upregulated in DH (FC 2.3562, FDR 0.018) yet down to normal level after TO (FC -2.7286, FDR 0.0106). AQP4 is a transport molecule, which has previously been shown to be down regulated by stretch force on fetal lung AT2 cells in vitro (Wang et al., 2006).
TO is however in essence performed to enhance alveologenesis and reverse vascular changes that are typical to this disease. We observed dysregulation of Connective Tissue Growth Factor (CTGF). CTGF indeed enhances alveologenesis and microvascularisation during late lung development. CTGF was earlier shown to be upregulated by tracheal occlusion in the saccular phase of E21 rats (Burgos et al., 2010; Mesas-Burgos et al., 2009). In line with that, CTGF was in our rabbits also firmly upregulated by TO (FC 2.8135, FDR 0.0151). The same dysregulation was observed for a number of other genes earlier associated with DH and TO. BMPR2 expression was up in DH (FC 1.9265, FDR 0.0228) and went down after TO (FC -2.7956, FDR 0.0097). This gene encodes a member of the bone morphogenetic protein (BMP) receptor family of transmembrane serine/threonine kinases and the ligands of this receptor are the BMPs, which are members of the TGF-beta superfamily, and are involved in embryogenesis. Disruption of BMPR2 activity and downstream signaling has been demonstrated in the nitrofen rodent model of CDH (Gosemann et al., 2013; Makanga et al., 2013). BMPR2 plays a key role in pulmonary vasculogenesis during the late stages of fetal lung development and is essential for control of endothelial and smooth muscle cell proliferation. Dysfunction of BMPR2 and downstream signaling have been shown to disturb the crucial balance of proliferation of smooth muscle cells contributing to the pathogenesis of human and experimental pulmonary hypoplasia. In line with that was downregulation of BMP7 in DH, yet significant upregulation by TO (FC 2.311, FDR 0.0163). BMP7 was reportedly down regulated in the lungs of the nitrofen rodent (Makanga et al., 2013).

We were however mainly interested in the RNA-Seq technology because it has the potential to identify new targets that are involved in the condition and that can respond to prenatal intervention. One example is that DH in our rabbits was associated with a significant increase in PDE5A gene expression. The latter has been well documented in rodents and sheep, and it has been tied to the postnatal problem of pulmonary hypertension. The interesting observation here is that TO dramatically reduced PDE5A gene expression (FC= -2.49; FDR = 0.0097). PDE5A gene expression may also be altered by an alternative prenatal intervention, i.e. the transplacental administration of drugs. Sildenafil, which is already used to treat persistent pulmonary hypertension of the newborn (Noori et al., 2007), was recently shown to effectively inhibit cGMP-specific phosphodiesterase Type 5 in utero. In the latter experiment nitrofen rodents were used, and we confirmed a similar effect in rabbits and others in the fetal lamb (Luong et al., 2011; Russo et al., 2015; Shue et al., 2014).
Future application: single cell type populations

When starting from our RNA-seq data, the Cytoscape plugin iRegulon predicted that over 20% of dysregulated genes were target genes of FOXJ1. FOXJ1 is a marker for ciliated cells; it is actually involved as a transcription factor in ciliogenesis. To our knowledge it has so far not been named in DH. Our experiments demonstrate that DH rabbit pups display upregulation of FOXJ1, whereas it was down regulated after TO. Other ciliated cell marker genes (CCDC19, SPAG17, CCDC39, LRRIQ1, EFHC1, LRR2C3, TTC18, WDR16, FANK1, ENKUR, CCDC113) were also upregulated in DH lungs compared to controls, and downregulated by TO to levels lower than in controls. These findings may point to an increased numbers of ciliated cells in DH lungs, which is reversed by TO. Loss of ciliated cells after TO has been earlier described in the upper conducting airways in fetal lambs with CDH though that study did not quantify changes in the more distant airways (Deprest et al., 2000). The study of effects of DH and TO on the variety of individual lung cell types can benefit from modern technology such as single cell RNA-seq (Treutlein et al., 2014). Those investigators applied this to the developing distal lung epithelium in the mouse which revealed a number of putative markers next of ciliated cells, Clara cells, AT1 cells and AT2 cells. A number of genes in our data are also present within the top 30 putative markers for each lung cell type (both known and novel) previously identified in the Treutlein et al. study (Table 2). For instance, also Clara cell markers SCGB1A1, PPAP2B, and KDR (VEGFR2) were upregulated in DH, and subsequently down regulated, after TO. These findings are consistent with findings in rodents (Asabe et al., 1998) (Santos et al., 2007).

Discordant findings and limitations.

The above is a simplified interpretation and biased presentation of our results. First we identified a number of discrepancies in results between rabbits and other models. For example, AQP4, TGFBR3, EPAS1 and TGFB-I were upregulated in rabbits with CDH, whereas they were down regulated in the nitrofen rat. In rodents, DH is induced early in gestation by a teratogen interfering directly with the RA pathway, whereas in rabbits a surgical defect is created beyond mid gestation. This means that in both models different mechanisms or pathways are involved. Therefore, prenatal interventions may have different effects. For instance maternal retinoic acid administration in rabbits does not affect the gross anatomical, morphological or proliferative characteristics (Gallot et al., 2008), whereas it rescues lungs in nitrofen rats (Montedonico et al., 2006). Next to that there are also significant differences between the time points in gestation and/or lung development at what time prenatal interventions are done.
Actually, the largest effect measured was that for WNT inhibitor factor 1 (WIF-1) gene expression (up in DH (FC 8.8107, FDR 0.0153) and down after TO (FC -4.1432, FDR 0.0272)). However, WIF-1 down regulation was earlier reported in the nitrofen rodent lungs during the saccular stage of lung development (Fujiwara et al., 2012). WNT signaling plays an essential role in embryonic development. WIF1 is a target gene of SMAD1 in the developing lung epithelial cells and acts to inhibit WNT proteins. SMAD1 plays a key role in organogenesis including lung development and maturation, and SMAD1 knockout mice display reduced sacculation, which is an important feature of pulmonary hypoplasia. How this needs to be tied together with our rabbit data remains unclear.

In addition to that we observed changes that were different from what was previously described in clinical specimens. One example is that of the endothelin receptors EDNRA and EDNRB, which play a complex role in vascular tone. They are over-expressed in the thickened media of the pulmonary arteries of newborns with CDH (de Lagausie et al., 2005), as well as in the nitrofen rodent model (Dingemann et al., 2010). Binding of endothelin 1 to its receptor EDNRA on pulmonary artery smooth muscle cells results in vasoconstriction, whereas binding to the receptor EDNRB present on endothelial cells, results in vasodilation mediated by endogenous nitric oxide. In our rabbit model however EDNRB was not differently expressed in DH.

We also observed discrepancies between the effects of DH and TO within the rabbit model. For instance c-fos induced growth factor (FIGF, also known as vascular endothelial growth factor D, VEGFD) was down regulated in CDH as well as following TO. This is a member of the PDGF/VEGF family activating VEGFR2 and VEGFR3 receptors. As it was not differentially expressed in the controls we assume it was not involved. Conversely, VEGFR2 was upregulated in DH (FC 1.7819, FDR 0.0281) and down to normal levels after TO (FC -2.321, FDR 0.0121), which demonstrates the complex interactions of these factors.

We further had a number of discrepancies between RNA-Seq and qPCR, to which we pointed already when reporting on results for RFX3, LRRQL1, BMPR2 and PDE5A. Such discrepancies have been observed in previous reports (Duressa et al., 2013; Shi and He, 2014). This lack of concordance may be due to the poor knowledge of the rabbit genome and the absence of any SNP database which could result in a poor primer design for certain genes.

Additionally we acknowledge several other limitations to our study. The sample size is small yet it matches financial limitations and is in line with the sample size of other studies analyzing the gene expression changes by TO in DH. Another is that the study of gene expression in whole
lung samples will not detect changes occurring in individual cell types or different areas along the airways. These add to the generic limitations of translational studies in animal disease models, of which the clinical relevance remains uncertain. Another limitation of the study is that the surgical model does not fully recapitulate the etiology of the disease.

In conclusion we first describe the pattern of dysregulated pulmonary gene expression in lungs of fetal rabbits with surgically induced DH. Interestingly, tracheal occlusion, which is currently investigated as a prenatal surgical method to reverse pulmonary hypoplasia, is associated with a gene expression pattern that is comparable to what was observed in littermates with normal lung development.
Material and Methods

Surgery

The animal experiments were approved by the Ethics Committee of the Faculty of Medicine of the KU Leuven and all animals were treated according to current guidelines on animal wellbeing. In total 12 does of New Zealand white rabbits (Oryctolagus cuniculus) were operated and premedicated after weight estimation with ketamine 50 mg/kg (Ketamine 1000 CEVA®, CEVA Santé Animale, Brussels, Belgium), xylazine 6 mg/kg (Vexylan®; CEVA Santé Animale), and buprenorphine 0.03 mg/kg (Vetergesic®; Reckitt Benckiser Healthcare, Brussels, Belgium), all injected intramuscularly. General anesthesia was maintained using isoflurane 1.5% (Isoba® Vet; Abbott Laboratories Ltd., Queensborough, Kent, UK) in oxygen at 1 L/minute via a facemask. Maternal heart rate and oxygen saturation were monitored with a pulse oxymeter (Nellcor® N-20P; Nellcor Inc., Haasrode, Belgium). Physiologic body temperature was maintained by a heating pad. The doe was placed in the supine position, the abdominal wall was shaved and disinfected with povidone iodine (Isobetadine®; Asta Medica, Brussels, Belgium) and covered with sterile drapes. Aseptic conditions were maintained throughout all surgical procedures. We operated on two fetuses in each doe, located in the middle of the left or right uterine horn. First diaphragmatic hernia (DH) was created at 23 days gestational age (GA). A second operation was performed at 28 days GA on the previously operated fetuses, either tracheal occlusion (DH+TO group), or sham neck dissection and skin closure (DH group). The latter is performed to include the effects of the fetal surgery per se. Before and after surgery, animals were housed in separate cages at normal room temperature and daylight, with free access to food and water. At 30 days GA, the does were euthanized with an intravenous bolus of a mixture of embutramide 200 mg, mebezonium 50 mg, and tetracaine hydrochloride 5 mg (0.3 mL/kg T61; Marion Roussel Hoechst, Brussels, Belgium) after previous premedication as described above. All operated fetuses as well as one control unoperated littermate (control group) were euthanized in utero and harvested by cesarean section to obtain non-ventilated lungs.

Harvesting and sample preparation

During necropsy, the fetus and its lungs were weighed on a scale measuring accurately up to 0.001 g (HF 2000; A&D Instruments, Haasrode, Belgium) and the lung-to-body weight ratio (LBWR) was calculated.
Lung development is reflected by the lung-to-body weight ratio and pathologists have defined critical cut offs for pulmonary hypoplasia. We minimized contamination of the pulmonary tissue with other sources of DNA/RNA by flushing the lungs with saline by vascular access through the right ventricle of the fetal heart. The parenchymal lung tissue i.e. without the trachea was stored in RNAlater (Qiagen Benelux B.V., Venlo NL) at -4°C immediately.

**RNA isolation**

RNA isolation was performed on the entire left lung within 4 hours of harvesting using the RNeasy mini kit (Qiagen Benelux B.V., Venlo NL). Tissue lysis and homogenization was performed in 1200 µL Buffer RLT using the TissueLyser system (Qiagen Benelux B.V., Venlo, The Netherlands). Following tissue disruption and homogenization, samples were centrifuged for 3 minutes at 14,000 rpm in a benchtop microcentrifuge. Lysate was transferred to fresh tubes and an equal volume of 70 % ethanol was added. 600 µL of sample was added twice to a spin column, with 2 RNeasy spin columns used per sample. Following wash steps RNA was eluted in 50 µL RNAse-free H2O.

Total RNA quantification was performed using the Nanodrop 1000 spectrophotometer (Thermo Scientific, Aalst, Belgium). RNA integrity was assessed using the RNA 6000 Nano Kit and the Bioanalyser (Agilent Technologies, Diegem, Belgium) according to the manufacturer’s recommendations.

**RNA sequencing Library Preparation**

One µg of total RNA was used as input material for sequencing library preparation, which was performed with the TruSeq RNA Library Preparation Kit (Illumina, Eindhoven, NL) according to the manufacturers protocol. Fragmentation was performed for 6 minutes. 8 PCR cycles were used for the PCR enrichment step. Samples were indexed to allow for multiplexing. Sequencing libraries were quantified using the Qubit fluorometer (Life Technologies Europe BV, Gent, Belgium). Library quality and size range was assessed using the Bioanalyser (Agilent Technologies) with the DNA 1000 Kit (Agilent Technologies) according to the manufacturers recommendations.
RNA sequencing

Individual libraries were diluted to a final concentration of 2nM and pooled for sequencing. Pooled libraries were sequenced in a single lane of an Illumina HiSeq2000 flow cell generating single end 50bp reads. At least 10 million reads were obtained for all samples (range 12-32 million reads).

Data Analysis

Fastq files were mapped to the rabbit genome and transcriptome (OryCun2.0 retrieved from Ensembl, release 69) with the Array Studio software (OmicSoft, Cary, NC, USA) using default parameters for single short read data. Between 72% and 78% of reads mapped uniquely to the reference genome and/or transcriptome, and a further 6% to 8% of reads mapped non-uniquely. Between 16% and 18% of reads were unmapped. Expression values were calculated per gene; normalized to ‘reads per kilobase per million reads’ (RPKM) values as previously described (Mortazavi et al., 2008). Data was filtered to remove those genes with expression of <1RPKM in all samples. The RPKM expression values for the remaining 11,267 genes were subsequently Log2 transformed for downstream principal component analysis (PCA), hierarchical clustering, and for the subsequent calculation of fold changes by statistical inference analysis. All data are made available via ArrayExpress (accession number: E-MTAB-3452).

Hierarchical clustering

Hierarchical clustering was performed using the Array Studio software application, selecting ‘Complete’ for link option, and ‘Correlation’ for distance option. The ‘Correlation’ setting is also named "Centered Pearson" 1 - corr(x , y), and the source algorithms are further described in http://cran.r-project.org/web/packages/amap/amap.pdf. Clustering is based upon the Log transformed RPKM gene expression values.

Downstream Analysis

Downstream pathway analysis and creation of networks was performed using the Ingenuity Pathway Analysis (IPA) software application (Ingenuity Systems, Redwood City, CA, USA).
The iregulon plugin was used with the Cytoscape analysis software (http://www.cytoscape.org/) for prediction of upstream transcription factors of interest (Janky et al., 2014). The short time-series expression miner (STEM) application (Ernst and Bar-Joseph, 2006) was used to identify gene clusters with similar changes in expression between control and CDH, and following TO.

Real Time Quantitative PCR (qPCR)

In order to confirm RNA sequencing findings, we selected for qPCR (quantitative polymerase chain reaction) 12 genes with FC> 2 or FC< (-2) and the house-keeping genes SDHA, TOP1 and GAPDH that were previously shown to be among the most stable genes in CDH rabbit model and after TO (Vuckovic et al., 2013). PCR primers were designed using Primer 3 software and Primer-BLAST. The absence of secondary structure was checked by the Vector NTI program. All the primers were synthesized by Integrated DNA Technologies.

For a given gene, all samples were analyzed in the same qPCR run. Ten µL reactions were prepared using 2× LightCycler 480 SYBR Green I Master (Roche Applied Science) with 2 µL of a 1:10 cDNA dilution and a final concentration of 300 nM of each primer. Data were analysed using the LightCycler 480 Software (Roche Applied Science). The results were quantified using the comparative threshold cycle method as described by Livak and Schmittgen (Livak and Schmittgen, 2001). In addition, serial dilutions were used to create standard curves for relative quantification and the expression of each gene was normalized to at least two of the three house-keeping genes’ expression. The sequences and the different primers and the corresponding PCR efficiencies are provided in supplementary table 3.
Acknowledgments


Competing interests:

No potential conflicts of interest are to be disclosed.

Author contributions:

ACE, PB, JT, JRV and JD did the study design and manuscript writing. ACE, PDK, ME and JT performed the animal experiments. Sequencing and analysis was done by PB, ACE, MK, JFF, JRV and JD.

Funding:

JD receives a fundamental clinical research grant of the Fonds Wetenschappelijk Onderzoek Vlaanderen (FWO; 18.01207). ACE is supported by the Erasmus + Program of the European Union (Framework agreement number 2013-0040; contract 1011990). This publication reflects the views only of the authors, and the European Commission cannot be held responsible for any use, which may be made of the information contained therein.
References


**Figures**

**Figure 1:** Heatmap generated by unsupervised hierarchical clustering of individual sample level data (Log2 transformed RPKM expression values per gene) for the 641 unique genes found to be dysregulated in any group comparison, demonstrating that the different groups (DH, WT, and TO) cluster together. Samples are clustered on the horizontal axis and genes on the vertical axis. Gene expression values are shown as red for high expression and green for low expression, intensity reflects the level. This reveals three large gene clusters, labelled as 1, 2, 3. Information on Gene names, Fold Changes and FDR values for the groups comparisons, and Gene Cluster and STEM profile number are provided in Supplementary Table 1.
Figure 2: Shows the 4 significant gene profiles identified using the STEM application. Three timepoints are visible on the plots, representing changes in expression of the respective gene clusters from controls to DH, and DH treated with TO, respectively. Profile number is displayed in the top left, p-value bottom left.
## Tables

**Table 1:** In this table the three different groups (control, DH, DH+TO) are shown, containing fetal bodyweight, total fetal lung weight, weight of the left lung (LL), weight of the right lung (RL) and LBWR after harvesting at 30 days' GA. Excluded samples are specified.

<table>
<thead>
<tr>
<th>DH/DH+TO/Contr</th>
<th>excluded</th>
<th>Total lung weight</th>
<th>LL weight</th>
<th>RL weight</th>
<th>Bodyweight</th>
<th>LBWR</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH</td>
<td>yes</td>
<td>0.438</td>
<td>0.142</td>
<td>0.296</td>
<td>50.03</td>
<td>0.009</td>
</tr>
<tr>
<td>DH</td>
<td>yes</td>
<td>0.558</td>
<td>0.172</td>
<td>0.386</td>
<td>46.62</td>
<td>0.012</td>
</tr>
<tr>
<td>DH</td>
<td>yes</td>
<td>0.546</td>
<td>0.212</td>
<td>0.334</td>
<td>33.35</td>
<td>0.016</td>
</tr>
<tr>
<td>DH</td>
<td></td>
<td>0.506</td>
<td>0.121</td>
<td>0.385</td>
<td>36.14</td>
<td>0.014</td>
</tr>
<tr>
<td>DH</td>
<td></td>
<td>0.334</td>
<td>0.104</td>
<td>0.230</td>
<td>41.75</td>
<td>0.008</td>
</tr>
<tr>
<td>DH</td>
<td></td>
<td>0.399</td>
<td>0.128</td>
<td>0.271</td>
<td>36.27</td>
<td>0.011</td>
</tr>
<tr>
<td>DH</td>
<td></td>
<td>0.390</td>
<td>0.141</td>
<td>0.249</td>
<td>43.33</td>
<td>0.009</td>
</tr>
<tr>
<td>DH+TO</td>
<td>yes</td>
<td>0.376</td>
<td>0.093</td>
<td>0.226</td>
<td>29.16</td>
<td>0.013</td>
</tr>
<tr>
<td>DH+TO</td>
<td>yes</td>
<td>0.434</td>
<td>0.127</td>
<td>0.272</td>
<td>30.36</td>
<td>0.014</td>
</tr>
<tr>
<td>DH+TO</td>
<td>yes</td>
<td>0.310</td>
<td>0.105</td>
<td>0.205</td>
<td>29.94</td>
<td>0.010</td>
</tr>
<tr>
<td>DH+TO</td>
<td>yes</td>
<td>0.421</td>
<td>0.115</td>
<td>0.306</td>
<td>31.23</td>
<td>0.013</td>
</tr>
<tr>
<td>DH+TO</td>
<td></td>
<td>0.583</td>
<td>0.207</td>
<td>0.376</td>
<td>34.29</td>
<td>0.017</td>
</tr>
<tr>
<td>DH+TO</td>
<td></td>
<td>0.627</td>
<td>0.118</td>
<td>0.509</td>
<td>39.19</td>
<td>0.016</td>
</tr>
<tr>
<td>DH+TO</td>
<td></td>
<td>0.738</td>
<td>0.081</td>
<td>0.657</td>
<td>36.90</td>
<td>0.020</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>0.546</td>
<td>0.235</td>
<td>0.311</td>
<td>36.41</td>
<td>0.015</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>0.583</td>
<td>0.229</td>
<td>0.354</td>
<td>34.29</td>
<td>0.017</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>0.512</td>
<td>0.257</td>
<td>0.276</td>
<td>34.13</td>
<td>0.015</td>
</tr>
<tr>
<td>Control</td>
<td></td>
<td>0.459</td>
<td>0.223</td>
<td>0.236</td>
<td>32.79</td>
<td>0.014</td>
</tr>
</tbody>
</table>
Table 2: Table of marker genes for different lung cell types which are revealed to be significantly dysregulated in one or more comparisons (DH vs control; DH+TO vs DH; DH+TO vs control). Marker genes are defined from the top 30 known and novel putative marker genes reported by Truitlein et al.

<table>
<thead>
<tr>
<th>Markers of Lung Cell Types</th>
<th>DH vs Control</th>
<th>DH+TO vs DH</th>
<th>DH+TO vs Control</th>
<th>STEM Profile</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>FC</td>
<td>FDR</td>
<td>FC</td>
<td>FDR</td>
</tr>
<tr>
<td>AT2 Cell Markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD36</td>
<td>-1.574</td>
<td>0.0645</td>
<td>-1.6024</td>
<td>0.0651</td>
</tr>
<tr>
<td>CHI3L1</td>
<td>-2.6009</td>
<td>0.0656</td>
<td>1.3452</td>
<td>0.5505</td>
</tr>
<tr>
<td>DLK1</td>
<td>3.1566</td>
<td>0.0754</td>
<td>1.6257</td>
<td>0.4332</td>
</tr>
<tr>
<td>AT1 Cell Markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLIC5</td>
<td>1.4758</td>
<td>0.0638</td>
<td>-2.6291</td>
<td>0.0097</td>
</tr>
<tr>
<td>AHNAK</td>
<td>1.5615</td>
<td>0.386</td>
<td>-4.6513</td>
<td>0.0266</td>
</tr>
<tr>
<td>LIMCH1</td>
<td>1.3993</td>
<td>0.2078</td>
<td>-2.2892</td>
<td>0.0229</td>
</tr>
<tr>
<td>Clara Cell Markers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCGB1A1</td>
<td>1.2778</td>
<td>0.3533</td>
<td>-8.2686</td>
<td>0.0078</td>
</tr>
<tr>
<td>PPAP2B</td>
<td>1.6128</td>
<td>0.0436</td>
<td>-2.1248</td>
<td>0.0142</td>
</tr>
<tr>
<td>KDR (VEGFR2)</td>
<td>1.7819</td>
<td>0.0281</td>
<td>-2.321</td>
<td>0.0121</td>
</tr>
<tr>
<td>Ciliated Cell Markers</td>
<td>FOXJ1</td>
<td>CCDC19</td>
<td>SPAG17</td>
<td>CCDC39</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td></td>
<td>1.5702</td>
<td>1.4553</td>
<td>2.0341</td>
<td>2.0141</td>
</tr>
<tr>
<td></td>
<td>0.2284</td>
<td>0.2369</td>
<td>0.024</td>
<td>0.031</td>
</tr>
<tr>
<td></td>
<td>0.0968</td>
<td>0.0366</td>
<td>0.0138</td>
<td>0.0107</td>
</tr>
<tr>
<td></td>
<td>1.2808</td>
<td>1.5805</td>
<td>1.1792</td>
<td>1.5522</td>
</tr>
<tr>
<td></td>
<td>1.3025</td>
<td>0.6416</td>
<td>0.2945</td>
<td>0.6012</td>
</tr>
</tbody>
</table>
Translational Impact:

**Clinical issue:** Congenital diaphragmatic hernia (CDH) is a congenital malformation that can be nowadays treated prenatally by fetoscopic endoluminal tracheal occlusion (FETO). This treatment is currently under investigation in a multicenter randomized controlled trial. However the changes in gene expression induced by this treatment remain unknown. The rabbit model for CDH is well described and was used in this study for pulmonary transcriptome analysis after surgical creation of diaphragmatic hernia (DH) and consecutive fetal tracheal occlusion.

**Results:** In this study we saw that the gene expression profile of fetuses treated with tracheal occlusion after DH induction was more comparable to unaffected controls then to those with DH only. We also performed a pathway analysis and provide with this study a database that could be used to further improve the prenatal treatment of CDH.

**Implications and future directions:** Future prenatal treatments for CDH will have to be compared with FETO for their changes in pulmonary gene expression. There is an urgent need to decrease the invasiveness of fetoscopic procedures and to improve the efficiency of FETO. The herein provided database could be used to develop adjuvant therapies or transplacental therapies to fulfill that need.