Integrative analysis of the developing diaphragm provides a better understanding of congenital diaphragmatic hernia

by

Juan Felipe Garcia Rivas

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Department of Physiology
University of Alberta

Abstract

Introduction: Congenital Diaphragmatic Hernia (CDH) is a condition that affects 2.3 births every 10,000 and is characterized by an incomplete formation of the diaphragm. As a result of the incomplete development of the diaphragm, the abdominal organs will protrude into the thoracic cavity, impeding lung growth and causing lung hypoplasia. Due to the complexity of this condition, there are still gaps in the understanding of the etiology of CDH. Two main explanations are currently used to explain the pathogenesis of CDH: genetic and environmental. The retinoid hypothesis has tried to explain the etiology of this condition, and it states that abnormal retinoic acid signaling leads to the formation of diaphragmatic hernias. The goal of this thesis was to better understand the etiology of CDH, by trying to better understand both the genetic and environmental explanations. To fill the gaps in the knowledge in the etiology of CDH, we assessed (i) the levels of expression of multiple CDH and retinoid-associated genes in the pleuroperitoneal folds (PPF), and (ii) if blocking the retinoic acid signaling cascade in the developing diaphragm will cause CDH.

Methods: To answer the first questions, we performed single cell transcriptomics analysis in the developing diaphragm at gestational day (GD)13.5. Clustering analysis was performed utilizing Seurat in R. Expression analysis of curated retinoid and CDH-associated genes was performed in our scRNA-seq sample to determine the levels of expression of these genes in the different clusters present in the developing diaphragm. To test for the second hypothesis, we generated a conditional mutant mouse model expressing a truncated version of the retinoic acid receptors in the developing diaphragm. We also examined if *Dhrs3*$^{-/-}$ fetuses develop diaphragmatic hernias. The fetuses of our *Prrx1-Cre:Rar*^{dn} and *Dhrs3*$^{-/-}$ mouse models were collected and dissected via light microscopy to assess the incidence of diaphragmatic defects. The lung volumes of the

Prrx1-Cre:Rar^{dn} fetuses was also determine by utilizing contrast enhanced micro-Computing Tomography (micro-CT).

Results: Based on our single cell expression analysis, we were able to identify 18 different clusters and 10 distinct cell populations in the developing diaphragm. Expression analysis of our sample revealed that expression of both retinoid and CDH-associated genes is mostly seen in the mesenchymal and mesothelial components of the primordial diaphragm. A total of 215 CDH-associated genes and 30 retinoid-associated genes were analyzed. Our mutant mouse models gave us varying results. We found that the *Prrx1-Cre:Rar*^{dn} fetuses present CDH, with a penetrance of 100%. The herniations were seen mostly on the left side of the diaphragm, and they normally took more than 50% of that side. Our micro-CT analysis revealed that the *Prrx1-Cre:Rar*^{dn} fetuses have lower left lung volumes compared to control fetuses. This difference was not seen in the right lungs. In contrast, our *Dhrs3*$^{-/-}$ mouse model was not able to produce CDH in the fetuses.

Conclusion: To our knowledge, this is the first time scRNA-seq analysis has been performed in the developing diaphragm. We demonstrated that the non-muscular mesenchyme is the primary hotspot for expression of both CDH- and retinoid-associated genes. The retinoid hypothesis was also validated with our conditional animal model, as blocking the retinoic acid signaling cascade utilizing our *Prrx1-Cre:Rar*^{dn} model led to the formation of CDH in these fetuses.

Preface

This thesis is an original work by Juan F. Garcia Rivas. No part of this thesis has been previously published.

Some of the research conducted for this thesis forms part of a research collaboration. Chapter 2 has been submitted for publication to Pediatric Research. The manuscript ID is PR-2023-0366. For chapter 3, library preparation and cell sorting were performed by Dr. Michael Wong at the High Content Analysis Core at the University of Alberta. For chapter 4, micro-CT analysis was performed in collaboration with Dr. Michael Doschak and Dr. Adibehalsadat Ghazanfari from the department of pharmacology at the University of Alberta.

Acknowledgments

This thesis could not have been possible without the support of many people. I am extremely grateful to my supervisor, Dr. Robin Clugston, for his continued supervision, support and advice. I am also grateful to my committee members, Dr. Silvia Pagliardini and Dr. Meghan Riddell, for their expertise and advice during this time. Special thanks to my parents and brother for letting me follow my dreams, and for being my rocks during this process. I would also like to thank all my friends for being my family away from home. Finally, I would like to thank all the animals that were sacrificed for the purpose of this research.

Table of Contents

List of Tables

List of Figures

Abbreviations

CRABP – Cellular Retinoic Acid Binding Protein

CDH – Congenital Diaphragmatic Hernia

CRL – Crown-Rump Length

GD – Gestational Day

Ht – Heart

L - Left

LRAT – Lecithin Retinol Acyltransferase

Li – Liver

Lu – Lungs

Micro-CT – contrast enhanced Micro-Computer Tomography

PBS – Phosphate Buffered Saline

PPF – Pleuroperitoneal Folds

R - Right

R/ALDH – Aldehyde/Retinaldehyde Dehydrogenase

RA – Retinoic Acid

RAR – Retinoic Acid Receptors

Rar^{dn} - RAR dominant negative mutant

RARE – Retinoic Acid Response Element

RBP – Retinol Binding Protein

RDH – Retinol Dehydrogenase

RE – Retinyl Esters

REH – Retinyl Esters Hydrolase

RRD/DHRS – Retinol Dehydrogenase/Reductase

RXR – Retinoid X Receptors

UMI – Unique Molecular Identifiers

Chapter 1: General Introduction

1.1 Congenital Diaphragmatic Hernia (CDH)

Congenital Diaphragmatic Hernia (CDH) is a life-threatening birth defect with high mortality that affects 2-3 in 10,000 newborns (1–3). CDH is characterized by an incomplete formation of the diaphragm and associated lung hypoplasia, which contributes to severe neonatal respiratory distress (4,5). In high-resource settings, at least 1 in 5 newborns with CDH will not survive the neonatal period (~20-30% mortality), with outcomes significantly worse in low-resource settings (>90% mortality) (6–8). While it is rare, there is a considerable cost incurred when providing newborn CDH patients with intensive care, requiring significant healthcare resources (3,9–11). This burden on the healthcare system and patient families is compounded when one considers that over half of surviving CDH patients have complex long-term morbidity (12–15). To fully grasp the clinical picture of CDH's impact, the reader is referred to the recent review by Zani and colleagues, which provides a comprehensive overview of CDH from a clinical perspective (5).

The etiology of CDH is as complex and remains largely unknown (5,16). There are different aspects of CDH to consider here, including clinical presentation and type of diaphragm defect. For example, CDH can present as an isolated defect (~50-60% of cases) or complex CDH, where additional abnormalities are present (1,17,18). Moreover, CDH can be thought of as an umbrella term incorporating different types of diaphragm defect (19,20). This includes the most commonly occurring Bochdalek CDH that is characterized by a hole in the posterolateral diaphragm (70% of cases), as well as anterior/parasternal holes (Morgagni hernias, ~10-27% of cases), diaphragm eventrations (~2-3% of cases), and central tendon defects (~1-2% of cases) (19,21,22). As discussed below, it is important to consider that etiologic factors may vary between these different presentations of CDH. It is generally considered that the two main factors that contribute to this birth defect are genetic and/or environmental. Approximately 30-

40% of CDH cases have an identifiable genetic cause (3,23). Of this number, chromosomal defects account for about 10% of CDH cases, and *de novo* mutations account for 10-22% of cases (24–26). Environmental risk factors for CDH are varied, and include associations with maternal age, alcohol use, and smoking (3,16,27–30). As the major focus of this review, one of the leading hypotheses to explain the etiology of CDH is the so-called Retinoid Hypothesis. This hypothesis was crystallized in the landmark paper published by Greer, Babiuk, and Thebaud in 2003, stating that "abnormalities linked with the retinoid signaling pathway early in gestation may contribute to the etiology of CDH" (31). The Retinoid Hypothesis is centered on the importance of the signaling molecule retinoic acid as a potent regulator of mammalian development (32–34).

1.2 Diaphragm Development and CDH

To better understand the etiology of CDH, it is necessary to know how the development of the diaphragm works. The first stage in diaphragm development is the formation of the septum transversum, a thin membrane that lays on top of the liver and functions as a transient structure that divides the pleural and peritoneal cavities (35). Due to the lack of markers, this structure has not been well studied, and therefore its role in diaphragm development and CDH is poorly understood (36,37). The second structure are the Pleuroperitoneal Folds (PPF). This pair of transient structures are pyramidal in shape and are located between the pleural and peritoneal cavities. This tissue can be seen during weeks 4-6 in human gestation, gestational day (GD) 14-16 in rats, and GD 11.5-14.5 in mice (35,36). Both myogenic cells as well as phrenic axons migrate into the PPF, and later proliferate to form the fully muscularized diaphragm by GD 16.5 in mice and GD 17 in rats. It is important to note that the expansion of the PPF gives rise to two distinct domains. The first one, which occurs laterally in both of the PPFs gives rise to the

muscular tissue of the diaphragm. The central domain, where both PPFs meet, is non-muscularized, and gives rise to the central tendon (36). Muscle precursor cells arise from the somites, beginning at GD 9 in mice. By GD 11.5, the muscle precursor cells have fully migrated into the PPFs (37). As the PPF expands, it carries the myogenic cells, regulating muscle morphogenesis. A diagram of the location of the PPFs is presented in Figure 1.1. Genetic studies utilizing markers of the PPF, specifically *Gata4*, have shown that this tissue gives rise to both the central tendon as well as non muscular fibroblasts in the adult diaphragm (38,39). It is also known that defects in the PPF can give rise to malformations in the mature diaphragm, the most common one being CDH. For example, studies examining muscle proliferation demonstrated that in nitrofen-exposed mice there is no disturbances in myogenesis, suggesting that the non-muscular mesenchyme of the PPF is the source of CDH (40). Furthermore, teratogen-exposed *c-met* null mutants still develop CDH, even in the absence of muscular tissue (41). Mutant mice models for multiple CDH-associated genes have demonstrated that it is mutations in genes expressed in the mesenchymal and mesothelial components of the PPF that give rise to diaphragmatic defects (38,42,43). These studies also demonstrated that defects can be seen prior to full muscularization of the diaphragm, at GD12.5, further linking defects in the PPF as the main culprit in the formation of CDH. The reader is instructed to read the following literature in order to better understand diaphragm development during embryogenesis (35,37).

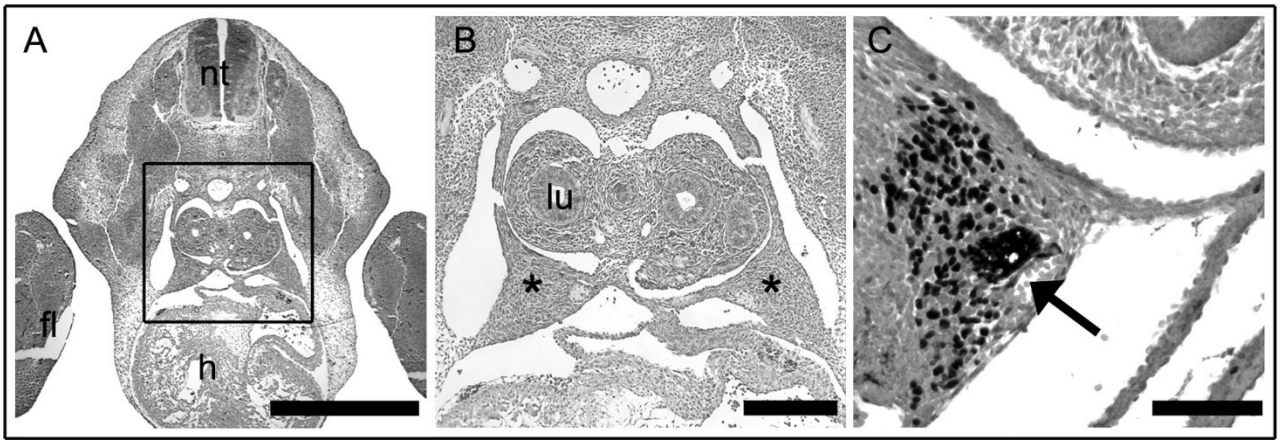


Figure 1.1: Diaphragm development is controlled by the PPF. Location of the PPFs during embryogenesis can be seen in A. Higher focus is shown in B, with both PPFs shown with an asterisk. Finally, C shows the location of the phrenic nerve with an arrow, as well as muscle precursor cells in black via immunological staining of *Pax3.* This image was adapted from Greer & Clugston, 2007 (35).

1.3 Vitamin A Metabolism

To better understand the retinoid hypothesis, a basic understanding of the vitamin A signaling pathway is needed. Vitamin A and other pro-retinoid compounds are obtained through the diet and can be stored in the form or retinyl esters (RE) via Lecithin Retinol Acyltransferase (LRAT) in the liver (44). When retinol is needed REs can be broken down via Retinyl Ester Hydrolase (REH), and retinol itself can be transported in the bloodstream while being bound to retinol binding protein 4 (RBP4) (45). RBP4-bound retinol can be taken into the cell via the transmembrane transporter STRA6 (46). Once inside the cell, retinol can be metabolized into retinaldehyde, a reaction that is catalyzed by members of the retinol dehydrogenase (RDH) family (47). This reaction is reversible, as retinaldehyde can be reduced into retinol by members of the retinol dehydrogenase/reductase (RRD/DHRS) family. Retinaldehyde can be further oxidized into retinoic acid (RA), which is the active component in the signaling cascade (47). This reaction is the rate-limiting step and is performed by enzymes of the aldehyde/retinaldehyde dehydrogenases family (R/ALDH). It is important to mention that inside the cell, retinol and its derivatives are also bound to different proteins. Retinol is bound to RBP1 (48), whereas RA is bound to members of the cellular retinoic acid binding protein (CRABP) family (49).

Different forms of RA send different signaling cues by acting on different receptors. All-*trans* RA acts as a ligand for the retinoic acid receptors (RARs). These receptors can homodimerize or heterodimerize with each other or with another group of receptor isoforms, retinoid X receptors (RXRs) (50). The RXRs are nonspecific and have multiple ligands, including 9-*cis* retinoic acid (51), although it is unsure if this compound occurs endogenously (52). Both families of receptors work in a similar fashion, by binding RA and dimerizing on DNA regulatory sites known as retinoic acid response elements (RAREs), leading to modulation of gene expression (53). A

schematic of basic vitamin A metabolism can be seen in Figure 1.2. RA itself regulates the gene expression of multiple different genes that are involved in a variety of pathways like limb formation, lung and heart organogenesis, among others (54). RA is an active metabolite of dietary vitamin A, and while its complex biology is beyond the scope of the current review, the reader is directed to the following reviews that comprehensively describe vitamin A metabolism and retinoic acid signaling (47,55,56). For the purpose of this thesis, whenever RA is mentioned, we are referring to all-*trans* RA.

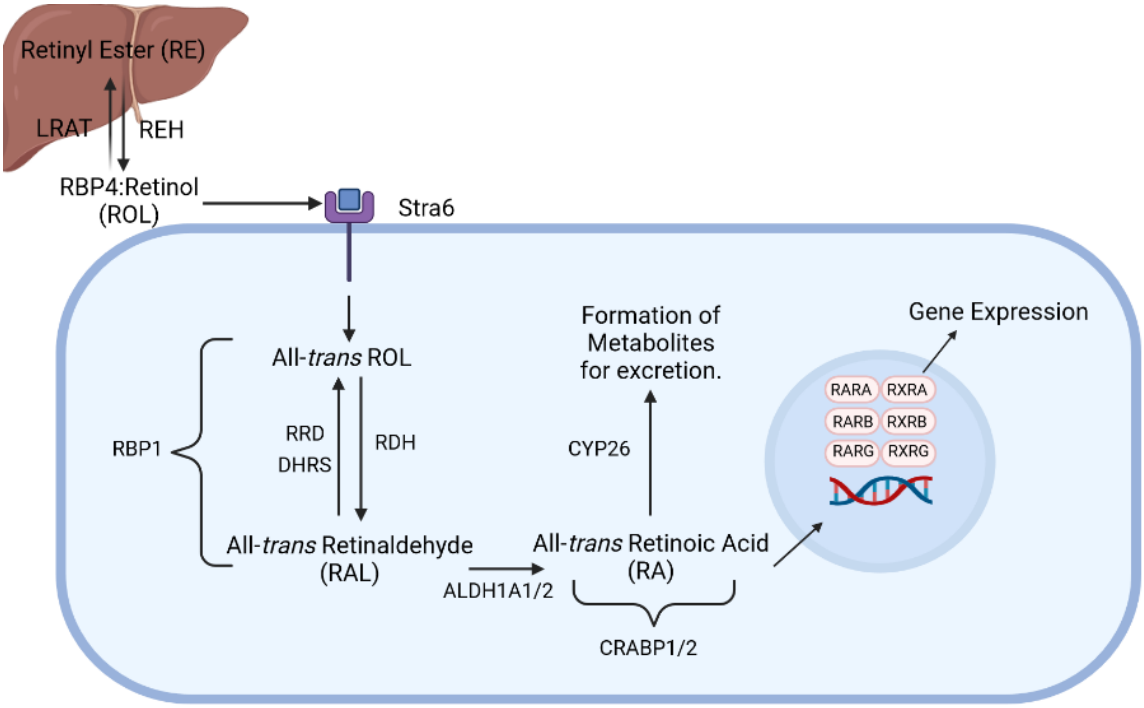


Figure 1.2 Basic Vitamin A metabolism and signaling pathway. Retinols are stored in the liver in the form of RE. When needed, RE is broken down into retinol and transported in the bloodstream bound to RBP4. Retinol-bound RBP4 can enter the cell via Stra6, where it goes through a series of oxidation reactions to generate RA. RA can bind the RARs to modulate gene expression. Created with BioRender.com.

1.4 Hypotheses and Goal of this Thesis

The overall goal of this project was to better understand the etiology of CDH by performing integrative analysis on the developing diaphragm. To achieve this, we were interested in proving the following two hypotheses: (i) the non-muscular mesenchyme of the PPF will have the highest level of expression of CDH-associated genes and retinoid-related genes and (ii) that blocking the RA signaling cascade in the developing diaphragm will cause CDH. Both hypotheses aim to strengthen the theory that the PPF is the main tissue involved in the development of diaphragmatic hernias, while also providing more evidence to support the retinoid hypothesis of CDH.

1.5 Experimental Plan

1.5.1 Hypothesis 1 – Gene expression in the developing diaphragm

The first hypothesis was generated since there are previous studies that have linked the non-muscular mesenchyme of the PPF with the development of CDH. Previously, it was shown that CDH can occur even in the absence of muscle precursor cells, and independent of myogenesis (41). Furthermore, animal models that target genes in the PPF known to cause CDH when mutated have shown the development of diaphragmatic hernias (38,57). Based on these studies, we suggest that the non-muscular mesenchyme of the developing diaphragm will have the highest expression of retinoid and CDH-associated genes. To test this, we performed scRNA-seq analysis on a PPF sample from a mouse fetus, and we determined the distinct populations of cells present in this sample. Expression analysis was then performed to assess the levels of expression of multiple genes in the different cell populations of the PPF.

1.5.2 Hypothesis 2 – RA signaling in the mesenchyme of the developing diaphragm

The second hypothesis mostly seeks to further support the retinoid hypothesis, by testing if blocking the RA signaling cascade in the mesenchymal component of the diaphragm will cause CDH. As stated above, we were interested in the non-muscular mesenchyme since it has been heavily implicated with the formation of diaphragmatic defects. To test for this, we generated a conditional mutant mouse model utilizing a *Prrx1-Cre* driver that has been previously described to target the mesenchyme of the developing diaphragm (57). To block RA signaling, we used a well-characterized RAR dominant negative mutant (Rar^{dn}), which is known to block RA signaling by disrupting the function of all three RAR (58). By mating both of these mutant mice, we generated *Prrx1-Cre:*Rar^{dn} offspring, and the incidence of CDH in these fetuses was analyzed via dissection. To determine the degree of lung hypoplasia, contrast enhanced Micro-Computer Tomography (micro-CT) was performed.

1.6 Rationale and Significance

Although CDH is a condition that has been highly studied, there are still many gaps in the knowledge of the etiology of this disease. This thesis hopes to close some of said knowledge gaps, by exploring the role of the non-muscular mesenchyme of the developing diaphragm. This thesis will also provide valuable new knowledge in the field, as (i) it is the first time that scRNA-seq has been performed in the PPF, (ii) we generated a brand-new mouse model that presents 100% penetrance of CDH and (iii) we added more evidence in support of the retinoid hypothesis. Overall, the data generated in this thesis not only provides a better general knowledge about the developing diaphragm, but also highlights the importance that retinoids have in the formation of diaphragmatic hernias, which can have translational potential for improving CDH incidence.

Chapter 2: Literature Review

2.1 History of the Retinoid Hypothesis

The Retinoid Hypothesis was based on research dating back to the 1940s (59), and drew together several threads of evidence to justify the notion that altered retinoid signaling may cause CDH (31). The foundation for the Retinoid Hypothesis included lines of evidence from animal models of CDH, as well as emerging clinical data. In brief, these studies included: 1) the observation that the offspring of vitamin A deficient rats had a high incidence of diaphragmatic defects, and when vitamin A was re-introduced into the diet, the rate of herniation decreased (60–62). 2) Compound Retinoic acid receptor (*Rar*) knock-out mice have a low incidence of CDH (63,64). 3) Nitrofen, a teratogen widely used to study CDH in rodents, inhibits the retinoic acid synthesizing enzyme RALDH2 (also called ALDH1A2) *in vitro* (65), and suppresses activation of a *Rar* reporter gene (*RARE-LacZ*) *in vivo* (65,66). 4) Nitrofen-induced CDH in rats can be reduced by the co-administration of large doses of vitamin A (67). And 5) in a small cohort of infants with CDH (n = 11), circulating levels of retinol and its carrier protein in the blood (RBP) were decreased by ~50% in comparison with controls (22). The balance of this evidence led the authors to conclude that altered retinoid signaling may be a contributing factor in the etiology of CDH. As reviewed below, the formulation of the Retinoid Hypothesis had a significant impact on the field and its direction of research.

2.2 New insights from animal models of CDH

Animal models of CDH continue to be an important tool in improving our understanding of how CDH develops and its causes. As discussed below, recent discoveries have helped to define the retinoic acid signaling pathway in the developing diaphragm, as well as provide important new insight into the role of vitamin A and its derivatives in the development of CDH gleaned from teratogenic and genetic models.

2.2.1 Defining the retinoic acid signaling pathway in the developing diaphragm

As the Retinoid Hypothesis has developed it has become increasingly important to define the retinoic acid signaling pathway in the developing diaphragm, including identifying which specific components of the pathway are expressed in the developing diaphragm, and in which specific cells. Given our knowledge of CDH pathogenesis, this work has primarily focused on the pleuroperitoneal folds (PPFs); key structures in the developing diaphragm that are known to be disrupted in animal models of CDH (35,40). Immunohistochemical and transcriptomic studies have helped to define the retinoic acid signaling pathway in the PPF (68,69) ,with additional insight gained from other studies (42). Figure 2.1 describes our current understanding of the retinoic acid signaling pathway in the PPF, although we recognize that this pathway does not discriminate between different cellular subtypes within this structure, an important knowledge gap in the field given the suspected importance of specific cell types in the cellular origins of CDH, a topic beyond the scope of the current review (35,70).

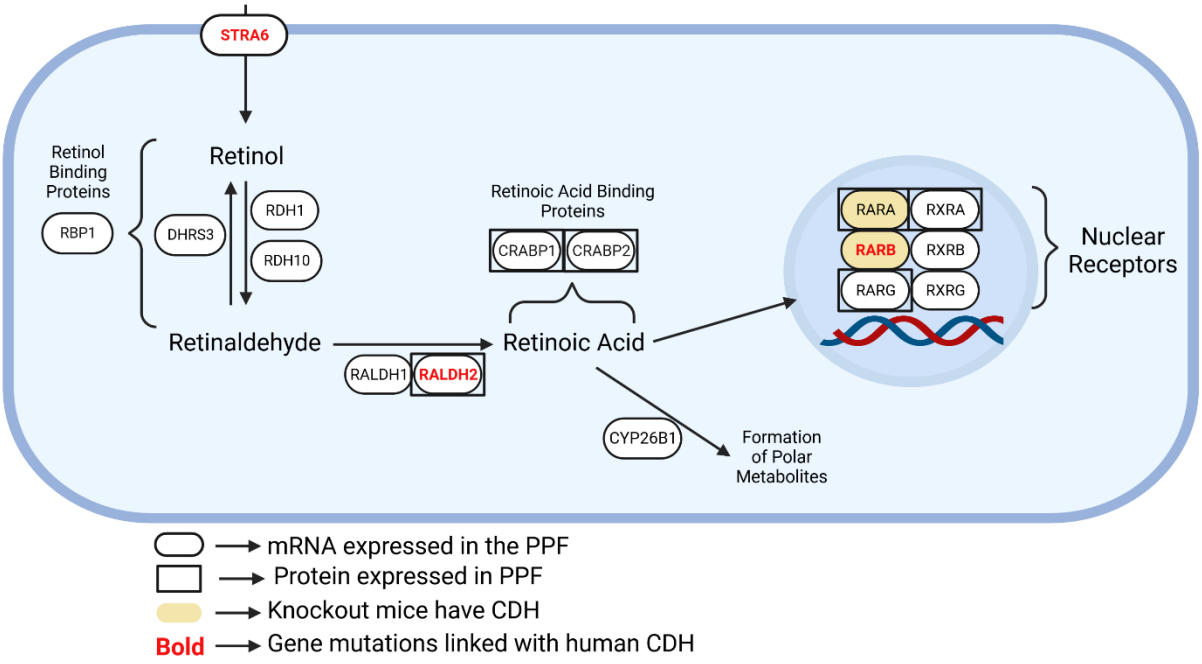


Figure 2.1 Retinoid metabolism and signaling in the developing diaphragm. This schematic represents our current knowledge regarding retinoic acid metabolism and signaling in the pleuroperitoneal fold (PPF), a key structure in the developing diaphragm. All shown genes are known to be expressed in the PPF based on transcriptomic analyses. The schematic also highlights cases where protein expression in the PPF has been assessed, as well as instances where mouse or human gene mutations are associated with abnormal diaphragm development and CDH. Created with BioRender.com.

2.2.2 Teratogenic models of CDH

Key studies in the original formulation of the Retinoid Hypothesis were based on data generated using the nitrofen model of CDH. More recent work has further explored how nitrofen produces CDH through the disruption of retinoic acid signaling, and how exogenous retinoids can rescue teratogen-induced CDH. The effect of teratogen-exposure on the expression level of key components of the retinoic acid signaling pathway has been described, and direct targeting of RAR signaling has strengthened the Retinoid Hypothesis.

2.2.2.1 Mechanisms of teratogen-induced CDH

Nitrofen is a well-established model to induce CDH in rodents (71–73). The Retinoid Hypothesis was founded on reports that nitrofen can inhibit RALDH2 and suppress activation of RAR signaling (65,66). Noble and colleagues conducted a series of mechanistic studies to better understand how nitrofen disrupted retinoid signaling to induce CDH (74). This work provided further evidence that nitrofen inhibits RALDH2 using a *RARE-LacZ* reporter system in P19 cells, which could be overcome by adding excess enzyme substrate in the form of retinal (74). It was subsequently shown in rats, that nitrofen treatment on gestational day 9 significantly decreased fetal retinoic acid levels between gestational day 11 and 13, which coincides with a critical time in diaphragm development and provided unequivocal evidence that nitrofen treatment could inhibit retinoic acid synthesis *in vivo* (74). This decrease in retinoic acid synthesis complemented the previously observed decrease in activation of the *RARE-LacZ* reported gene in nitrofen-treated mice (66). Indeed, nitrofen and three other CDH-inducing teratogens that inhibit RALDH2 all decrease embryonic retinoid acid signaling throughout the critical period in diaphragm development (gestational day 12-14), as subsequently shown in *RARE-LacZ* mice

(68). While alternative mechanisms have been suggested (75,76), the evidence suggests that nitrofen primarily induces CDH by inhibiting RALDH2, thereby decreasing retinoic acid synthesis and inhibiting downstream signaling through the RARs.

2.2.2.2 Exogenous retinoids can rescue teratogen-induced CDH

Powerful evidence to support nitrofen's mechanism of action and the importance of retinoid signaling in diaphragm development came from the observation that administration of vitamin A could lower the incidence of nitrofen-induced CDH in rats (67). Using a dose of 15,000 IU of vitamin A, Thebaud *et al.* decreased the incidence of nitrofen-induced CDH by over half (67). Using the same dose of vitamin A, other groups confirmed that co-administration of vitamin A could lower the incidence of nitrofen induced CDH in rats, with these studies also shifting their focus to potential beneficial effects of vitamin A on the developing lung, as discussed below (77,78). Based on the model that nitrofen inhibits retinoic acid synthesis, subsequent studies tested the hypothesis that administration of retinoic acid would be more efficacious than vitamin A (retinol) administration in preventing teratogen-induced CDH (79). These studies showed that while vitamin A administration (25,000 IU vitamin A) reduced the incidence of CDH by ~40%, retinoic acid had a more potent rescue effect, lowering the incidence of CDH by ~80% (79). The potent effect of direct supplementation with retinoic acid to prevent CDH was further emphasized in teratogen-treated mice (68). Here, a high incidence of CDH (~60% CDH) was almost completely abrogated by co-administration with retinoic acid (~1% CDH), with the authors further showing that while teratogen treatment decreased embryonic *RARE-lacZ* expression, RAR signaling was restored by retinoic acid treatment in accord with the decrease in CDH incidence (68).

A limitation of the rescue studies discussed above was the high dose of vitamin A (e.g., 15,000-25,000 IU vitamin A) required to reverse the effects of nitrofen. Interestingly, based on the observation that low maternal dietary vitamin A intake is a risk factor for developing CDH in humans (80,81) , it has been shown that altered dietary vitamin A intake can modulate the incidence of teratogen-induced CDH, such that mice consuming a vitamin A deficient diet have a higher incidence of CDH (82). This study is important because it shows that high, pharmacological doses of vitamin A are not required to prevent teratogen-induced CDH, rather, it can be achieved through increased dietary vitamin A intake. As discussed below, this is important in the context of low maternal dietary vitamin A intake in humans as a potential risk factor for CDH.

Taken together, several studies have shown the efficacy of high doses of vitamin A to prevent teratogen-induced CDH, with more recent work suggesting that adequate dietary vitamin A intake can achieve a similar result. These studies highlight the importance of retinoic acid signaling in the development of teratogen-induced CDH and the feasibility of preventing CDH, at least in animal models.

2.2.2.3 Directly targeting retinoic acid receptor signaling induces CDH

As discussed above, there is strong evidence to suggest that nitrofen induces CDH by inhibiting retinoic acid synthesis via RALDH2. Other potential mechanisms of nitrofen's action have been suggested (74–76), representing a potential limitation in the interpretation of these studies. On the other hand, studies using the pan-RAR antagonist BMS493 provide direct evidence in support of the Retinoid Hypothesis, showing that specific inhibition of retinoic acid signaling is a potent inducer of Bochdalek CDH in rats (68). Indeed, a 5 mg/kg dose of BMS493

induced almost 100% herniation when given to pregnant dams on gestational day 11. BMS493 administration was further used to define the critical period of diaphragm development in rats, which spanned gestational day 8 to 13. Remarkably, separate critical periods were identified for the left and right hemi-diaphragms, with the left side of the diaphragm being preferentially affected with BM3493 administration between gestational days 8 and 9, the right side preferentially affected between gestational days 11 and 13, and bilateral hernias produced with administration on gestational day 10. Thus, direct targeting of RAR signaling on different gestational days induced CDH in a temporally and spatially restricted manner (68). While BMS493 was first used to induce CDH in rats, subsequent research has shown that this compound can also induce CDH in mice, albeit at a lower incidence (~6%) (83). Similarly, <1% of mice treated with the pan-RAR antagonist BMS-189453 also developed CDH (84).

2.2.3 Genetic models of CDH

The observation that *Rar* knock-out mice developed CDH was integral to the original formulation of the Retinoid Hypothesis (31). As discussed below, there has been a surprising lack of studies genetically dissecting the retinoic acid signaling pathway in the developing diaphragm; however, several other mouse models with an indirect relationship to retinoic acid signaling pathway have been described that can be interpreted as being supportive of the Retinoid Hypothesis.

2.2.3.1 Genetic models directly related to retinoic acid signaling

The only direct genetic evidence supporting the Retinoid Hypothesis obtained from mice with targeted mutations in the retinoic acid signaling pathway come from studies of *Rar* compound mutant mice (Figure 2.1). There are three *Rars* encoded in the mammalian genome,

Rara, *Rarb*, and *Rarg* (32,53). Single knock outs for each *Rar* do not produce diaphragm defects; however, a low incidence of CDH has been reported in *Rara*$^{-/-}$*:Rarb*$^{-/-}$ compound mutant mice, which was not observed in *Rara*$^{-/-}$*:Rarg*$^{-/-}$ or *Rarb*$^{-/-}$*:Rarg*$^{-/-}$compound mutants (63,64). Interestingly, diaphragm defects were also reported in *Rxra* transgenic mice that lack their activation function domain (AF-2), highlighting the importance of this heterodimeric partner of the RARs (85). To our knowledge, there are no other published studies describing diaphragm defects in mice carrying mutations in genes directly associated with the retinoic acid signaling pathway. Indeed, many knock-out mice in this pathway are grossly normal and viable (e.g. *Lrat, Crbp1,* and *Rbp4*) (44,48,86). The lack of a diaphragm phenotype in these mice may reflect robustness/compensation in the retinoic acid signaling pathway, or evidence that they are not essential in diaphragm development. Conversely, some knock-out mice in this pathway are embryonic lethal (e.g. *Raldh2, Cyp26a1,* and *Dhrs3*) (87–89) precluding an assessment of the formation of the full muscularized diaphragm. In this latter case, experimental approaches using conditional deletion of specific genes could allow testing of their hypothetical importance in the developing diaphragm.

2.2.3.2 Genetic models indirectly related to retinoic acid signaling

In addition to *Rar* compound mutant mice, numerous other genetic models of CDH have been described in mice (90). As discussed below, this includes diaphragm defects in mice deficient for several genes that can be indirectly linked to retinoic acid signaling.

This link is perhaps most clear with *Wt1*, which encodes a transcription factor with a broad expression pattern, including the PPF (43,57,91). The original description of *Wt1*$^{-/-}$ mice emphasized its role in early urogenital development but also mentioned diaphragm defects that

were later found to be similar to Bochdalek CDH (92,93). Because of its broad importance in embryonic development (94), *Wt1*$^{-/-}$ mice typically do not survive into late gestation, limiting the ability to study diaphragm development (93). This limitation has been overcome by conditional deletion of *Wt1*. Carmona *et al.*, used an enhancer of *Gata4* to conditionally delete *Wt1* in the lateral plate mesoderm, leading to a ~80% incidence of Bochdalek CDH (42). Similarly, Cleal *et al.*, conditionally deleted *Wt1* in the PPF to generate offspring with an 80-90% penetrance of Bochdalek type diaphragm defects (57). Regarding retinoic acid signaling, WT1 regulates the expression of the retinoic acid synthesizing enzymes RALDH2 (95), and *Wt1* itself is thought to be a retinoic acid target gene (96). Most interestingly, dietary retinoic acid supplementation has been shown to relieve the incidence of CDH in mice with conditional *Wt1* deletion and decrease the size of diaphragmatic defects in embryos that still displayed CDH (42).

Like WT1, COUP transcription factor 2 (COUP-TF2, encoded by *Nr2f2*) is a developmentally important transcription factor with links to the retinoic acid signaling pathway (97,98). It has been shown that COUP-TF2 is upregulated by retinoic acid, and it can regulate gene expression by modulating the dimerization of RAR/RXR receptors (96,99–102). Global *Coup-tf2*$^{-/-}$ mice are embryonically lethal; however, conditional deletion of *Coup-tf2* in the PPF induces diaphragm defects that are similar to Bochdalek CDH in ~50% of offspring (103). Two other genes that encode transcription factors that functionally interact with each other and have been linked to retinoic acid signaling and CDH are *Gata4* and *Fog2* (also known as *Zfpm2*). GATA4 is involved in many different developmental processes, the most important one being proper cardiac development (104–106). *Gata4*$^{-/-}$ mice are embryonic lethal, but heterozygous mutants develop CDH in ~30% of offspring, which presents along the ventral midline of the diaphragm and impacts the central tendon (107). FOG2 is thought of as a co-factor of the GATA

transcription factors and has been shown to interact with GATA4 in heart development (108). Mice with a mutation in *Fog2* present diaphragm abnormalities characterized by incomplete muscularization of the posterolateral diaphragm in 100% of mutant mice (109). Regarding the Retinoid Hypothesis, both *Gata4* and *Fog2* are thought to be retinoic acid target genes (96,99,110,111), and it has been shown that retinoids can regulate downstream gene expression by interacting with GATA4 and FOG2 (112). Interestingly, FOG2 has also been shown to interact with COUP-TF2, suggesting complex interactions with multiple transcription factors that are important in diaphragm development (113).

The presence of diaphragm defects in *Wt1*, *Coup-tf2*, *Gata4*, and *Fog2* mutant mice indicates their individual importance in normal diaphragm development. Links between these genes and the retinoic acid signaling pathway provide supporting evidence for the Retinoid Hypothesis, as also recently discussed by others (114). Future studies integrating our understanding of how these genes interact should provide insight into the cellular pathogenesis and etiology of CDH.

2.3 New Insights from human cases of CDH

When Greer and colleagues originally formulated the Retinoid Hypothesis there was sparse evidence linking it with CDH in infants (31). This represented a limitation of the hypothesis; however, in the intervening 20 years new insights driven by advances in clinical genetics, as well as epidemiological studies have strengthened the relevance of the hypothesis to CDH in humans.

2.3.1 Retinoid status of infants with CDH

The only clinical data in support of the Retinoid Hypothesis when it was first formulated came from a small study of vitamin A status in 11 cases of CDH (22,31). This study found that newborns with CDH had a ~50% decrease in circulating retinol and retinol binding protein (RBP) concentrations compared to healthy controls (22). A follow-up study including 22 newborns with CDH confirmed this observation, reporting significantly lower retinol and RBP concentrations in newborns with CDH compared to matched controls; however, the magnitude of this decrease was relatively smaller (~20-25% decrease) (115). Importantly, neither study found a decrease in maternal markers of vitamin A status, suggesting that maternal vitamin A deficiency was not a contributing factor to CDH in these patient populations. Thus, these two studies suggest that the vitamin A status of newborns with CDH is somehow impaired, although the reason for this remains elusive and requires further study. To our knowledge, there has been no other comprehensive examination of vitamin A status in CDH. One autopsy study including nine cases of CDH attempted to indirectly measure retinoid stores in the liver and lung via CRBP1 immunoreactivity (116). These results were interpreted in the context of "retinoic acid status", which we have questioned has incorrect (117), nevertheless this study hints at the possibility of impaired vitamin A status of fetuses with CDH. A physiological decline in circulating maternal retinol levels has been reported in human pregnancy, with a rebound to non-pregnant levels after birth (118–120). A similar pattern is seen in pregnant mice and rats (121,122). This has led to the suggestion that these declines correspond with a critical period in diaphragm development and may render this structure more susceptible to aberrant retinoic acid signaling (31,114). As above, variations in circulating maternal retinol levels, fetal retinoid status, and their importance in CDH require further study.

2.3.2 The genetics of CDH and the Retinoid Hypothesis

As discussed above and reviewed by others at length, the genetic etiology of CDH is complex and continues to be an active area of research and discovery in the field (16,90,123). When the Retinoid Hypothesis was first proposed our understanding of the genetics of CDH was limited; however, since then there has been significant progress in our overall understanding of the genetics of CDH, as well as how this new knowledge relates to the Retinoid Hypothesis (114,124). Similar to the mouse genetics, two themes emerge from the human genetics of CDH as it relates to the Retinoid Hypothesis: a subset of genes associated with CDH are directly involved in retinoic acid signaling, while others can be indirectly linked to retinoic acid signaling either as target genes or factors that potentially interact with RAR signaling.

2.3.3 Human genetic disorders directly linking retinoid acid signaling and CDH

Goumy *et al.*, systematically discussed the retinoic acid signaling pathway and drew links of varying strength with several CDH-associated genes (124). The strongest evidence linking the retinoic acid signaling pathway and CDH in humans exists for *STRA6*, *RALDH2*, and *RARB* (Figure 2.1). *STRA6* (OMIM: 610745) encodes the membrane receptor for RBP, which acts as a transporter for the cellular uptake of retinol (125). *STRA6* mutations cause Matthew Wood syndrome, which includes microphthalmia, pulmonary hypoplasia, cardiac malformations, as well as diaphragm defects variously reported as CDH/diaphragm eventration in ~25% of cases (126–129). Diaphragm defects are frequently associated with *STRA6* mutations, and this gene is known to be expressed in the developing diaphragm, strongly linking *STRA6* with CDH (69,129). *ALDH1A2 (RALDH2*; OMIM: 603687) encodes a protein with the same name in the aldehyde dehydrogenase family, which synthesizes retinoic acid (130). Mutations in this gene

cause a variety of different phenotypes, including tetralogy of Fallot, absent thymus, kidney defects, as well as diaphragmatic eventration/Bochdalek hernia (131,132)**.** A total of four known patients with mutations in this gene have diaphragm defects, it is expressed in the developing diaphragm, and its inhibition causes CDH in rodents, strongly linking *RALDH2* with CDH (65,69,74,131–133)**.** *RARB* (OMIM: 180220) encodes for the β isoform of the RAR, a nuclear receptor activated by retinoic acid that regulates the expression of many different genes (134)**.** Mutations in this gene are associated with diaphragmatic hernias as well as eventrations (135–137). *RARB* mutations have been reported in 9 patients with CDH, its mutation is linked with CDH in mice, and it is expressed in the developing diaphragm, strongly linking it with CDH (135,136)**.** As discussed by Goumy *et al.* and others, *RBP1*, *RBP2*, *RBP5*, and *LRAT* have all been linked with CDH, although the evidence is relatively weak at this time (124,138). These links are primarily driven by the proximity of these genes to chromosomal regions recurrently deleted in individuals with CDH, but to our knowledge no specific gene mutations have been described in them directly linking them to CDH.

2.3.4 Human genetic disorders indirectly linking retinoid acid signaling and CDH

Genes that have been linked with CDH in humans that are indirectly involved in retinoic acid signaling are numerous and can be broadly categorized as genes whose protein products interact with RAR function or they are retinoic acid target genes.

Regarding genes encoding proteins that are thought to interact with RAR function (and as discussed above) there is significant evidence linking *Wt1*, *Coup-tf2*, *Gata4*, and *Fog2* to CDH in animal models. Interestingly, all these genes are also associated with CDH in humans. Mutations in *WT1* (OMIM: 607102) are associated with several syndromes that include diaphragm defects

within their phenotypic spectrum, including Denys-Drash syndrome, Frasier syndrome, and Meacham syndrome (139). Amongst several other genes, *COUP-TF2* (OMIM: 107773) is located within the 15q26.1-26.2 cytogenetic hotspot for syndromic CDH (140). The occurrence of CDH in mice with a conditional deletion of *Coup-tf2* and the later discovery of specific mutations in this gene in individuals with CDH support its importance in diaphragm development (103,141). *GATA4* (OMIM: 600576) is found within the CDH critical region 8p23.1 (138), with subsequent reports of *GATA4* variants in humans causing CDH (142). Similarly, *FOG2* (OMIM: 603693) was originally linked to CDH as it is in the critical region 8q22-8q23 (138), with later studies showing specific mutations/deletions of this gene were recurrently associated with CDH (109,143). *WT1*, *COUP-TF2*, *GATA4* and *FOG2* are all expressed in the developing diaphragm, and all are strongly linked to CDH via both mouse and human mutations. As discussed above and in the context of the Retinoid Hypothesis, all four genes have also been linked to retinoic acid signaling, providing supporting evidence for the importance of this pathway in normal diaphragm development.

Interestingly, *COUP-TF2*, *GATA4* and *FOG2* are all linked to retinoic acid signaling via interactions with the RARs. While the evidence is less strong, there are other genes that can be linked to the retinoic acid signaling pathway in this way. Predicted pathogenic variants in *NSD1* have been identified in a cohort of fetuses with CDH (144). This gene is also associated with Beckwith-Wiedemann syndrome that infrequently includes diaphragm defects (144–146). Moreover, *NSD1* is another protein that can interact with nuclear receptors, including the RARs (147). Similarly, while there has only been one known report linking it with CDH, *SIN3A* encodes a transcriptional corepressor that interacts with RARs (26,148). Last, while its exact

interaction with RAR signaling remains unclear, *KIF7* and retinoic acid signaling have been linked to the development of diaphragm defects in mice and three cases of CDH (149–151).

In addition to genes that may interact with retinoic acid signaling at the level of the receptor, diaphragm defects have been described in association with mutations in numerous retinoic acid target genes. In our past analysis of CDH-associated genes, we identified an inclusive list of 218 genes linked with CDH, 44 of which were retinoic acid target genes (21) This category includes genes regulated by retinoic acid that also have direct roles in retinoic acid signaling (e.g. *STRA6*, *RALDH2*, and *RARB*), and indirect interactions with retinoic acid signaling (e.g. *WT1*, *COUP-TF2*, *GATA4*, and *FOG2*). Thus, it is possible to envisage how aberrant retinoic acid signaling in the developing diaphragm could impact the expression of multiple genes linked with diaphragm development and precipitate CDH.

As we learn increasingly more about the genetics of CDH, a common pathway that has emerged is retinoic acid signaling, providing support for the Retinoid Hypothesis in human CDH (21). As discussed above, many CDH-associated genes can be linked with the retinoic acid signaling pathway, the challenge for the future will be to determine how these genes contribute to diaphragm development and whether it is possible to construct a gene regulatory network to show the relationship between these genes and how aberrant retinoic acid signaling can lead to abnormal diaphragm development.

2.3.5 Epidemiological studies and the Retinoid Hypothesis

Population-based studies have given some insight into the link between CDH and the Retinoid Hypothesis, this primarily relates to the importance of adequate maternal dietary vitamin A intake. One of the first epidemiological studies to link vitamin A intake and CDH

leveraged maternal nutrient intake data between 1997 and 2003 from the US National Birth Defects Prevention Study (152). Vitamin A intake below the 10th percentile was associated with isolated CDH in women who did not use periconceptional vitamin supplements (retinol intake, odds ratio [95% CI] = 2.1 [1.1, 3.9]), and in women who did use periconceptional vitamin supplements (total vitamin A intake, odds ratio [95% CI] = 1.7 [1.2, 2.6]). The authors concluded that this data supported the Retinoid Hypothesis, but it is important to note that an expanded analysis of the US National Birth Defects Prevention Study including maternal nutrient intake data from 1997 to 2011 did not find a significant association between vitamin A intake and CDH (153). Despite this lack of agreement, two other studies from The Netherlands and Japan have shown that low maternal vitamin A intake confers an increased risk of CDH (80,81). In a Dutch population, Beurskens *et al.*, showed that maternal dietary vitamin A intake below the recommended daily intake (<800 µg vitamin A per day) in normal weight mothers was significantly associated with an increased risk of CDH (odds ratio [95% CI] = 7.2 [1.5, 34.4]). In a Japanese population, Michikawa *et al.*, showed that in a similar group of normal weight mothers, high total maternal dietary vitamin A intake was associated with a reduced risk of CDH (odds ratio [95% CI] = 0.6 [0.3, 1.2]). Taken together there is collective epidemiological evidence that low maternal dietary vitamin A intake is a risk factor for CDH, although this concept requires further exploration.

While there is evidence to suggest that maternal dietary vitamin A intake is an important factor in the etiology of CDH, it has been observed that there is no evidence to suggest an increased incidence of CDH in countries with high rates of vitamin A deficiency (115), although a counterpoint to this argument is that birth defect registries in these countries are inadequate, and data regarding the global incidence of CDH is incomplete (3). While considering maternal

dietary vitamin A intake as a risk factor for CDH it is important to highlight that even in developed countries inadequate dietary intake can be prevalent in the general population, and has been reported in 15.5% of pregnancies in the USA (n = 1,003), and 10% in Poland (n = 1,764) (154,155) . In a separate study, ~7% of women of childbearing age in the USA had low serum retinol concentrations, a factor that was linked to lower socioeconomic status (156). While overt vitamin A deficiency might not be a major contributor to the occurrence of CDH, we believe that inadequate dietary vitamin A intake is a risk factor that may combine with other underlying factors (genetic or environmental) to cause CDH (82). Indeed, we echo the recent remarks made by Gilbert and Gleghorn (114), supporting adequate maternal vitamin A intake during pregnancy at the population level to help prevent CDH, and possibly other birth defects.

Regarding other environmental risk factors that may intersect with the Retinoid Hypothesis, it is interesting to note that others have drawn links between maternal alcohol use and cigarette smoking and altered retinoid signaling (157). While these links are speculative and require further study, there is experimental evidence that links both alcohol exposure and cigarette smoke exposure to alterations in vitamin A metabolism (157–159).

2.4 The Retinoid Hypothesis and the lung in CDH

In CDH, incomplete development of the diaphragm allows the abdominal contents to herniate into the thoracic cavity, damaging the growing lungs (5). While beyond the scope of the current review, it is important to acknowledge that retinoic acid signaling is important in lung organogenesis (160–162). In 2000, Keijzer *et al.* proposed the Dual-hit Hypothesis to explain pulmonary hypoplasia in teratogenic models of CDH (163). This hypothesis posited that lung formation was affected twice in CDH, once before the diaphragm develops, and another after

abnormal diaphragm formation and herniation. Given the known role of retinoic acid signaling in lung development, and its emerging role in the diaphragm, it is possible to surmise that the Retinoid Hypothesis is consistent with a dual-hit model of lung damage in CDH. In this scenario, abnormal retinoic acid signaling could contribute to abnormal lung development (first hit), and this would be compounded by the effects of abnormal retinoic acid signaling on diaphragm formation leading to abdominal organ herniation and further damage to the lungs (second hit), as explored by others (164–166). Whether altered retinoic acid signaling contributes to both abnormal diaphragm and lung development in CDH remains to be established, nevertheless there is a literature exploring antenatal vitamin A treatment in preclinical models of CDH to improve lung development (167–169).

2.5 Different types of diaphragm defects and the Retinoid hypothesis

As introduced, there are different types of diaphragm defects in CDH. It should be highlighted that the Retinoid Hypothesis is most strongly linked to Bochdalek CDH. Most of the evidence supporting the Retinoid Hypothesis comes from teratogenic, dietary, and genetic models that recapitulate this type of CDH. Moreover, in our analysis of CDH-associated genes, we identified a link between genes involved in retinoid signaling and Bochdalek CDH, but not the other types of diaphragm defects (21). Thus, while the Retinoid Hypothesis helps explain the etiology of Bochdalek CDH, there is little experimental evidence to support that this hypothesis also explains the etiology of diaphragm eventration, central tendon defects, and Morgagni hernias. While Bochdalek CDH is the most commonly occurring and clinically relevant type of CDH, there is a gap in our understanding of the etiology of other rare types of CDH that should be the focus of future research.

2.6 Alternative hypotheses to explain the etiology of CDH

While the Retinoid Hypothesis primarily explains Bochdalek CDH, it should also be acknowledged that this hypothesis may not be a unifying explanation for all cases of Bochdalek CDH, this is most evident when considering CDH-associated genes. For example, in their permissive list of 218 CDH-associated genes, Dalmer *et al*. only identified 24.7% with a discernible link to retinoid signaling (21). Indeed, variants in *LONP1* have been reported to be the most commonly occurring in CDH (170), but this gene has no discernible link to retinoid signaling. This raises the questions of alternates to the Retinoid Hypothesis. For example, altered thyroid hormone signaling was suggested to contribute to CDH, although experimental evidence does not support this (74). Similarly, while it has been suggested that maternal intake of nutrients associated with one-carbon metabolism might be associated with CDH (152,153), direct assessment of biomarkers of this pathway were not associated with the risk of CDH (171). Other alternative hypotheses that remain to be fully explored include a possible role for maternal dietary vitamin D intake (172,173), as well as an emerging interest in extracellular vesicles and micro RNAs (5,174).

2.7 Future perspectives on the Retinoid Hypothesis

As we accumulate knowledge in support of the Retinoid Hypothesis it is important that we consider its limitations and be cognizant of alternative hypotheses. As described in the text above, there are many gaps in our knowledge regarding the Retinoid Hypothesis. Here, we discuss three questions that we believe should guide future studies into the etiology of CDH and the Retinoid Hypothesis.

First, how does the Retinoid Hypothesis relate to CDH in humans? The strongest evidence for the Retinoid Hypothesis comes from animal models, but there is a growing literature from CDH in humans that supports the hypothesis. In order to further strengthen this evidence base, we think there should be a continued emphasis on studying the importance of maternal dietary vitamin A intake as a risk factor for CDH, more attention paid to markers of maternal and fetal vitamin A status in cases of CDH, a focus on gene-nutrient interactions at both the maternal and fetal level, and last, a continued exploration of intersections between the retinoid signaling pathway and genes associated with CDH.

Second, how does the Retinoid Hypothesis explain abnormal diaphragm development at a cellular level? It remains to be determined how disturbed retinoid signaling precipitates abnormal diaphragm development at a cellular level. This is confounded by an incomplete understanding of the cellular pathogenesis of CDH, as well as the normal contribution of retinoid signaling to diaphragm development. It is thought that the non-muscular mesenchymal cells of the PPF are important in the development of CDH (39,41,133), but how the retinoid signaling pathway operates in these cells, and what is the fate of the cells when retinoid signaling is disrupted remains largely unexplored. Future studies using animal models of CDH can help address these gaps, supplemented by studies in PPF-derived cell cultures (175), and patient-derived fibroblasts (176).

Zani and colleagues recently emphasized that by improving our understanding of CDH's etiology we may improve its diagnosis and allow earlier interventions to improve CDH outcomes (5). This leads to our third question; how do we leverage our understanding of the Retinoid Hypothesis to help lessen the impact of CDH? Primarily because of its known importance in lung development, retinoid administration is already being studied as a tool to improve lung

development in preclinical models of CDH (167–169), and improving our understanding of underlying defects in the retinoid signaling might allow this approach to be optimized. Taking this interventionist approach a step further, if early genetic testing revealed mutations in CDH-associated genes that are linked with retinoid signaling, it might be possible to intervene with retinoids to help ameliorate damage to the developing diaphragm and/or lungs, lessening the impact of CDH. At the population level, we and others believe that attaining optimal maternal dietary vitamin A intake may help lessen the impact of CDH (82,114). This would require Public Health efforts highlighting the importance of adequate periconceptional vitamin A intake in women, and their potential benefits with respect to CDH and other congenital anomalies.

Chapter 3: Single Cell Transcriptomic Analysis Provides Insight into Congenital Diaphragmatic Hernia

3.1 Introduction

The diaphragm is a mammalian muscle that serves two distinct purposes: serving as the primary muscle of respiration, and as a barrier between the abdominal and thoracic cavities (36,37). There are multiple developmental precursors of the diaphragm: the septum transversum, the somites, and the pleuroperitoneal folds (PPF). The septum transversum is the first structure seen in development and is present in all vertebrates. This thin layer separates the heart and the lungs from the abdominal organs (177). Although it is unclear if the septum transversum gives rise to any tissue that remain in adulthood, some studies have suggested that this tissue gives rise to the central tendon and non-muscular component of the diaphragm (178), although other studies contradict this theory (38). The somites are the source of the muscle cells in the diaphragm, and this has been confirmed by multiple expression and knockout studies (41,179,180). Finally, the PPFs are a pair of transient structures that are pyramidal in shape and are present between the pleural and peritoneal cavities on either side of the body. Genetic studies utilizing markers of the PPF, specifically *Gata4*, have shown that this tissue gives rise to both the central tendon as well as non muscular fibroblasts in the adult diaphragm (38,39). This tissue is also in charge of recruiting and helping with the migration of muscle cells during the morphogenesis of the diaphragm (39). It is also known that defects in the PPF can give rise to malformations in the mature diaphragm, the most common one being Congenital Diaphragmatic Hernia (CDH).

CDH is a condition that affects ~2.3 per 10,000 live births (3) and is characterized by an incomplete formation of the diaphragm, which results in the herniation of abdominal organs into the thoracic cavity. The mortality of this condition varies depending on the severity of the herniations, as well as if this condition is present with other comorbidities. Furthermore,

although the mortality of CDH in high income countries is approximately 30%, it can reach about 90% of cases in low-income countries (6,8). Although the etiology of CDH is not fully understood, there are two main explanations: genetic and environmental. It is known that around 40% of CDH cases have a specific genotype associated with them, and there are multiple genes for which mutations are known to cause CDH (17,138,181). The second explanation derives from the fact that there are multiple environmental factors that have been associated with a higher incidence of CDH, like maternal age, smoking and maternal diet (3,14,27,182). Retinoic acid (RA) has also been heavily associated with the incidence of CDH, to the extent that a review published in 2003 by Greer *et al.* gave rise to the retinoid hypothesis in CDH, which states that disruptions in the RA signaling cascade causes CDH (31). After the initial publication of the retinoid hypothesis, there has been more evidence linking RA and the incidence of CDH. There have been multiple studies that have associated low maternal vitamin A intake with a higher incidence of CDH in newborns (67,80–82). Furthermore, genetic expression studies have shown disruptions in RA signaling in both teratogen exposed animals that develop CDH, and human cases of CDH (68,166,183).

As the diaphragm arises from multiple embryological sources, studies of the pathogenesis of CDH have been performed to determine the origin of this condition. Multiple studies have shown that the PPF is involved in the formation of CDH. Analysis performed on nitrofen-exposed rat embryos shows that there are defects in the PPF at GD13.5. Furthermore, rats exposed to low maternal vitamin A intake, as well as *Wt1*$^{-/-}$ mice that develop CDH all show defects in the PPF (184). These results are further supported by genetic studies. Mutations in both *Gata4* and *Coup-tfII* cause defects in the PPF, that ultimately lead to the formation of CDH (38,103). Since the PPF is composed of three main cell types; nerve cells, muscle cells and non-

muscular mesenchyme, studies trying to determine which cell type gives rise to CDH have been performed. It is known that expression of *Pax3*, a known muscle marker, is not altered in nitrofen-induced CDH (185). Furthermore, multiple studies have shown that the non-muscular mesenchyme is the main driver in the formation of CDH (38,41,184).

Transcriptomic analysis has been performed in both the PPF and the diaphragm to better understand the genetic expression patterns of these tissues, and to derive candidate genes for the formation of CDH (69). However, this analysis was only able to show global expression patterns, and therefore was not able to assess the differences in gene expression across different cell types. Since the PPF is a cellularly diverse tissue, understanding which populations express different genes will provide better insight into the pathogenesis of CDH. In this study, we performed single cell transcriptomic analysis to better understand the different cellular populations present in the PPF, and the relationship that they have with RA signaling and CDH. We hypothesize that the non-muscular mesenchyme of the PPF will have the highest level of expression of CDH-associated genes and retinoid-related genes, and we hope that through this study we are better able to understand the etiology of CDH, and the relationship between distinct cellular populations in the PPF and the formation of diaphragmatic hernias.

3.2 Methods

3.2.1 Animal Care and Breeding

All experiments were approved by the University of Alberta Research Ethics Committee and were performed in accordance with the Canadian Council on Animal Care guidelines. BALB/c mice were obtained from The Jackson Laboratory (Bar Harbor, ME, USA). The animals were housed in the University of Alberta conventional animal facility, under standard environmental conditions. To generate timed pregnancies, one male mouse and two female mice

were set in a cage after 16:00 hrs and left overnight. The following morning at 9:00 hrs the presence of a vaginal plug was examined to assess for copulation. Each cage was left for a maximum of two nights. After a vaginal plug was seen, noon of that day was considered gestational day (GD) 0.5.

3.2.2 Dam Dissection and Fetus Collection

One dam at GD 13.5 was euthanized via anesthesia using isoflurane inhalation immediately followed by cervical dislocation. An abdominal incision was made in the mice, and the entire uterus was removed. The uterus was placed in a petri dish containing ice cold Phosphate Buffer Saline (PBS; Sigma-Aldrich, Oakville, ON). To dissect the individual fetuses, stereomicroscopy was used (Stemi 508, Zeiss, Oberkochen, Germany). To obtain the PPF, both the cranial and caudal section of the fetus were dissected, and the heart and lungs were removed. For our scRNA-seq sample, one fetus was utilized, and the PPF from each side of the body was obtained and pooled. A sample dissection can be seen in Figure 3.1. The PPFs were kept in ice cold PBS for cell suspension and isolation.

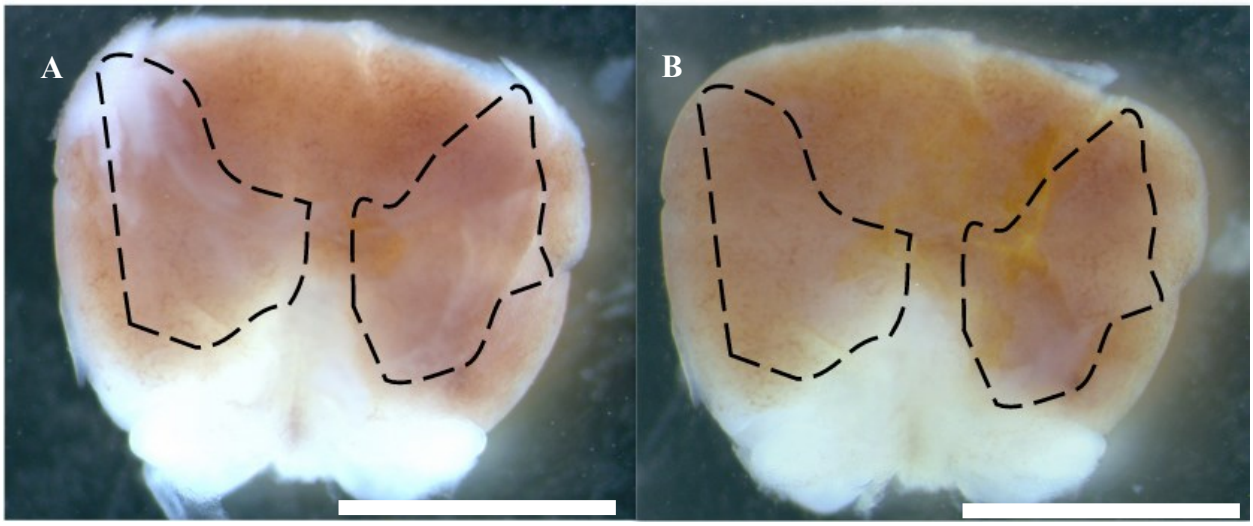


Figure 3.1: PPF dissection for scRNA-seq analysis. Superior view of the liver, before (A) and after (B) the PPF was removed. The dotted black line represents the area in which the PPF spanned the surface of the liver. The scale bars are 0.5 mm.

3.2.3 PPF cell suspension

Cell isolation was performed modifying a previously described method (175). The PPF tissue was transferred to DMEM (Thermo Fisher Scientific, Toronto, ON) with 10% FBS (Thermo Fisher Scientific) and 0.25% trypsin-EDTA (Thermo Fisher Scientific) and was incubated for 9 minutes at 37°C in 5% CO_2. Trypsin neutralization was done by adding 100 µL of DMEM media. The solution was centrifuged at 340 *g* for 5 minutes, and the supernatant was decanted. The cells were resuspended in 100 µL of DMEM. To determine the concentration and viability of the cells, a mixture containing 5 µL of the cell suspension as well as 1 µL of 0.4% trypan blue (Thermo Fisher Scientific) was added to a hemocytometer and the cells were counted using an inverted light microscope (Thermo Fisher Scientific). Both the number of alive and dead cells were counted in the corner quadrants of the hemocytometer to determine the cell concentration and viability of the suspension. The cell suspension was prepared for sequencing using the standard 10X Genomics technology (10X Genomics, San Francisco, CA; Figure 3.2). Library preparation was performed by the High Content Analysis core at the University of Alberta, and sequencing was done by Novogene via Illumina next generation sequencing (Sacramento, CA).

3.2.4 scRNA-seq data analysis

Feature-barcode matrices, read alignment and other analyses were performed utilizing the CellRanger software (10X Genomics). A total of 13,462 cells were in our PPF sample. Data clustering and visualization was performed in two different software packages. Clustering via K-means was performed utilizing Cloupe (10X Genomics). Cells were filtered based on Unique Molecular Identifiers (UMIs) (UMIs < 20,000) as well as features (50 < features < 2500). All secondary analysis performed for the K-means model was also performed in this software. The

second clustering analysis was performed using Seurat V4 in R (186,187). Filtering in Seurat was done by features, with cells being excluded if they had less than 200 or more than 2500 features, as well as mitochondrial counts (percent. mt <5). Identity of the different clusters was determined utilizing a three-step method: (i) PanglaoDB, a single-cell sequencing resource for gene expression data (188), (ii) gene ontology analysis to determine the top biological processes associated with each cluster (189,190) and (iii) expression analysis of known features for certain cell types.

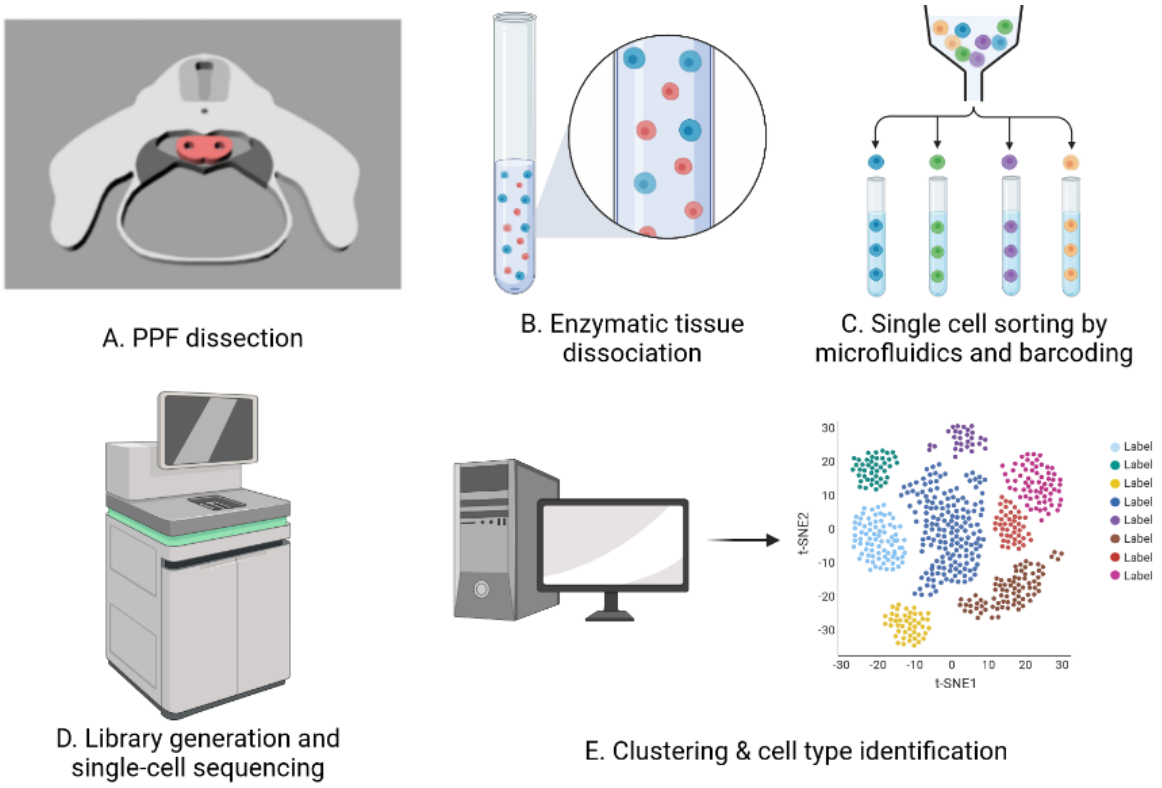


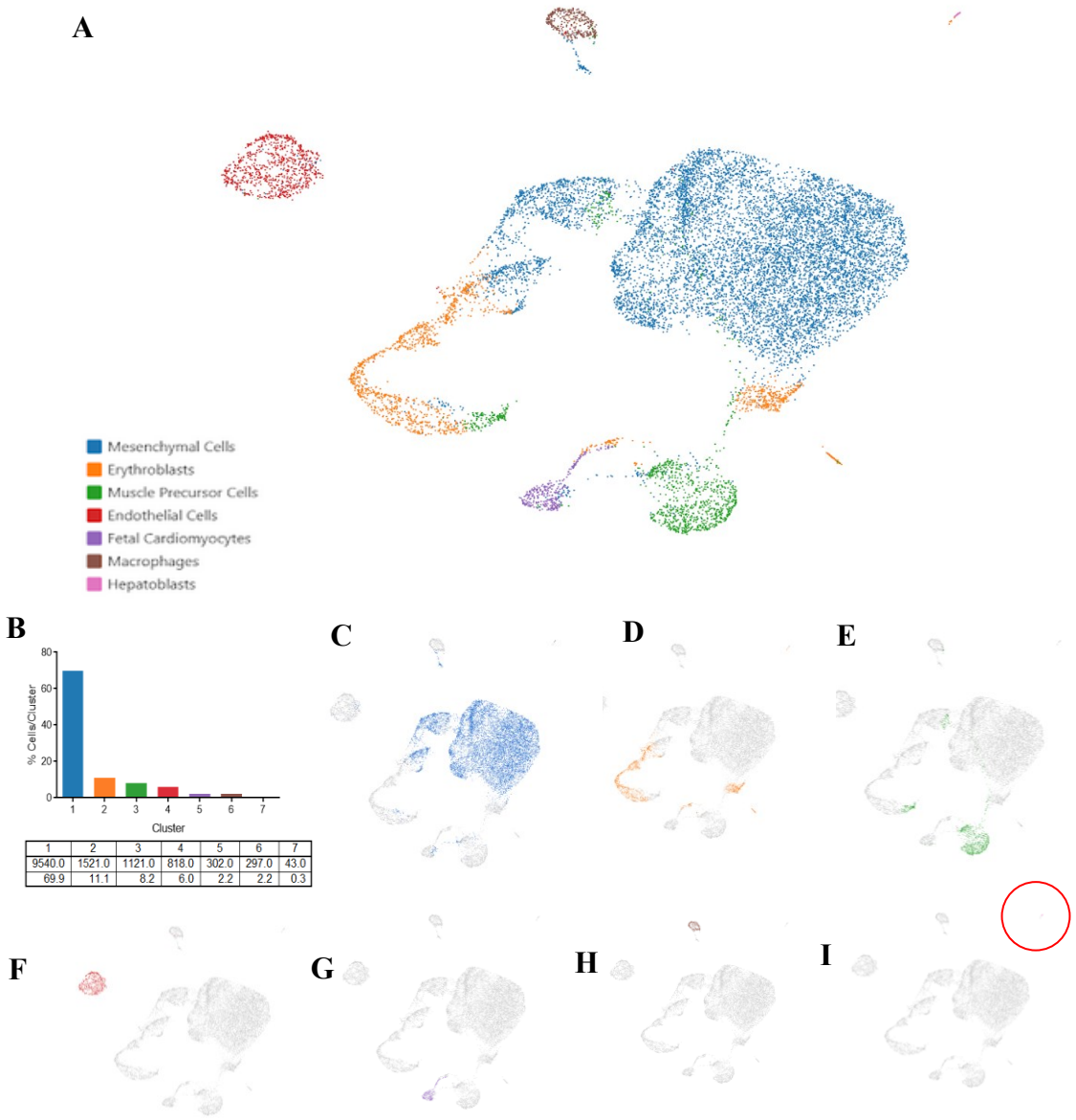
Figure 3.2: Summary of scRNA-seq workflow. After (A) the dissection of the PPF, cells were suspended via enzymatic dissociation (B). Individual cells were given a barcode and sorted (C). Sequencing was performed by Novogene (D). Bioinformatics analysis was performed to determine the different cell clusters present in our PPF sample (E). Created with BioRender.com

3.3 Results

3.3.1 Preliminary Clustering Reveals Distinct Cell Populations in the PPF

Preliminary K-means clustering was performed at first to determine the major cell populations present in our PPF sample. The total number of cells in our sample was 13,462. The clustering analysis can be seen in Figure 3.3. There was a total of 7 clusters based on our K-mean analysis, with nearly 70% of all cells belonging to the mesenchymal cell cluster. The other clusters listed based on the highest percentage of cells were: (i) Erythroblasts (11%), (ii) Muscle Precursor Cells (8.2%), (iii) Endothelial Cells (6.0%), (iv) Fetal cardiomyocytes (2.2%), (v) Macrophages (2.2%) and (vi) Hepatoblasts (0.3%). Annotation of the clusters was determined as described in the methods: Cluster 1 was classified as **Mesenchymal cells** (CL_0008019), based on automated cluster identification (top ranked term = "Fibroblasts"), complimentary coexpression data, and multiple relevant enriched GO Biological Processes (e.g. *animal organ morphogenesis*, and *tissue morphogenesis*). Note, in the context of embryonic development, the term mesenchymal cell was preferred over fibroblast. Cluster 2 was classified as **Erythroblasts** (CL_0000765), based on automated cluster identification (3rd ranked term = "Erythroid-like"), complimentary coexpression data, and multiple relevant enriched GO Biological Processes (e.g. *erythrocyte homeostasis*, and *erythrocyte differentiation*). Cluster 3 was classified as **Muscle precursor cells** (CL_0000680). The leading cell type based on automated cluster identification was "Unknown". Analysis of coexpressed genes also returned many poorly defined cell types; however, 9 of the top-expressed genes were classified as hallmark genes of myogenesis (*Ckb, Des, Myl1, Myog, Lama2, Actc1, Mylpf, Tnnt1, Eno3*) (191). Moreover, there were multiple enriched GO Biological Processes linked with muscle development (e.g. *muscle organ*

development, *muscle structure development*). As such, this cell cluster was classified as representing muscle precursor cells. Cluster 4 was classified as **endothelial cells** (CL_0000115). This classification is based on automated cluster identification (1st ranked term = "Endothelial"), complimentary coexpression data, and enriched GO Biological Processes (e.g. *angiogenesis*, *blood vessel morphogenesis*). Cluster 5 was defined as **Fetal cardiomyocytes** (CL_0002495). This classification is based on automated cluster identification (8th ranked term = "Cardiomyocyte"), complimentary coexpression data, and enriched GO Biological Processes (e.g. *heart contraction*, *heart process*). Cluster 6 was classified as **Macrophages** (CL_0000235). This classification is based on automated cluster identification (top ranked term = "Macrophages"), complimentary coexpression data, and enriched GO Biological Processes (e.g. *regulation of immune system process*, and *defense response to other organism*). Cluster 7 was classified as **Hepatoblasts** (CL_0005026). This classification is based on automated cluster identification (4th ranked term = "Hepatocytes") and complimentary coexpression data. It is important to note that the hepatoblasts cluster was removed from further analysis as it was considered sample contamination.



1	2	3	4	5	6	7
9540.0	1521.0	1121.0	818.0	302.0	297.0	43.0
69.9	11.1	8.2	6.0	2.2	2.2	0.3

Figure 3.3: K-means cluster analysis of the PPF. The seven clusters identified can be seen in the UMAP based on different colors, shown in the Figure legend (A). The distribution of cells per cluster is graphed along tabulated numbers showing the total number of cells, as well as the percentage of cells (B). UMAPs highlighting each individual cluster are shown (C-I). Cluster 1, 'Mesenchymal Cells' is in blue (C), cluster 2, 'Erythroblasts' is shown in orange (D), cluster 3, 'Muscle Precursor Cells', is green (E), cluster 4, 'Endothelial Cells' is red (F), cluster 5 'Fetal Cardiomyocytes' is purple (G), cluster 6 'Macrophages' is brown (H), and cluster 7 'Hepatoblasts' is pink (I; highlighted with a red circle).

3.3.2 Seurat Analysis Further Characterizes the Distinct Populations Present in the PPF

To further analyze our scRNA-seq sample derived from the PPF and identify subclusters of cells, UMAP clustering was performed utilizing the Seurat package in R. A second method of clustering was performed since K-means clustering has limitations in the numbers of clusters it can form. UMAP clustering was utilized since it does not have this issue, and it also shows the similarity between clusters (186,192). Clusters that are closer together are more similar than those farther apart. By utilizing this clustering method, a total of 18 clusters were identified. The UMAP clustering performed by Seurat is shown in Figure 3.4. Cluster identification was performed in the same three-step process described above. Clusters 1-4, as well as cluster 15, were identified as **Mesenchymal cells** (CL_0008019), due to automated cluster identification ("Fibroblasts" was in the top 3 terms for all clusters), coexpression data, and GO biological processes ("*Tissue morphogenesis*", "*Body Morphogenesis*", "*Cell-matrix adhesion*"). Clusters 5 and 13 were identified as **Erythroblasts** (CL_0000765) 1 & 2 respectively. This was based on coexpression data, as well as many relevant GO analysis hits (i.e., "*Erythrocyte development*", "*Erythrocyte differentiation*"). Cluster 6 was labelled **Epithelial Cells** (CL_0000066), based on coexpression data as well as the GO hit results ("*morphogenesis of an epithelium*", "*epithelial tube morphogenesis*"). Cluster 7 was named **Smooth Muscle Cells** (CL_0000192), due to the coexpression data, GO analysis results ("*circulatory system development*", "*blood vessel development*") and the fact that this clusters expresses *Myh11*, a known marker for smooth muscle tissue (193). Cluster 9 was labelled **Mesothelial Cells** (CL_0000077), due to the expression patterns of mesothelial gene markers like *Wt1* and*Upk3b* in this cluster (194,195). Cluster 10 and 16 were identified as **Fetal Cardiomyocytes** (CL_0002495), based on coexpression data, GO analysis hits ("*heart contraction*", "*cardiac myofibril assembly*"), as well as expression of *Myh6*, a known marker for cardiomyocytes, in both clusters (196). Clusters 11

and 17 were identified as **Endothelial Cells** (CL_0000115), due to automated cluster identification ("Endothelial Cells" was the top result for both clusters), coexpression data, GO analysis terms ("*Angiogenesis*", "*Vasculature Development*") and expression of multiple endothelial cell gene markers like *Cd34* and *Vwf* (197). Cluster 12 was identified as **Myoblasts** (CL_0000056) due to results from the GO analysis ("*Muscle cell differentiation*", "*Muscle cell development*"), as well as expression of *Myog* and *Myod1* (191). Cluster 14 was identified as **Macrophages** (CL_0000235) due to automated cluster identification (top hit for this cluster was "Macrophages"), coexpression data, as well as several relevant GO analysis hits ("*immune effector process*", "*innate immune response*").

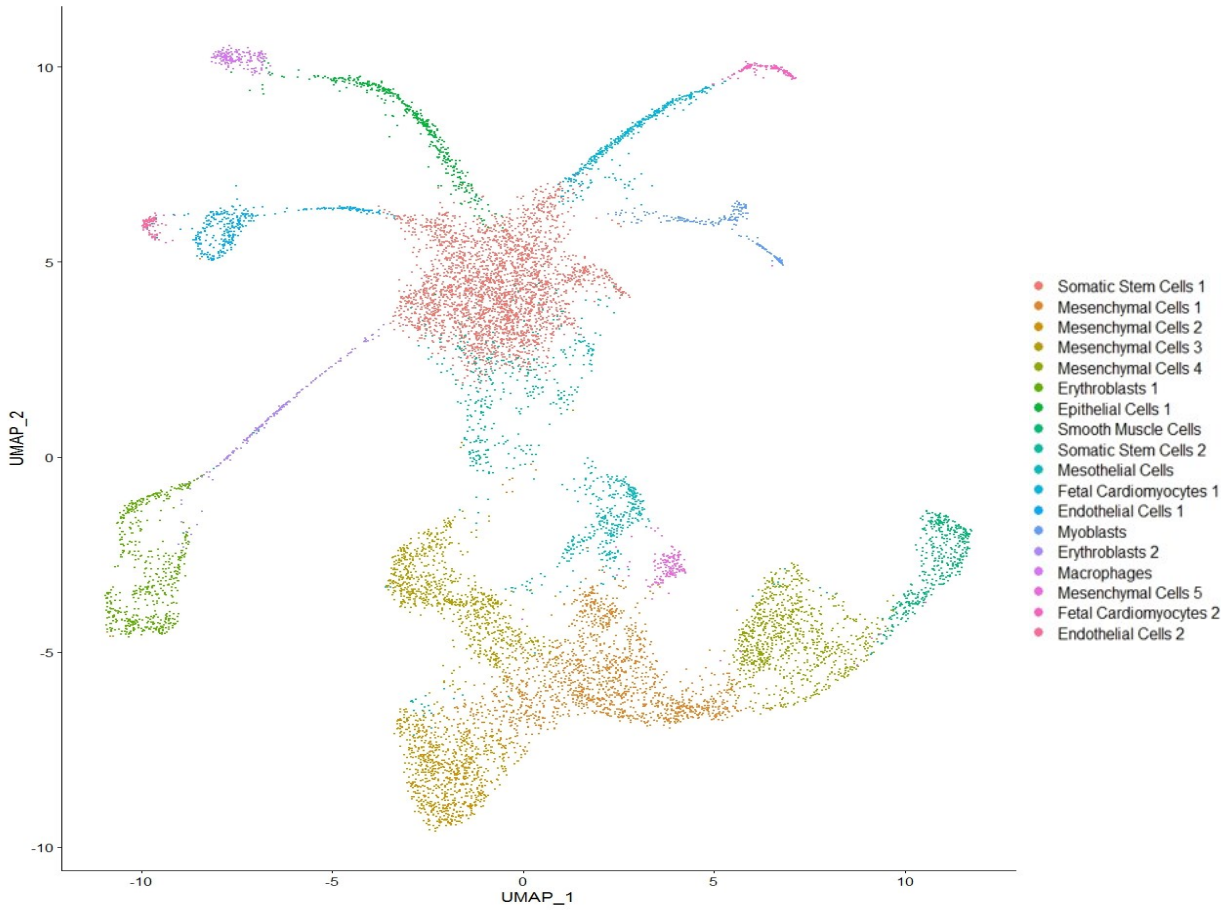


Figure 3.4: Secondary clustering analysis showcases distinct cell populations. UMAP clustering of the scRNA-seq data utilizing the Seurat package in R. A total of eighteen clusters were found. Cluster labels are organized in order of the number of cells per cluster. Each cluster is represented with a different color, which can be seen in the Figure legend on the left side of the UMAP.

As seen in Figure 3.4 there were a total of 18 different clusters, and 10 distinct cell populations. Some cell populations, like the mesenchymal cells, have multiple cluster IDs. We were also able to assess the top markers for each cluster, ensuring that each cluster had distinct expression patterns. A heatmap of the top 15 markers per cluster, as well as a dot plot of the top gene marker for each cluster is presented in Figure 3.5. As seen in the heatmap, the top 15 markers are enriched for each cluster, validating our clustering method. Based on the DotPlot, some of the top genes for some clusters are known markers for a specific cell type. For example, the top expressed gene for the Erythrocyte 1 & 2 clusters were *Hbb-bs* and *Hba-x* respectively, which are known markers for this type of cells (198).

As we performed two different clustering analyses, we were interested in investigating the cell fate across the different clustering methods. An alluvial plot representing this is presented in Figure 3.6. Some of the clusters, like the cardiac cluster in our K-means analysis, translated almost exclusively into either the fetal cardiomyocytes 1 or 2 clusters, demonstrating similarities in the different clustering methods. Furthermore, the alluvial plot shows that Seurat analysis was successful in determining different subclusters for a specific cell population of K-means. However, some clusters, like the hematocytes cluster in the K-means analysis, were divided into many different clusters for the UMAP performed with Seurat, demonstrating the limitations and biases present in the K-means analysis in Cloupe, since the number of clusters is predetermined.

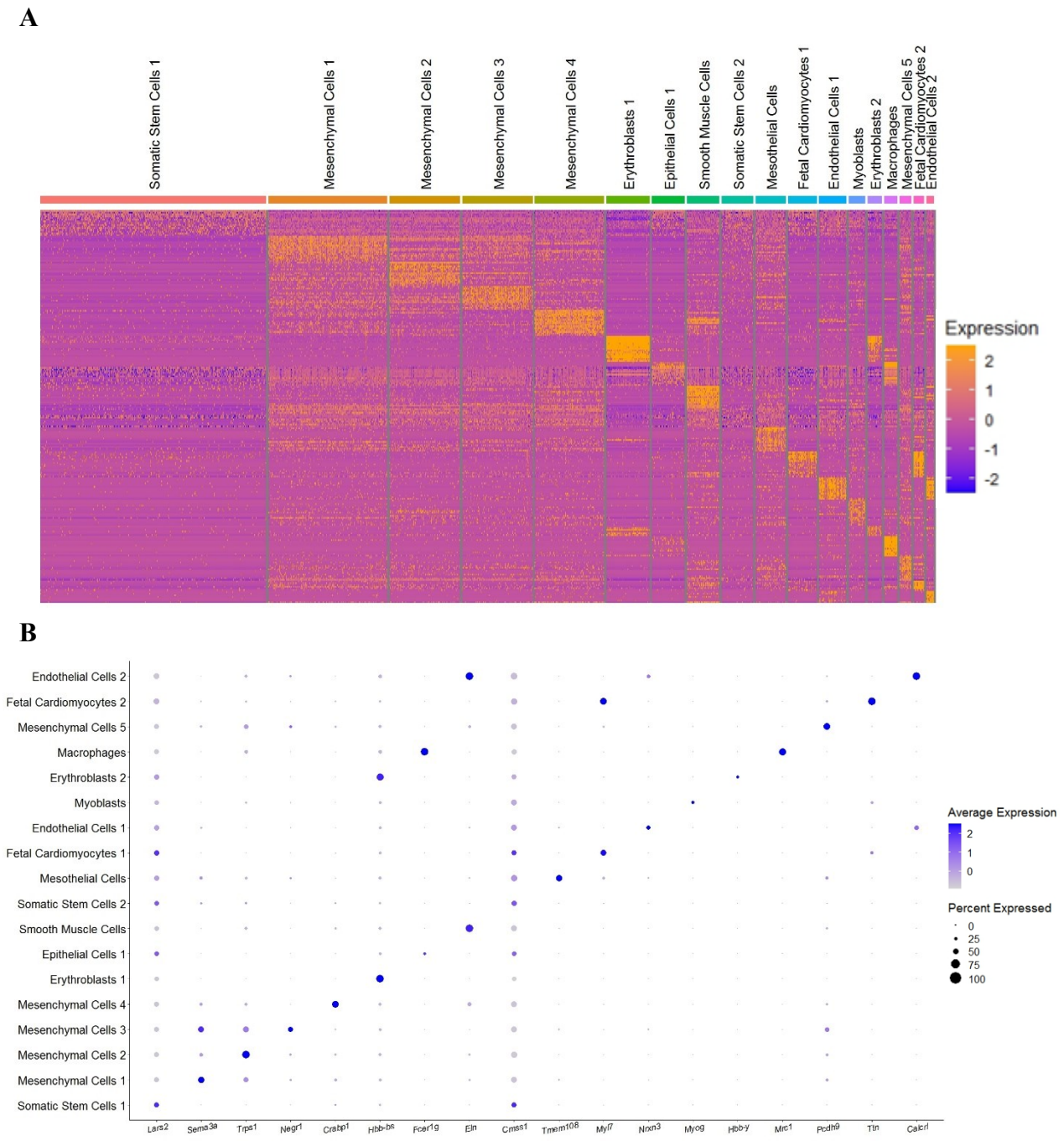


Figure 3.5: Clustering analysis reveals cell clusters with distinct gene expression patterns. The heatmap (A) presents the top 15 markers of each cluster based on the Seurat analysis. Orange represents higher levels of expression. The dot plot (B) shows the top expressed gene per cluster. The circle size represents the number of cells per cluster expressing a certain gene, whereas the color represents expression levels.

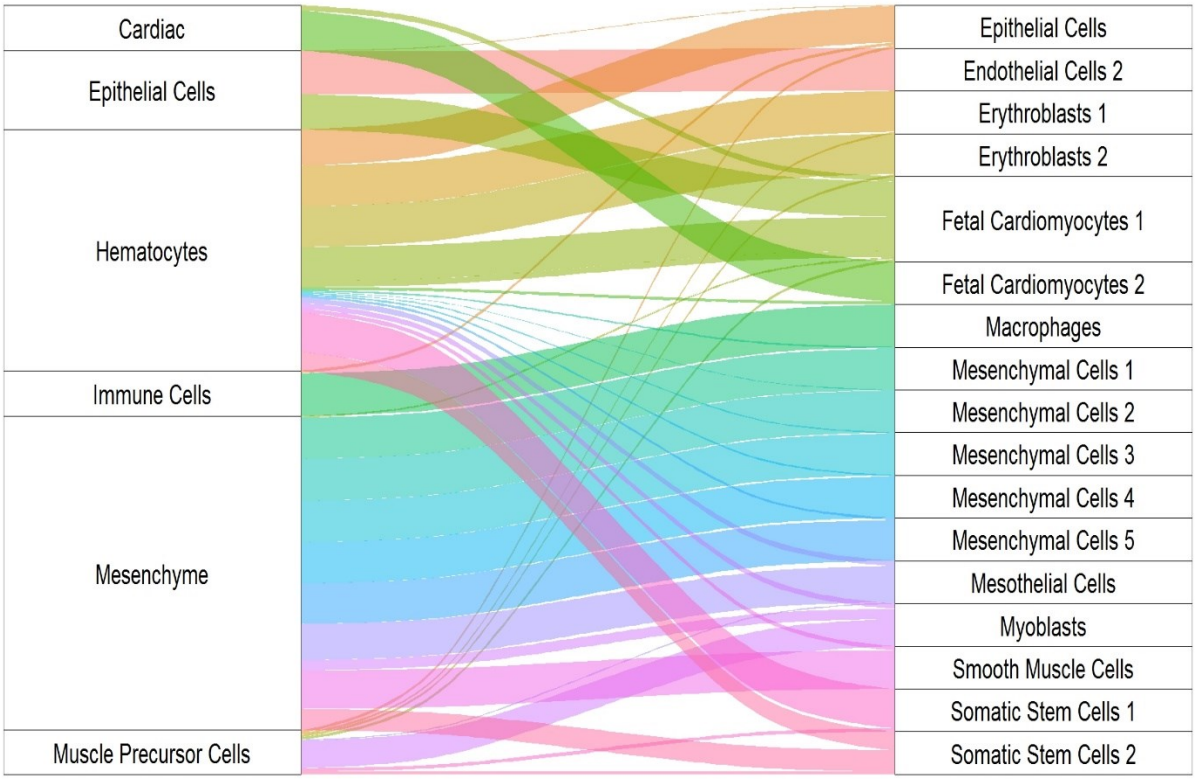


Figure 3.6: Cell identities are conserved following re-clustering of single cell expression data. An alluvial plot linking our K-means analysis (left) and re-clustering performed using Seurat (right) is shown. How the initial six clusters correspond to the 18 clusters generated by re-clustering is reflected in the lines linking either side of the alluvial plot. Wider lines represent a higher percentage of cells from a K-means cluster going into the cluster for the Seurat UMAP. Each color in the alluvial plot represents a different cluster found in the UMAP performed in Seurat.

3.3.3 Expression Analysis Reveals Important Clusters for CDH Pathogenesis

To determine if a cell population might have a higher impact on the pathogenesis of CDH, the expression of multiple CDH-associated genes was assessed in our PPF sample. We initially utilized a list of 31 different genes that have been highly associated with CDH in either humans or mice that was previously described in the literature (199). Expression levels of these genes were quantified and can be seen in a heat map in Figure 3.7A. Expression analysis demonstrated that these genes are mostly present in the mesenchymal component of the PPF, as well as in the 'Smooth Muscle Cells' cluster. Next, we broadened this analysis of CDH-associated genes by using an inclusive list of 218 genes that were previously identified to be associated with diaphragmatic hernias (21). Just like with our previous gene list, expression was primarily seen in the mesenchymal, mesothelial, and smooth muscle cell components of our PPF sample. The heatmap for this list of CDH-associated genes can be seen in Figure 3.7B. To continue our analysis, we determined the expression pattern of a subset of genes that are known to cause Bochdalek hernias in either mice or humas from the 31 genes highly associated with CDH. This subset contained the following genes: *Dnase2*, *Efnb1*, *Wt1*, *Gli2*, *Gli3*, *Lrp2*, *Msc*, *Nr2f2*, *Pdgfra*, *Rara*, *Rarb*, *Stra6*, *Tcf21*. The level of expression of each gene per cluster can be seen in Table 3.1. Violin plots were generated if more than 25% of cells of any cluster expressed a specific gene (Figure 3.8). As shown, *Wt1*, *Gli2*, *Gli3*, *Nr2f2*, *Pdgfra*, and *Tcf21* were primarily expressed in the mesenchymal and mesothelial components of the PPF sample (Table 3.1 and Figure 3.8). *Wt1* expression was highest in the 'Mesothelial Cells' (59% of cells), 'Mesenchymal Cells 1' (50% of cells) and 'Mesenchymal Cells 3' (46% of cells) clusters. *Gli2* expression was highest in the 'Smooth Muscle Cells' (38% of cells), 'Mesenchymal Cells 2' (37% of cells) and 'Mesenchymal Cells 4' (34% of cells) clusters. *Gli3* was expressed the highest in the 'Mesenchymal Cells 4', (59% of cells) 'Mesenchymal Cells 3' (57% of cells) and 'Mesenchymal

Cells 2' (56% of cells) clusters. *Nr2f2* expression was highest in the 'Endothelial Cells 1' cluster (55% of cells), followed by 'Mesenchymal Cells 2' (53% of cells) and 'Fetal Cardiomyocytes 2' (46% of cells). The top three clusters expressing *Pdgfra* were the 'Mesenchymal Cells 2' (74% of cells), 'Mesenchymal Cells 1' (51% of cells), 'Mesenchymal Cells 5' (47% of cells). Finally, less than 25% of cells expressed *Tcf21* in all clusters except for one, the 'Mesenchymal Cells 5' cluster, in which a total of 56% of cells expressed this gene. Based on our expression analysis it is clear that CDH-associated genes are primarily expressed in the mesenchymal and mesothelial components of the PPF, revealing cell populations that are of high importance in the pathogenesis of CDH.

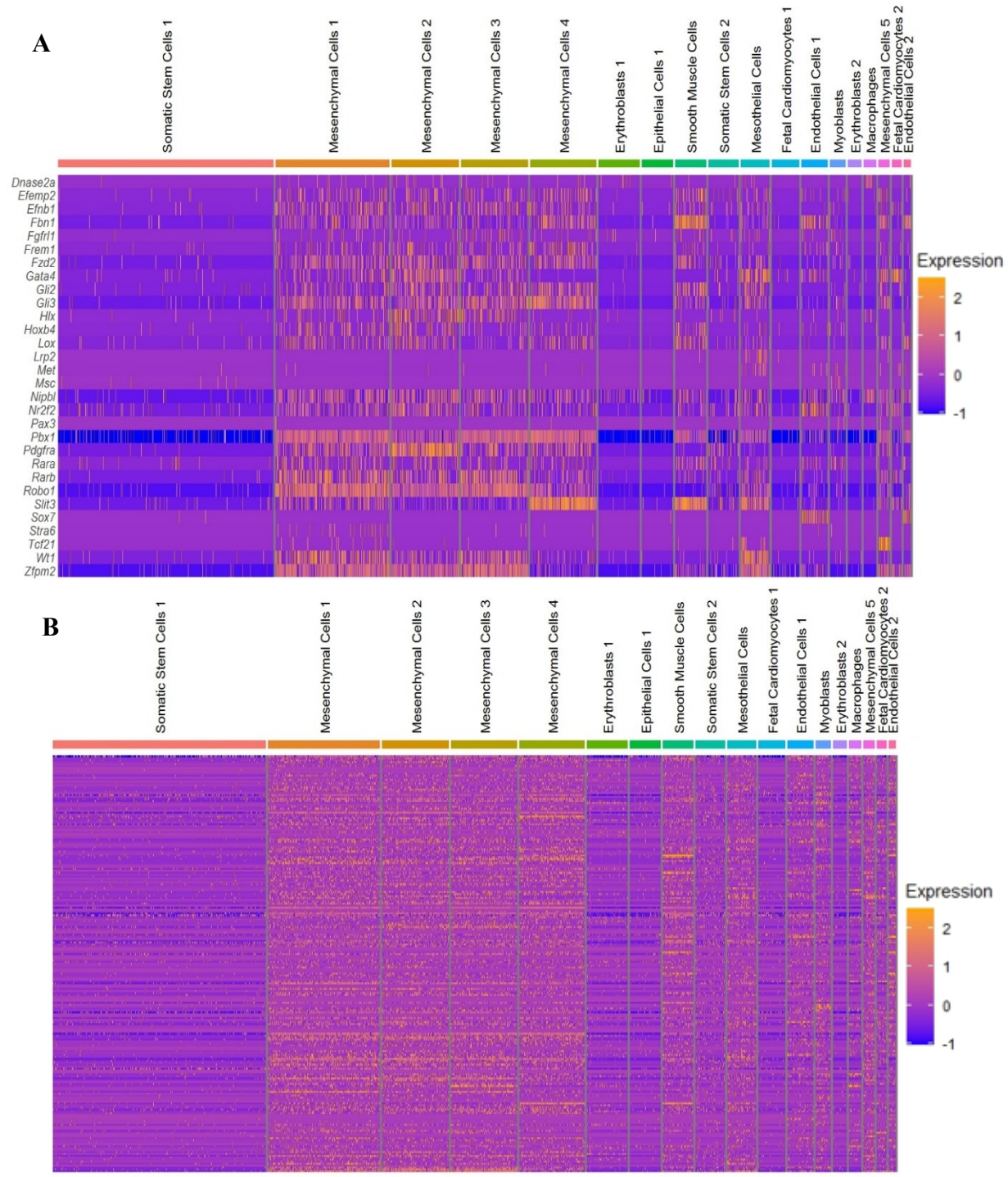


Figure 3.7: CDH-associated genes present varied levels of expression across the PPF. Heat maps for different CDH-associated genes. The expression levels for the 31 gene list are presented in A. B shows the heatmap from the 215 genes derived from gene ontology analysis. High levels of expression are seen in orange, whereas low levels are seen in blue.

	Percentage of cells expressing a gene (%)												
Cluster ID	*Dnase2a*	*Efnb1*	*Wt1*	*Gli2*	*Gli3*	*Lrp2*	*Msc*	*Nr2f2*	*Pdgfra*	*Rara*	*Rarb*	*Stra6*	*Tcf21*
Somatic Stem Cells 1	0.2	1.1	2.1	1.1	2.6	0.0	0.8	5.1	1.9	1.7	1.4	0.1	0.8
Mesenchymal Cells 1	3.9	24.8	50.3	30.0	56.7	0.2	0.2	37.7	51.9	18.0	52.5	5.8	4.0
Mesenchymal Cells 2	4.7	16.0	27.4	37.7	57.4	0.0	0.0	53.7	74.7	19.4	43.2	0.8	0.7
Mesenchymal Cells 3	2.6	19.7	46.2	27.2	59.8	0.3	0.2	32.7	45.2	15.1	50.8	2.3	0.3
Mesenchymal Cells 4	1.3	19.0	15.4	34.7	65.6	0.2	0.1	29.1	43.1	16.8	28.3	1.6	0.2
Erythroblasts 1	6.4	0.7	2.0	1.3	2.1	0.0	0.2	3.6	2.7	1.1	2.0	0.2	0.0
Epithelial Cells 1	0.5	1.4	1.4	0.9	0.5	0.0	0.0	1.9	0.5	1.2	0.9	0.2	0.2
Smooth Muscle Cells	1.7	21.6	1.7	38.2	40.0	0.7	0.2	23.7	7.1	17.3	15.6	1.2	0.2
Somatic Stem Cells 2	0.7	6.1	11.2	8.0	14.6	0.0	0.2	17.8	21.0	7.6	11.0	0.2	1.0
Mesothelial Cells	8.1	11.9	59.7	30.9	42.0	17.5	0.3	36.7	6.3	8.9	32.7	1.3	11.9
Fetal Cardiomyocytes 1	0.3	1.9	1.6	1.1	1.1	0.5	0.3	7.2	1.9	1.6	2.4	0.3	0.3
Endothelial Cells 1	1.4	2.8	4.2	0.3	15.1	0.3	0.3	55.2	6.2	11.8	13.7	0.6	0.6
Myoblasts	0.5	10.1	4.6	13.4	26.3	0.0	10.6	28.1	1.8	25.8	5.5	4.1	0.5
Erythroblasts 2	2.1	1.6	1.6	0.5	1.1	0.5	0.0	1.1	1.1	0.5	0.0	0.0	0.0
Macrophages	21.4	2.3	1.7	1.2	2.3	0.6	0.0	4.6	4.6	4.6	1.2	0.6	0.0
Mesenchymal Cells 5	1.3	19.2	11.5	21.2	59.0	0.6	0.0	29.5	47.4	11.5	48.7	0.6	59.6
Fetal Cardiomyocytes 2	0.0	22.6	2.3	10.5	6.8	0.0	0.0	46.6	18.0	10.5	18.0	0.0	1.5
Endothelial Cells 2	1.0	1.0	2.0	1.0	10.0	2.0	0.0	1.0	6.0	7.0	7.0	0.0	0.0

Table 3.1: Expression analysis for CDH-associated genes. Percentage of cells expressing the different Bochdalek CDH-associated genes across all the different clusters.

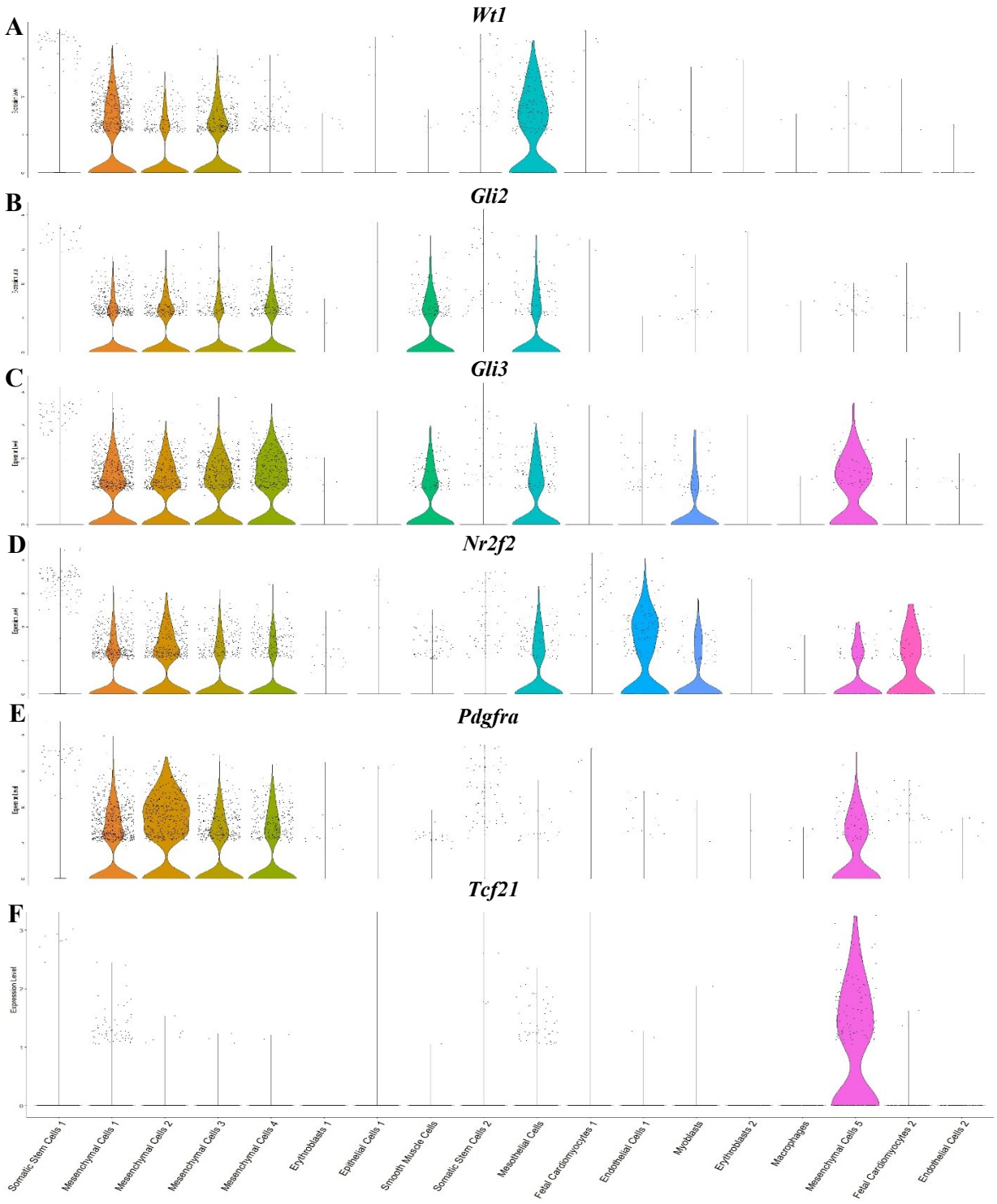


Figure 3.8: CDH-associated genes are primarily expressed in the mesenchymal component of the PPF. Violin plots for (A) *Wt1*, (B) *Gli2*, (C) *Gli3*, (D) *Nr2f2*, (E) *Pdgfra* and (F) *Tcf21*, which are all Bochdalek CDH-associated genes. Genes were chosen if they had more than 25% of cells expressing it in any cluster.

3.3.4 Expression Analysis Reveals Important Clusters for Retinoid Signaling in the developing diaphragm

Given the strong association between the retinoic acid signaling pathway and the development of CDH, we were interested in assessing the expression level of genes associated with this pathway in our scRNA-seq dataset. We derived a list of 24 genes associated with various aspects of the retinoic acid signaling pathway from the literature (200), 8 of which had a cluster in which at least 25% of cells were expressing said gene in our PPF dataset (Table 3.2). *Crbp1* was the top-expressed gene in the PPF, it was expressed preferentially in the 'Mesenchymal Cells 2' (84% of cells), 'Mesenchymal Cells 1' (79% of cells) and 'Mesenchymal Cell 5' (78% of cells) clusters. *Aldh1a2* show highest levels of expression in the 'Mesothelial Cells' (57% of cells), 'Mesenchymal Cells 1' (30% of cells) and 'Mesenchymal Cells 5' (11% of cells) clusters. *Crabp1* was expressed the most in the 'Mesenchymal Cells 4' (90% of cells), 'Mesenchymal Cells 1' (33% of cells) and 'Mesenchymal Cells 2' (29% of cells) clusters. *Crabp2* was highest expressed in the 'Mesenchymal Cells 2' (57% of cells), 'Mesenchymal Cells 4' (40% of cells) and 'Mesenchymal Cells 3' (37% of cells) clusters. For *Rara,* only the 'Myoblasts' cluster met the expression cutoff described, with 26% of cells in this cluster expressing this gene. *Rarb* expression was preferentially seen in the 'Mesenchymal Cells 1'(53% of cells), 'Mesenchymal Cells 3' (51% of cells) and 'Mesenchymal Cells 5' (49% of cells) clusters. *Dhrs3* was expressed the highest in the 'Macrophages' (39% of cells), 'Smooth Muscle Cells' (34% of cells) and 'Mesenchymal Cells 1' (33% of cells) clusters. *Rdh10* was mostly expressed in the 'Mesenchymal Cells 1' (27% of cells), clusters. Violin plots of the aforementioned genes are shown in figures 3.9-3.11. Similar to our analysis of CDH-associated genes, expression of retinoid related genes was primarily seen in the mesenchymal and

mesothelial components of our PPF sample, highlighting the importance of RA signaling in the same cell populations linked to the pathogenesis of CDH.

	Percentage of cells expressing a gene (%)							
Cluster ID	*Dhrs3*	*Rdh10*	*Rbp1*	*Aldh1a2*	*Crabp1*	*Crabp2*	*Rara*	*Rarb*
Somatic Stem Cells 1	1.91	0.72	23.05	3.62	19.97	5.02	1.67	1.40
Mesenchymal Cells 1	33.33	27.56	79.70	30.29	32.88	35.93	17.96	52.53
Mesenchymal Cells 2	13.41	20.61	84.62	7.74	29.55	52.45	19.41	43.18
Mesenchymal Cells 3	22.59	15.13	67.98	7.79	18.42	36.73	15.13	50.77
Mesenchymal Cells 4	21.31	7.99	73.81	5.55	90.57	40.40	16.76	28.30
Erythroblasts 1	1.25	1.96	8.21	1.61	10.00	2.68	1.07	1.96
Epithelial Cells 1	3.04	0.70	9.35	2.80	10.51	1.87	1.17	0.93
Smooth Muscle Cells	34.83	4.50	63.51	0.95	24.64	20.14	17.30	15.64
Somatic Stem Cells 2	7.56	5.85	35.37	7.07	19.51	14.15	7.56	10.98
Mesothelial Cells	14.94	8.61	35.95	57.22	9.62	9.11	8.86	32.66
Fetal Cardiomyocytes 1	0.80	0.53	9.89	2.14	12.03	1.87	1.60	2.41
Endothelial Cells 1	4.48	4.20	54.34	2.52	10.64	1.12	11.76	13.73
Myoblasts	5.53	1.84	68.66	1.38	6.91	13.36	25.81	5.53
Erythroblasts 2	1.60	1.60	11.76	2.67	6.42	1.60	0.53	0.00
Macrophages	39.31	5.20	7.51	2.31	11.56	2.31	4.62	1.16
Mesenchymal Cells 5	12.18	7.69	78.85	11.54	25.00	5.13	11.54	48.72
Fetal Cardiomyocytes 2	1.50	3.76	12.78	1.50	15.04	1.50	10.53	18.05
Endothelial Cells 2	12.00	19.00	54.00	0.00	8.00	15.00	7.00	7.00

Table 3.2: Expression analysis for retinoid-associated genes in the developing diaphragm. Percentage of cells expressing the different retinoid associated genes per cluster. Only genes that show expression in more than 25% of cells for any cluster are shown in this Table.

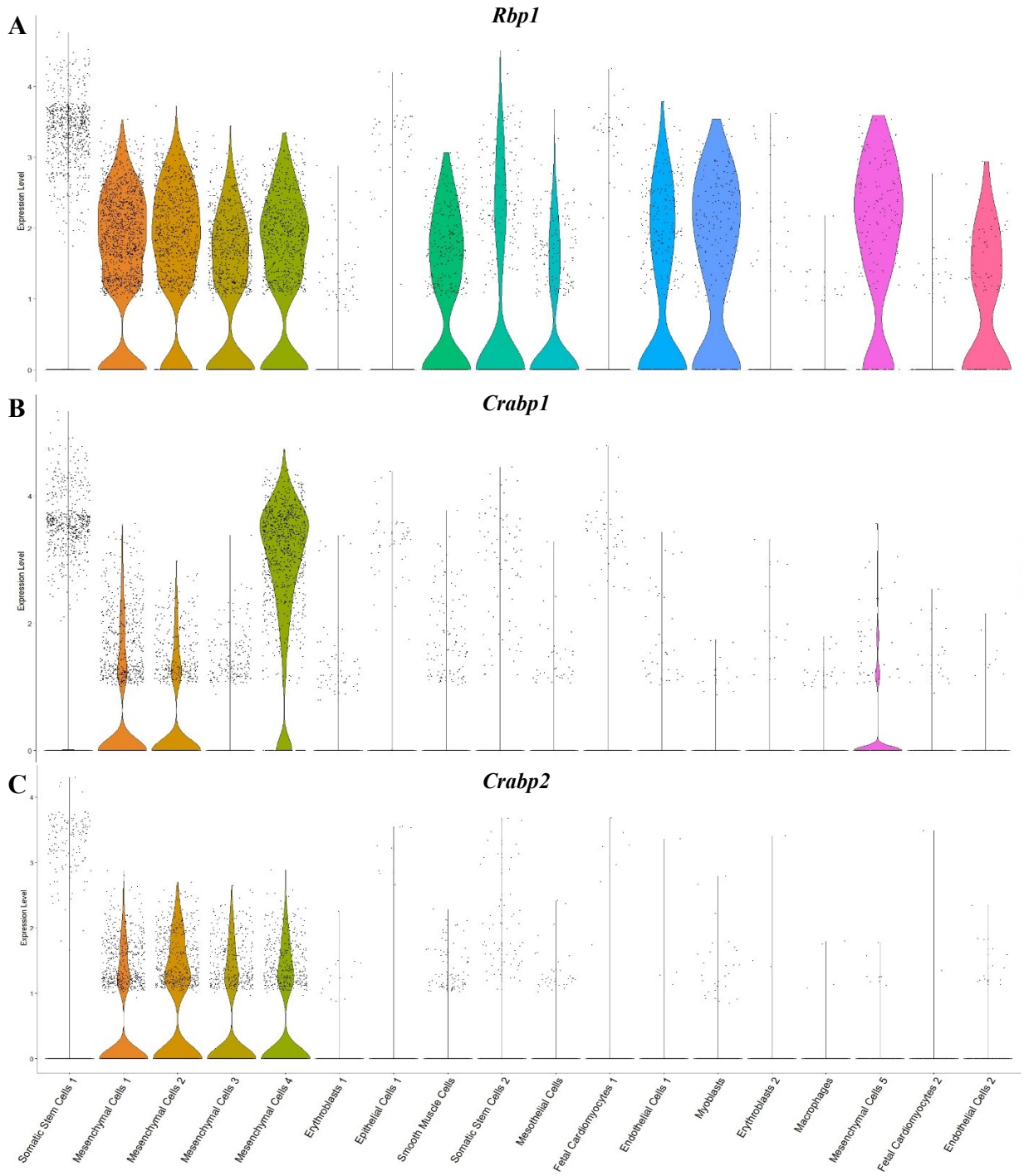


Figure 3.9: Genes of retinoid binding proteins show preferential expression in the mesenchymal and mesothelial components of the PPF. Violin plots for *Rbp1* (A), *Crabp1* (B) and *Crabp2* (C). Each dot represents a cell, whereas the width of the violin plot represents the frequency of the expression level per cluster.

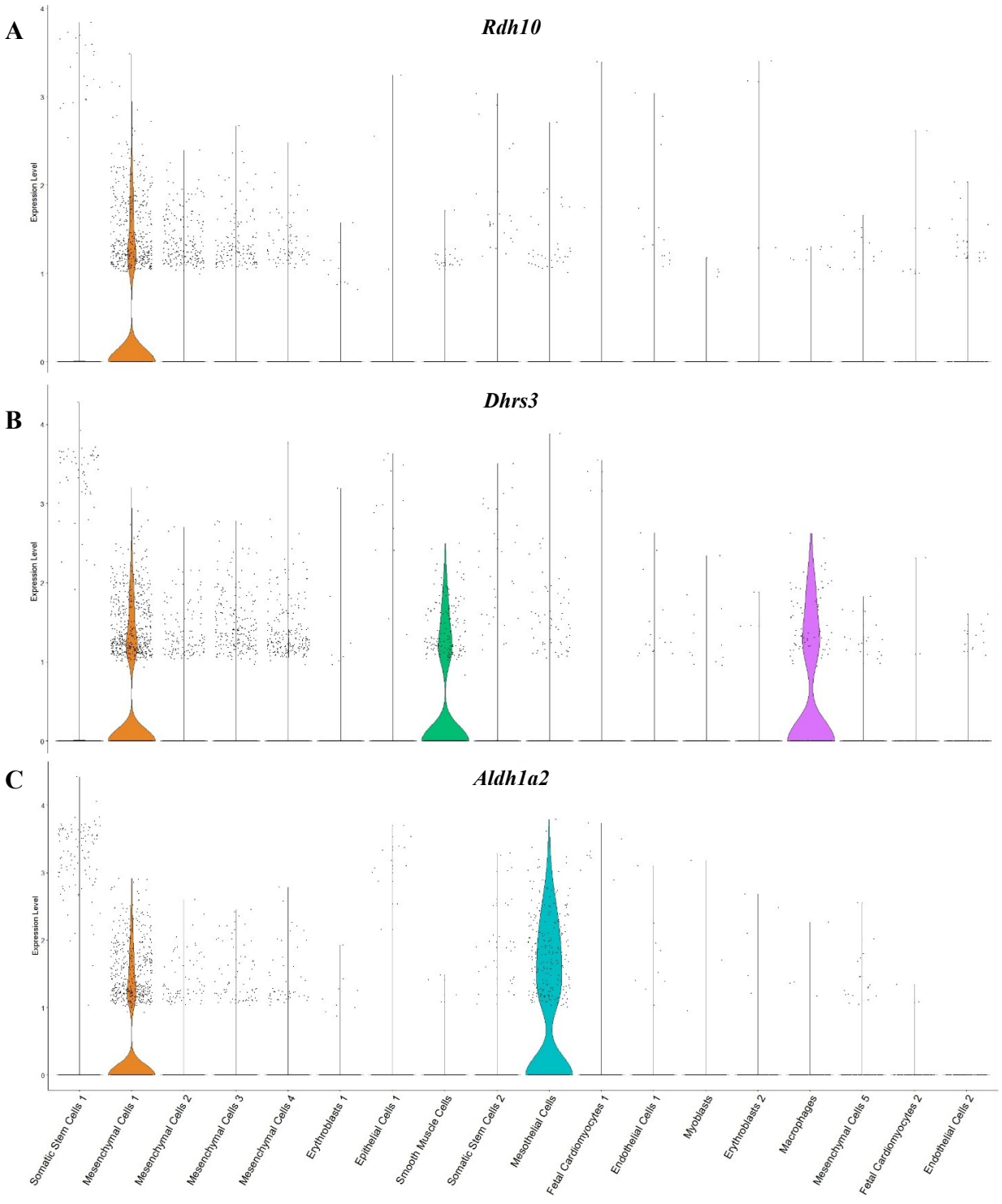


Figure 3.10: Genes of enzymes in the retinoid signaling pathway show preferential expression in the mesenchymal and mesothelial components of the PPF. Violin plots for *Rdh10* (A), *Dhrs3* (B) and *Aldh1a2* (C). Each dot represents a cell, whereas the width of the violin plot represents the frequency of the expression level per cluster.

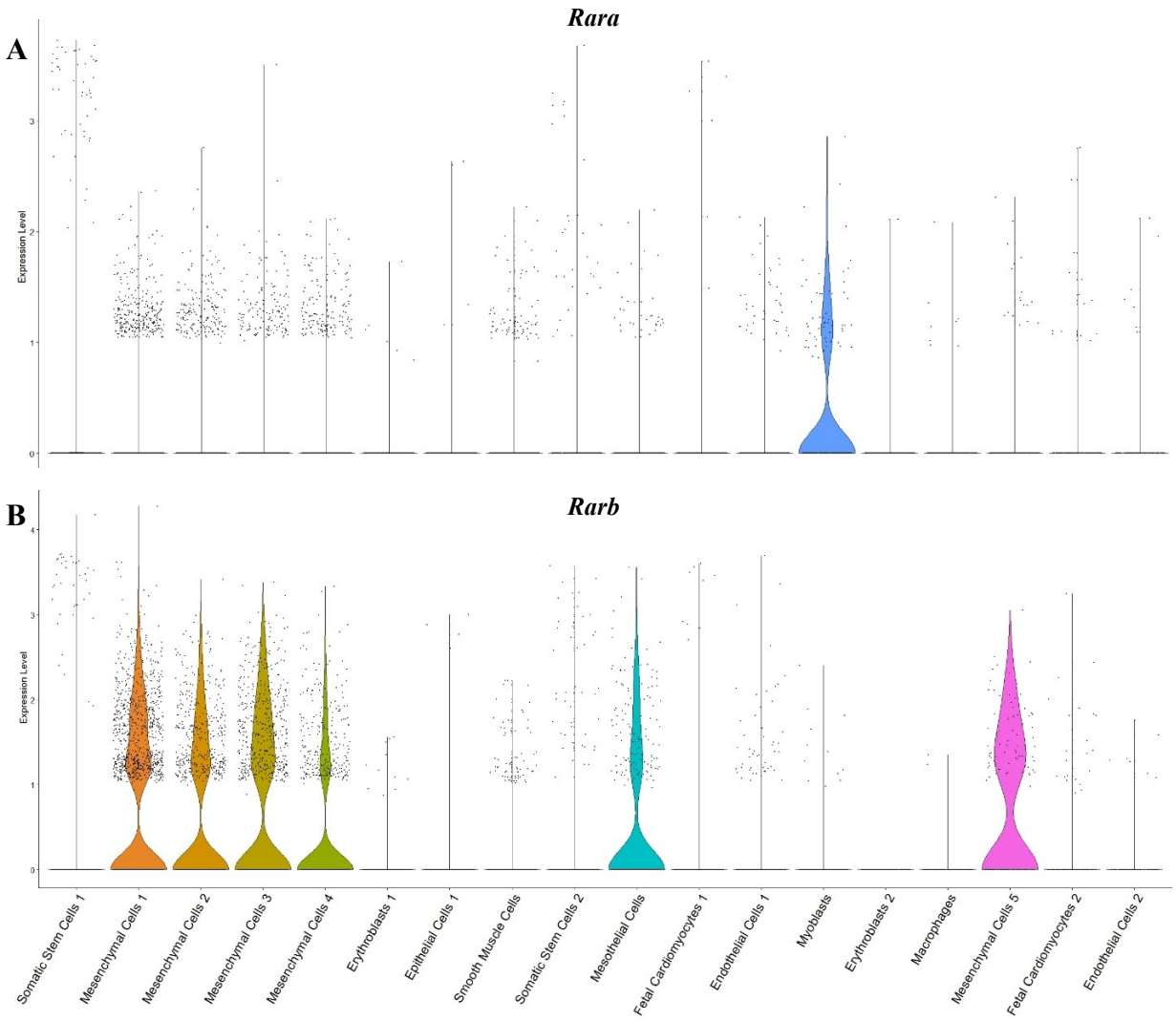


Figure 3.11: Genes of retinoic acid binding proteins show preferential expression in the mesenchymal and mesothelial components of the PPF. Violin plots for *Rara* (A) and *Rarb* (B). Each dot represents a cell, whereas the width of the violin plot represents the frequency of the expression level per cluster.

3.4 Discussion

To our knowledge, this is the first single cell transcriptomic analysis performed in the PPFs. Although transcriptomic analyses had been previously performed in both the PPF and diaphragm before (69), this microarray analysis was only able to determine the global levels of expression in these tissues, whereas our analysis was able to assess gene expression in individual cells.

Overall, we performed different clustering analyses in the PPF to determine the distinct cell populations present in this tissue. Seven distinct clusters were identified using the K-means analysis. This is due to the fact that this type of algorithm clusters the data set into K clusters, with K being a number that can be chosen by the user. Therefore, although K-means clustering is a valuable technique for low complexity models, the depth of analysis does not compare to other clustering algorithms, like UMAP, in samples with high complexity or a high number of cell populations (186,192). Although K-means clustering has its limitations, we were still able to identify the different major cell populations present in the developing diaphragm, including mesenchymal cells, muscle precursor cells, etc. Reclustering utilizing Seurat was performed to try and identify different subpopulations of these cell types, as well as to determine a wider range of cell populations than that possible with K-means analysis. A total of 18 clusters and 10 distinct cell populations were found. Some populations, like the mesenchymal cells, had multiple clusters. This might be due to the fact that the PPF is a developmentally active tissue, and some of these clusters might be in different developmental stages or have different functions once fully developed. Our alluvial plot shows that, even though the clustering algorithms performed gave us different results, most cells remained in similar clusters between our different analyses. The only clusters that show a wide array of difference between our clustering algorithms were the

‘Hematocytes’ and ‘Muscle Precursor Cells’ clusters in the K-means clustering. These two clusters were divided among many different clusters in the secondary clustering done in Seurat.

Based on the Seurat clustering, we were able to assess the levels of expression of both CDH- and retinoid-associated genes in the different cell populations, to determine what clusters might have a higher importance in the development of CDH. Although CDH-associated genes show expression in multiple clusters, the highest percentage of cells expressing Bochdalek CDH-associated genes was seen in the mesenchymal and mesothelial clusters. Genes like *Wt1, Gli2* and *Rarb* show expression in more than 50% of cells for at least one mesenchymal cluster. *Wt1* in specific also shows high levels of expression in the mesothelial cluster, with over 60% of cells in this cluster expressing this gene. This might indicate that these clusters are more important in the pathogenesis of CDH than the other clusters. This agrees with previous results in the literature, that have determined that the mesenchymal component of the PPF, which later gives rise to the central tendon and connective tissue, are the main drivers in the formation of CDH (38). It has also been shown that CDH can occur even in the absence of myogenic tissue (41), further implicating the importance of the non-muscular mesenchyme, not the muscle itself, in the pathogenesis of CDH. It is important to note that some genes show expression in clusters different from the mesenchymal and mesothelial clusters. *Nr2f2*, for example, shows the highest level of expression in non-mesenchymal clusters, like the ‘fetal cardiomyocytes’ and the ‘endothelial cells’ clusters. This might be due to the fact that there are different types of CDH, which might have different developmental origins, and some genes are associated with complex or multiple types of CDH. (46,47). Regarding the retinoid-associated genes, we found that the expression of these genes is enriched in the mesenchymal and mesothelial clusters, providing evidence that these clusters are the ones in charge of modulating the RA signaling cascade in the

PPF. As disruptions in the signaling cascade have been correlated with CDH, these results further sustain the idea that these cell populations are the main drivers in the formation of CDH.

Some genes in particular give better insight into the relationships between different cell populations and the pathogenesis of CDH. For example, *Aldh1a2* shows levels of expression almost exclusively in the mesothelial cluster, with over 50% of cells in this cluster expressing this gene. This gene is of interest for two main reasons; it is implicated in the oxidation of retinaldehyde into RA, and it has also been implicated in cases of human CDH when mutated (131,132). Furthermore, one of the top gene ontology results for this cluster was '*diaphragm development'*, highlighting the importance of this cluster in normal diaphragm morphogenesis, as well as how this cell population might be of importance in the development of diaphragmatic hernias. Other genes, like *Rara* and *Rarb*, which are both CDH and retinoid-associated genes (63,136), show highest expression in the mesenchymal clusters. Interestingly, there is a low level of coexpression of both *Aldh1a2* and *Rarb,* the main RAR isoform expressed in the PPF. This is of importance because *Aldha1a2* positive cells would be in charge of producing RA (33), which can act and induce signaling on *Rarb* positive cells. Only 9% of cells in all the mesenchymal clusters as well as the mesothelial cluster express both genes. The clusters with the highest level of coexpression were the 'Mesenchymal Cells 1' and 'Mesothelial Cells' clusters, with 18% of cells expressing both genes. This further support the idea that RA signaling during embryogenesis mostly occurs in a paracrine fashion based on expression gradients (201). Finally, some genes like *Wt1*, which is known to cause CDH when mutated, is a downstream target of the RA signaling cascade, while also providing feedback regulation by controlling the expression of *Aldh1a2* (42,95,184). These facts further strengthen the retinoid hypothesis of CDH, and further

explain the relationship between the retinoid signaling cascade and the formation of diaphragmatic hernias.

3.5 Conclusion and Future Directions

This first of a kind analysis was successful in determining the mosaic nature of the PPF, as there are many distinct cell populations in this tissue. By using multiple clustering analyses, we were able to gain a better insight into the distinct components of the PPF, while also determining how important these components are in the pathogenesis of CDH. Based on expression analysis, we found that the mesenchymal and mesothelial clusters show the highest level of expression of CDH- and retinoid-associated genes, and therefore might be the main drivers in the formation of diaphragmatic hernias. Since this is a transcriptomic analysis, there might be variations between our results and other studies that have measured protein expression in the PPF. For example, although both our results and the transcriptomic analysis performed by others show expression of *Rarb* in the PPF, protein expression analysis in this tissue was not able to detect expression of RARB (69,133). Therefore, further protein expression analysis in the PPF needs to be performed, to fully be able to assess the levels of expression of multiple CDH and retinoid- associated genes. Single-cell transcriptomic analysis of a PPF with CDH can also be performed, to see the differences in cell populations between our control sample and a CDH sample. It is expected that the mesenchymal and mesothelial clusters will show the highest differences between samples. Overall, as this is the first time that scRNA-seq was performed in the PPF, this study functions as a steppingstone to better understand this tissue, as well as the pathogenesis of CDH.

Chapter 4: A *Prrx1-Cre:Rar^{dn}* Animal Model Reveals the Importance of Retinoid Signaling in the Developing Diaphragm

4.1 Introduction

The diaphragm is a mammalian specific muscle with two main roles in an organism: serving as the primary muscle of respiration and providing a barrier between the abdominal and thoracic cavities (36). This muscle develops from multiple sources; including the somites, the septum transversum and the pleuroperitoneal folds (PPF); a transient structure that is crucial in the development of the diaphragm, as it helps with the migration of muscle precursor to form the fully developed diaphragm (70,179). In mice, the PPF first starts to appear at GD11.5, although it proliferates and fully migrates by GD13.5 (178). Multiple studies have shown that the non-muscular mesenchymal component of the PPF is the main driver in the formation of diaphragmatic defects, like Congenital Diaphragmatic Hernia (CDH) (41,202).

CDH is a life-threatening condition with high mortality that affects about 1 in 3,000 newborns (182). This defect is characterized by an incomplete formation of the diaphragm, as well as a high degree of lung hypoplasia (203). The rate of survival for this disease is close to 70%, although it is greatly dependent on severity (size of hernia and degree of lung hypoplasia), and complexity (presence of other abnormalities), among other factors (204). Multiple phenotypes of CDH can be seen depending on the location of the hernia. Bochdalek hernias, which are the most common type, develop in the postolateral side (205). Morgagni hernias develop in the anterior part of the diaphragm (205). Central tendon defects are also seen, although they are extremely rare (206). Finally, other phenotypes of this syndrome exist in which there is a weakening of the musculature of the diaphragm, known as muscular eventration (207).

Currently, there are two main explanations for the etiology of CDH: (i) genetic and (ii) environmental. About 40% of CDH cases have a specific genotype associated with it (69). For example, mutations in some specific genes like *WT1* and A*LDH1A2* have been identified in

patients with diaphragmatic hernias (69,208). The second explanation is related to the environmental conditions to which the mother and the fetus are exposed during development. Multiple factors, like maternal age, maternal smoking and fetal sex have been associated with a higher incidence of CDH (3,29,182). Additionally, multiple studies have shown that lower maternal vitamin A intake is associated with a higher incidence of diaphragmatic defects (81,82,133,209). A review performed in 2003 by Greer *et al.* led to the formulation of the retinoid hypothesis, which states that a disruption in the retinoid signaling pathway is involved in the incidence of CDH (31).

There have been multiple studies that have assessed the importance of the retinoid acid (RA) signaling cascade during the pathogenesis of CDH. For example, it has been shown that a teratogen used to induce CDH in animal models, Nitrofen, works by inhibiting ALDH1A2, the enzyme in charge of synthesizing retinoic acid from retinaldehyde (183,210). Furthermore, the family of Retinoic Acid Receptor (RAR), which modulate gene expression after binding RA, have been linked with the formation of diaphragmatic defects. Compound *Rar* mutant mice experiments demonstrated that, although singular knockouts of each RAR isoform do not cause CDH, compound knockouts of the different isoforms do (63,64). Specifically, compound knockouts of $Rara^{-/-}$:$Rarb^{-/-}$ show a low incidence of CDH, which is not seen in the $Rara^{-/-}$:$Rarg^{-/-}$ or $Rarb^{-/-}$:$Rarg^{-/-}$ compound mutants (63,64). Likewise, diaphragmatic defects were also seen in transgenic mice that lacked the functional domain of the *Rxra* receptor, showcasing the importance of this gene during diaphragm development (85).

Since retinoid signaling is of paramount importance during development, generating null mice for any gene in the retinoid signaling cascade can be challenging. Indeed, many mutant animal models of retinoid related genes are not viable, and therefore this can pose a challenge

when developing mouse models to study the role of RA signaling in CDH (88,89,211). Because of this, conditional mutant mouse models might provide an advantage in studying the effect that ablation of different genes have in the formation of the diaphragm. For example, multiple groups have shown that conditional mutations of *Wt1* in the non-muscular mesenchyme component of the PPF leads to the formation of diaphragmatic hernias in animal models (42,57,184). Keeping in mind that (i) the mesenchymal component of the PPF is hypothesized to be involved in the formation of CDH, and (ii) RA signaling is thought to be important in the formation of CDH, we hypothesize that a lack of signaling by the RARs in the mesenchymal component of the developing diaphragm will lead to the formation of CDH. To test this hypothesis, we generated a brand new conditional transgenic animal model that expresses a well-characterized RAR dominant negative mutant (Rar^{dn}), which is known to block RA signaling by disrupting the function of all three RARs, in the mesenchymal component of the diaphragm by using *Prrx1-Cre* mice.

4.2 Methods

4.2.1 Animal Care and Breeding

All experiments were approved by the University of Alberta Research Ethics Committee and were performed in accordance with the Canadian Council on Animal Care guidelines. The *Prrx1-Cre* mice were obtained from The Jackson Laboratory (Bar Harbor, ME, USA), and the Rar^{dn} mice were a generous gift from Dr. William S. Blaner (Columbia University in the city of New York). The characteristics of these mice have been detailed previously by Rosselot *et al.* (58). The animals were housed in the University of Alberta barrier animal facility. To generate timed pregnancies, one male *Prrx1-Cre* mice and two female Rar^{dn} mice were set up in a cage after 16:00 hrs and left overnight. The following morning at 9:00 hrs the presence of a vaginal

plug was examined to assess for copulation. Each cage was left for a maximum of two nights. After a vaginal plug was seen, noon of that day was considered gestational day (GD) 0.5.

4.2.2 Genotyping

To genotype the *Prrx1-Cre:Rar^{dn}* fetuses, DNA was extracted from the fetal tails and PCR reactions were performed utilizing *Cre* allele primers (Forward: 5'-ACC TGA AGA TGT TCG CGA TTA TCT-3', Reverse: 5'-ACC GTC AGT ACG TGA GAT ATC TT-3') (Invitrogen, Toronto, ON), as well as the *Rar^{dn}* allele primers (Forward: 5'-ATG GTG TAC ACG TGT CAC C-3', Reverse: 5'-CAC CTT CTC AAT GAG CTC C-3') (Invitrogen, Toronto, ON). Sex determination of the embryos was also performed utilizing *Sry* primers (Forward: 5'-CGT GGT GAG AGG CAC AAG T-3', Reverse: 5'-AAG GCT TTT CCA CCT GCA TCC CA-3'). PCR reactions were set by adding 2 µL of DNA, 0.5 µL of each primer, as well as 21.5 µL of PCR super-mix (Thermo Fisher Scientific, Toronto, ON). The ProFlex PCR system (Applied Biosystems, Toronto, ON) was used for the reactions. Gel electrophoresis was performed utilizing a 1.2% agarose gel, which was ran at 100 V for 90 minutes for both the *Cre* and *Rar^{dn}* primers. For *Sry* a 2% agarose gel was utilized, and gel electrophoresis was ran at 100 V for 120 minutes. The PCR products were visualized utilizing RedSafe dye (Intron Biotechnology, Seoul, KR), as well as the ChemiDoc Touch Imaging System (BioRad, Mississauga, ON). The *Cre* allele PCR product had a size of 374 bp, and the *Rar^{dn}* allele had a product of 204 bp. A representative agarose gel used for genotyping can be seen in Figure 4.1. The *Sry* allele PCR product had a size of 106 bp.

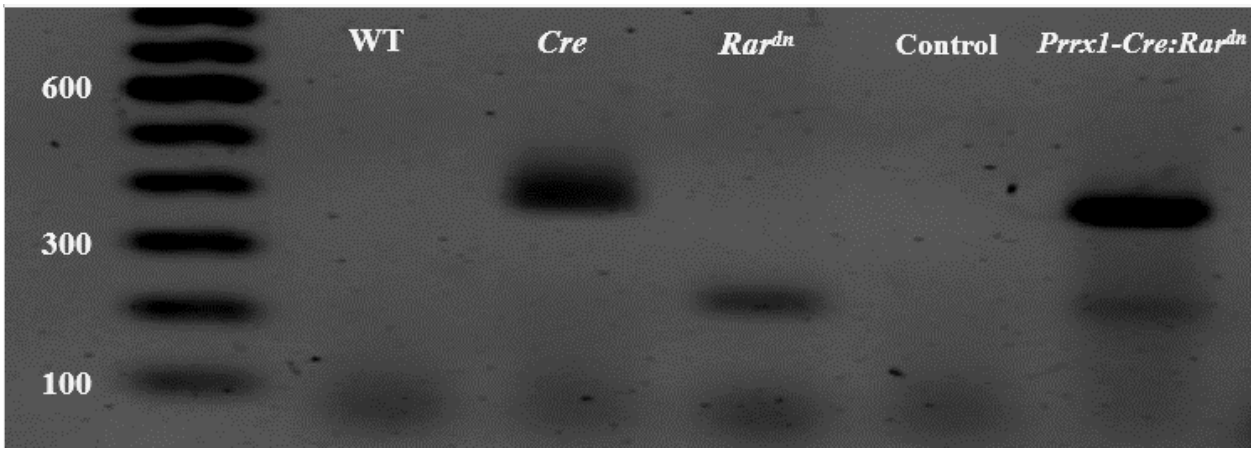


Figure 4.1: PCR Genotyping of the different fetuses. A representative agarose gel used for genotyping the different fetuses is shown. Wild type (WT) fetuses show no bands, and single transgenic mice, which carry only one of the alleles, only present one bad, either at 374 bp for Cre^{+} mice, or 204 bp for Rar^{dn} mice. Genotyping for *Prrx1-Cre:*Rar^{dn} show two distinct bands, at 374 bp and 204 bp respectively.

4.2.3 Dam Dissection and Fetus Collection

Dams at GD 13.5-15.5 were euthanized via anesthesia using isoflurane inhalation immediately followed by cervical dislocation. An abdominal incision was made in the abdomen, and the entire uterus was removed. Both the number of fetuses and resorptions were recorded. The uterus was placed in a petri dish containing ice cold Phosphate Buffer Saline (PBS) (Sigma-Aldrich, Oakville, ON). To dissect the individual fetuses, stereomicroscopy was used (Stemi 508, Zeiss, Oberkochen, Germany). The Crown-Rump Length (CRL) of each fetus was recorded. To obtain the diaphragm, both the cranial and caudal section of the fetus were dissected, and the heart and lungs were removed. Afterwards, the diaphragm was gently collected from atop the liver to assess for the incidence of diaphragmatic defects. The diaphragms were stored in formalin for later analysis. The degree of herniation was graded using letters, ranging from A-D, following guidelines described previously in the literature (212). A schematic of each herniation type can be seen in Figure 4.2. A heat map of the herniations was made utilizing Heatmapper (213).

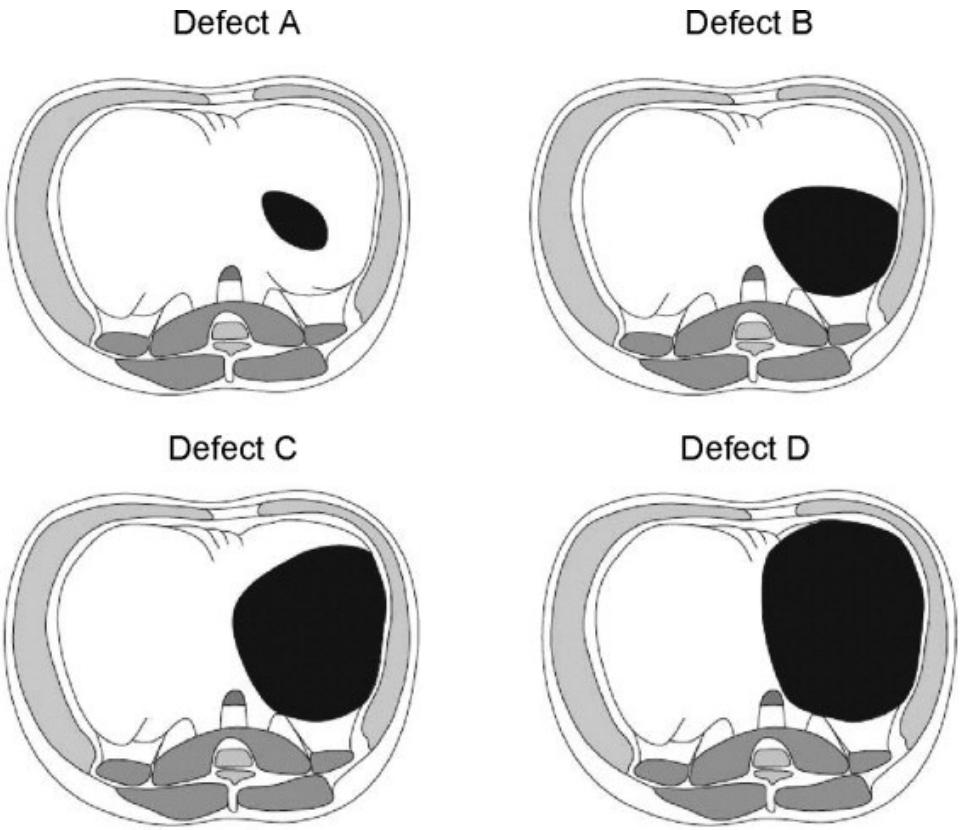


Figure 4.2: Schematic representation of different hernia sizes based on intraoperative findings of postolateral hernias. Type A refers to a small hernia that does not touch the body wall. Type B refers to a hernia that is less than 50% of one side of the diaphragm and it touches the body wall. Type C is a hernia that is more than 50% of one side of the diaphragm and it touches the body wall. Type D refers to full herniation in one side of the diaphragm. Figure reproduced from Tsao and Lally, 2008 (212).

4.2.4 Histology Analysis

Fetuses at GD 13.5 were taken to the HistoCore at the University of Alberta for processing, embedding, and H&E staining. The histology sections were visualized with a Axio Zeiss Vert A1 inverted light microscope (Zeiss). Images were acquired with a Zeiss Axiocam 208 color camera (Zeiss)

4.2.5 Contrast enhanced micro-Computed Tomography (Micro-CT)

Fetuses stored in formalin were rinsed with deionized water for 24 hours followed by immersion in a 5% Mercury Chloride (Acros organic, Toronto) solution for 3 days. Scans were performed on a Bruker Skyscan 1176 Micro-CT with the following scanning parameters: 90 kV X-ray tube voltage, 278 μA X-ray current, 1 mm Al filter, 300 ms integration time, frame averaging with n=3, and an angular rotation step of 0.7 degrees. Scans were visualized utilizing data viewer (Bruker), and the long volume was calculated utilizing the 3D analysis function in CTan (Bruker). The 3D renders were made utilizing Ctvox (Bruker). Statistical analysis of the volumes was performed utilizing Prism. To determine the significance between the lung volumes, student t-tests were performed, where a P value < 0.05 was considered significant.

4.2.6 scRNA-seq expression analysis

Expression data for *Prrx1* as well as the different RAR isoforms is based on the single cell transcriptomic data presented in the earlier chapter. The UMAPs and expression data were obtained following the methods previously described in chapter 3.

4.3 Results

4.3.1 Expression of Prrx1 and the Retinoic acid receptors in the developing diaphragm,

To generate our conditional transgenic mouse model, *Prrx1-Cre* mice were chosen due to a variety of different reasons. First, *Prrx1* has been shown to be expressed in the non-muscular mesenchyme of the PPF. *Prrx1-Cre* mice crossed with Cre-responsive reporter $Rosa^{LacZ}$ show labelling in the PPF but not other transient structures like the septum transversum (38). Furthermore, *Prrx1* labelling also shows that this gene is still expressed in structures that are derived from the PPF, like diaphragmic fibroblasts as well as the central tendon (38). Second, *Prrx1-Cre* mice have been previously used to generate a *Wt1* conditional knockout in the developing diaphragm that develops CDH (57). Third, our scRNA-seq data shows that expression of *Prrx1* is primarily seen in the mesenchymal component of the PPF. Similarly, expression of the RARs, in particular *Rara* and *Rarb* is also seen in the same cell populations. scRNA-seq expression data for both *Prrx1* as well as the different RARs can be seen in Figure 4.3.

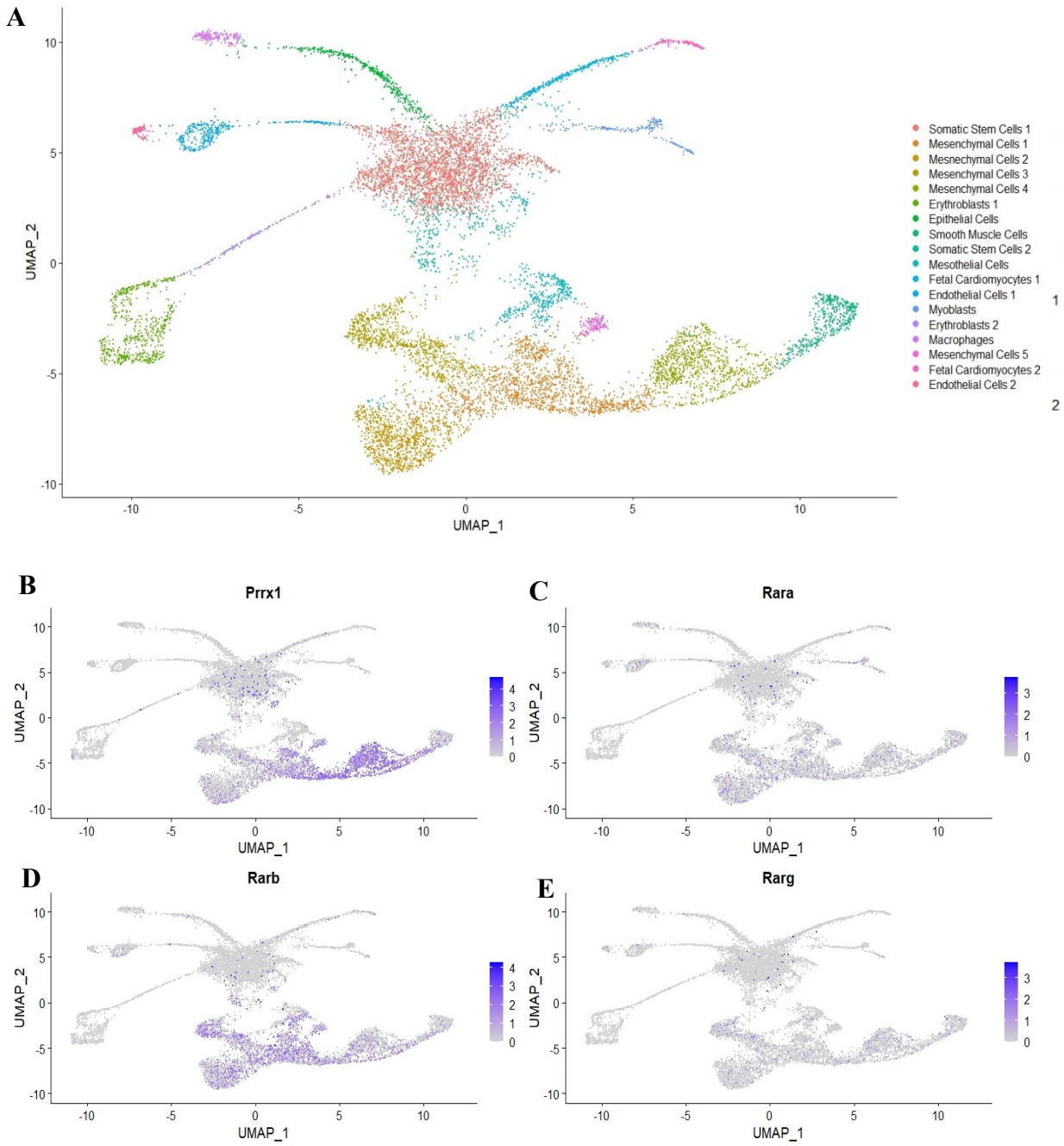


Figure 4.3: Single cell expression data for *Prrx1* and the different RAR isoforms in the developing diaphragm. The UMAP in A shows the different cell populations present in the developing diaphragm, with each cluster being represented by a different color based on the legend. B-E show feature plots for *Prrx1* (B), *Rara* (C), *Rarb* (D) and *Rarg* (E). Cells colored in blue express the different genes, where a darker blue represents higher levels of expression.

As seen in Figure 4.3 expression for both *Prrx1* as well as the three RAR isoforms is primarily confined to the mesenchymal clusters of the PPF. The three different RAR isoforms did not have similar expression levels. Based on the feature plots, *Rarb* was the main isoform expressed in the PPF, followed by *Rara.* The last isoform, *Rarg*, had low to no levels of expression in all cell populations. It is important to note the percentage of cells expressing each specific gene per cluster. In the case of *Prrx1*, clusters 'Mesenchymal Cells 1' and 'Mesenchymal Cells 4' had the highest percentage of cells expressing this gene, with values of 57% and 83% respectively. For these clusters, the percentage of cells expressing *Rara* was 17% and 16% respectively, 52% and 28% respectively for *Rarb*, and 9% and 14% respectively for *Rarg*. The percentage of cells expressing *Prrx1* as well as the respective RAR isoforms per cluster can be seen in Table 4.1. Furthermore, the patterns of expression for clusters in which any of the genes are expressed in more than 25% of the cells is shown in Figure 4.4. Finally, analysis of the scRNA-seq data was performed to determine coexpression of *Prrx1* with each of the different RAR isoforms. There are low levels of coexpression of *Prrx1* with any of the isoforms. Out of the total 4823 cells present in the mesenchymal and mesothelial clusters, 19% express both *Prrx1* and *Rarb*, 8% express both *Prrx1* and *Rara*, and only 5% express both *Prrx1* and *Rarg.*

	Prrx1	***Rara***	***Rarb***	***Rarg***
Somatic Stem Cells 1	5.29	1.67	1.40	0.89
Mesenchymal Cells 1	57.52	17.96	52.53	9.14
Mesenchymal Cells 2	31.41	19.41	43.18	8.29
Mesenchymal Cells 3	21.71	15.13	50.77	15.79
Mesenchymal Cells 4	83.35	16.76	28.30	14.10
Erythroblasts 1	2.50	1.07	1.96	0.36
Epithelial Cells 1	0.70	1.17	0.93	0.70
Smooth Muscle Cells	43.60	17.30	15.64	9.95
Somatic Stem Cells 2	16.34	7.56	10.98	1.71
Mesothelial Cells	6.58	8.86	32.66	4.56
Fetal Cardiomyocytes 1	3.48	1.60	2.41	0.53
Endothelial Cells 1	3.36	11.76	13.73	3.64
Myoblasts	3.23	25.81	5.53	4.15
Erythroblasts 2	1.60	0.53	0.00	0.53
Macrophages	4.62	4.62	1.16	3.47
Mesenchymal Cells 5	53.21	11.54	48.72	5.13
Fetal Cardiomyocytes 2	3.01	10.53	18.05	0.00
Endothelial Cells 2	7	7	7	11

Table 4.1: Expression of *Prrx1* and the RAR in the developing diaphragm. Cells expressing *Prrx1* and the RARs per cluster. The highest level of expression for *Prrx1* was seen in the 'Mesenchymal Cells 4' cluster, where 83% of cells expressed this gene. For *Rara* the highest expression was seen in the 'Myoblast; cluster, with 25% of cells expressing this isoform. The 'Mesenchymal Cells 1' cluster had a total of 52% of cells expressing *Rarb*, which was the highest amongst all clusters. Finally, expression of *Rarg* was the highest in the 'Mesenchymal Cells 3' cluster, with a total of 15% of cells expressing this gene.

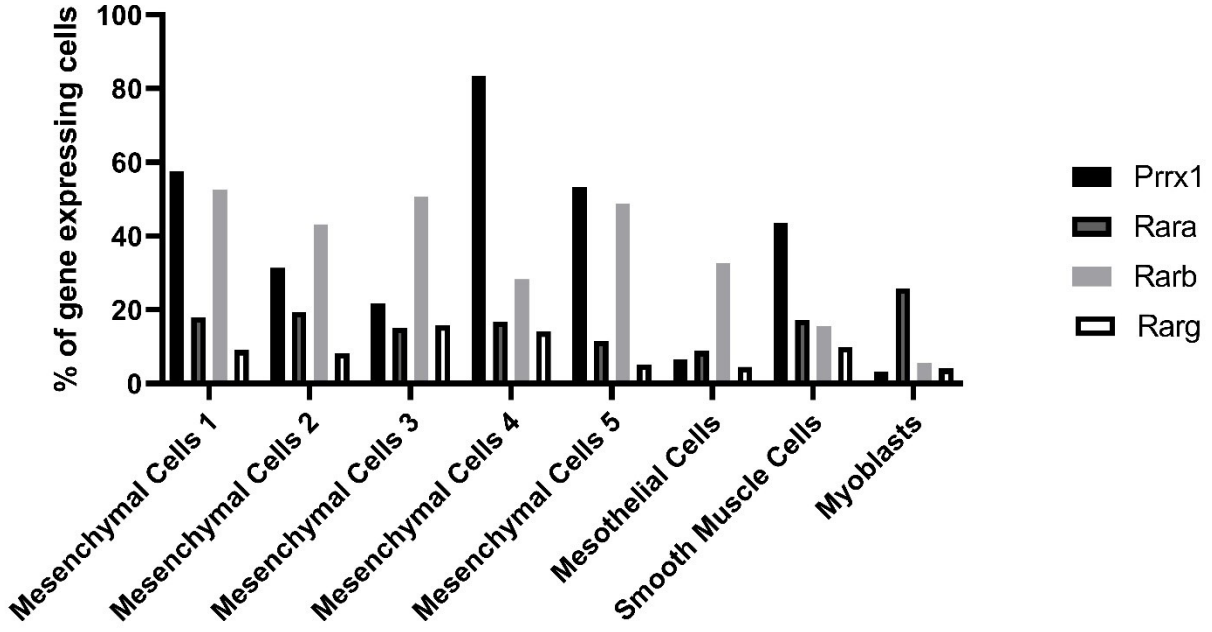


Figure 4.4: *Prrx1* and RAR isoform expression is enriched in the mesenchymal component of the developing diaphragm. Expression of *Prrx1* or the RAR isoforms is above 25% in eight distinct clusters, with the highest level of expression seen in the mesenchymal cell clusters.

4.3.2 Prrx1-Cre:Rar^{dn} mice develop CDH

Both microscopy as well as Micro-CT analyses reveal that the *Prrx1-Cre:Rar^{dn}* fetuses develop diaphragmatic hernias. A total of 17 *Prrx1-Cre:Rar^{dn}* double transgenic fetuses from 7 separate litters were analyzed, with all of them presenting diaphragm defects, meaning our model had 100% penetrance of CDH. None of the 44 control fetuses, which were either WT or had only one of the alleles, had CDH. The distribution of the herniations, as well as the type of herniations seen can be seen in Figure 4.5. The general characteristics of the fetuses can also be seen in Table 4.2. Overall, hernias in the posterior left side of the diaphragm were more common, as roughly half of the *Prrx1-Cre:Rar^{dn}* fetuses had left hernias. This was followed by bilateral hernias, with an incidence of 35%. Finally, although right sided hernias were seen, they were rare, with an incidence of ~15%. Most of the hernias seen in the *Prrx1-Cre:Rar^{dn}* fetuses were either grade B, meaning that less than 50% of one side of the diaphragm was affected, or grade C, meaning that more than 50% of that side of the diaphragm was affected. A total of 7 grade B and 11 grade C hernias were seen. Some Grade D hernias, which represent full herniation in one side of the diaphragm, were seen, although they were not as common, as only 25% of this type of herniation. Interestingly, in bilateral hernias, the left side of the diaphragm always had an equal or higher degree of herniation than the right side. It is important to note that in all of the hernias seen in the *Prrx1-Cre:Rar^{dn}* fetuses, the liver was pushing into the thoracic cavity. A heat map representing the areas of the diaphragm most can be seen in Figure 4.6D. Representative examples of the different types of herniations seen in our animal model can be seen in Figure 4.6A-C. Finally, although most of the fetuses presented actual diaphragmatic herniation, one of the *Prrx1-Cre:Rar^{dn}* fetuses showed an eventration, which speaks to the wide array of diaphragmatic defects that are co-described as CDH.

The embryological origins of CDH have been traced back to abnormalities in the PPF (38,179). Our histology analysis revealed that the diaphragmatic defects in these fetuses can be seen early in development, at GD13.5. Compared to the control fetuses, in the *Prrx1-Cre:Rar^{dn}* fetuses the PPF is almost completely absent, and this could explain why diaphragm development is disrupted in this mouse model of CDH. Histology analysis for both the WT as well as the *Prrx1-Cre:Rar^{dn}* fetuses can be seen in Figure 4.6G-H.

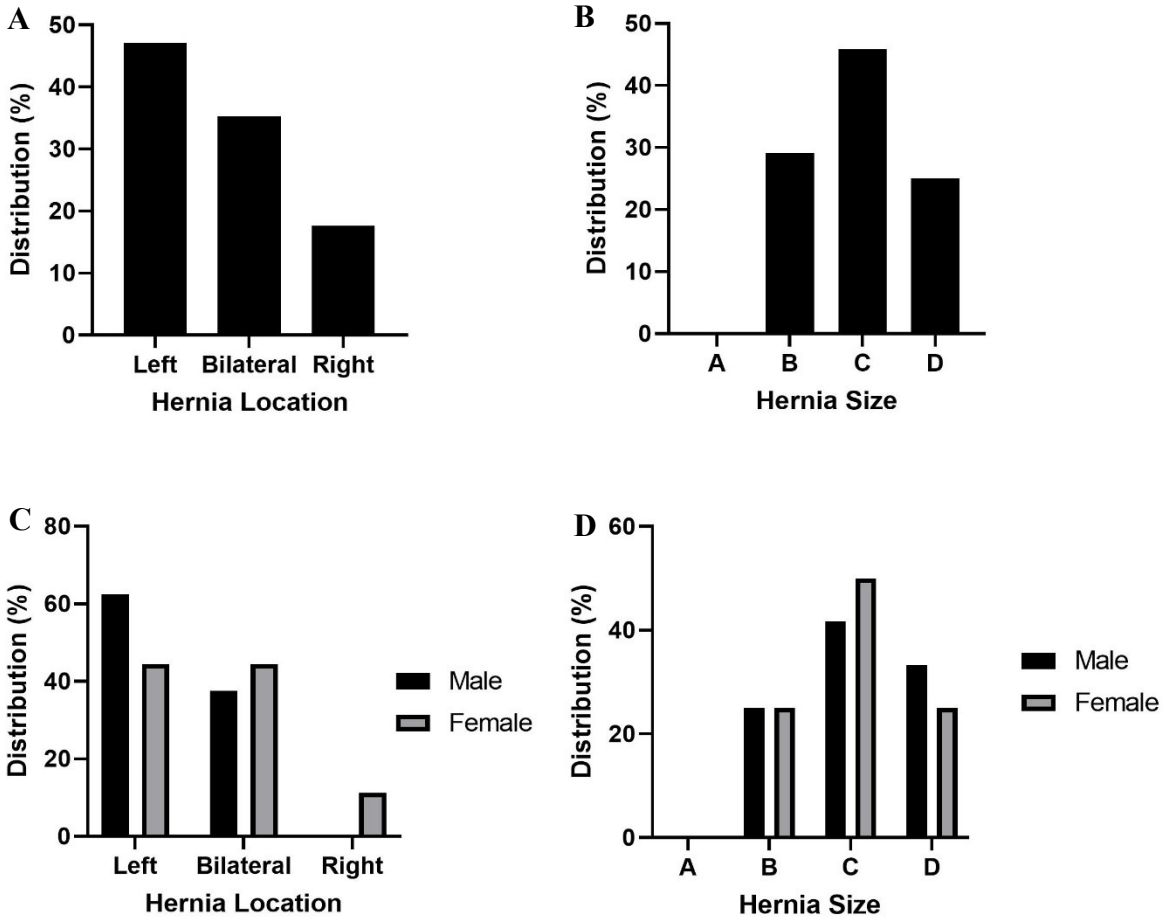


Figure 4.5: *Prrx1-Cre:Rar^{dn}* develop CDH. Bar graphs showing the location (A; n = 17) as well as the type of herniation (B; n = 24) seen in the *Prrx1-Cre:Rar^{dn}* fetuses. Fetuses with bilateral hernias were given two different scores for each side of the diaphragm. In C (n = 8 for males; n = 7 for females) and D (n = 12 for males and females), the location and the type of herniation of the hernias are divided by sex.

			Genotype			
Litter No.	**Maternal Age (days)**	**No. of fetuses**	***WT***	***Cre***	***Rar^{dn}***	***Prrx1-Cre:Rar^{dn}***
7	85.5±30	63	20	8	18	17

Table 4.2: General characteristics of the fetuses analyzed. Litter ID, maternal age, *Cre* and *Rar^{dn}* genotyping are all shown in the Table.

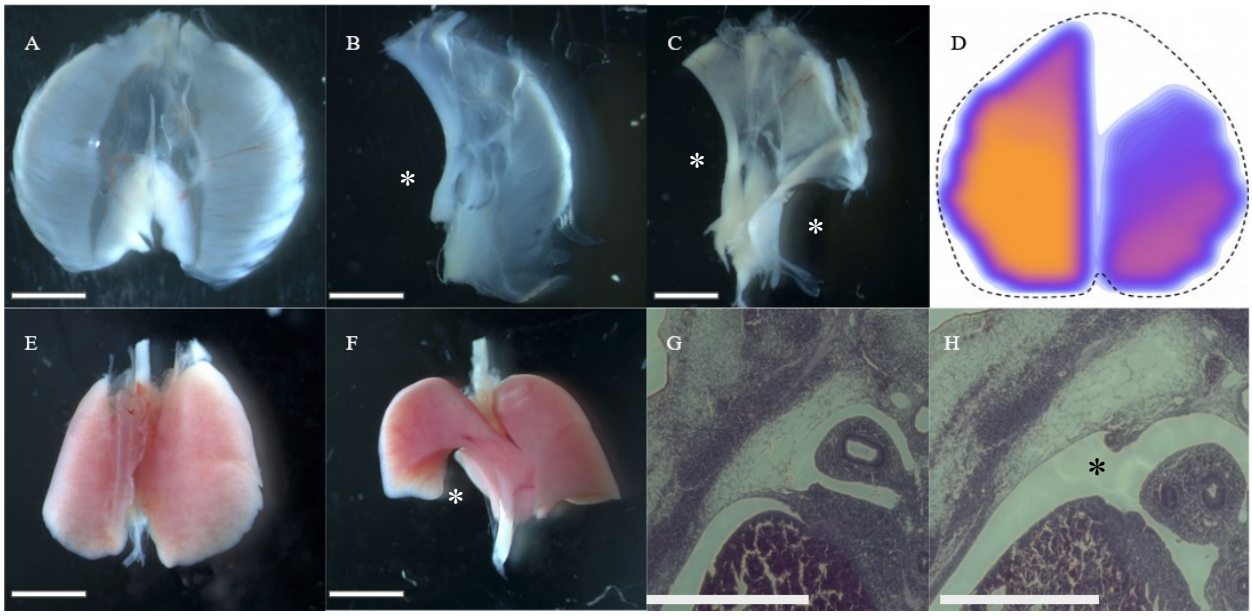


Figure 4.6: *Prrx1-Cre:Rar^{dn}* mice present CDH as well as lung hypoplasia. A shows a diaphragm of a WT fetus, B and C show the diaphragms of *Prrx1-Cre:Rar^{dn}* fetuses, where the herniations is shown with an asterisk. The diaphragm shown in B has a grade C left hernia. The diaphragm in C has a grade B hernia on the right side, and a grade C hernia on the left side. The heatmap showing the distribution of all hernias found in our animal model is seen in D. Areas in orange had a higher incidence of CDH than areas in blue. E & F show the lungs of a WT fetus and a *Prrx1-Cre:Rar^{dn}* fetus respectively. Hypoplasticity is shown with an asterisk in F. G & H show histology for a WT fetus and a *Prrx1-Cre: Rar^{dn}* fetus respectively. The lack of the PPF is shown with an asterisk. All scale bars are 0.5 mm.

4.3.3 Micro-CT analysis reveals lung hypoplasia in the Prrx1-Cre:Rardn fetuses

Micro-CT analysis was also successful in allowing us to see the anatomical differences in the liver, lungs and diaphragms of the *Prrx1-Cre:Rar^{dn}* fetuses. As seen in Figure 4.7, both the sagittal, coronal, and transverse views of the *Prrx1-Cre:Rar^{dn}* fetus show anatomical differences normally seen in CDH compared to the control. For the specific *Prrx1-Cre:Rar^{dn}* fetus shown, micro-CT clearly shows that the herniation is only present in the left side of the diaphragm. This was later confirmed via dissection which can be seen in Figure 4.8. Furthermore, visualization via micro-CT allowed us to observe the liver protruding into the thoracic cavity, compressing the ipsilateral lung. The degree of lung hypoplasia, as well as the difference in liver position can also be seen in the 3D-renders presented in Figure 4.9A-B. Compared to the control fetuses, it can be seen that in the *Prrx1-Cre:Rar^{dn}* fetuses the liver is elevated, as it is herniating into the thoracic cavity and causing lung hypoplasia. This liver elevation is only seen in the left side of the fetus, since the right side of the diaphragm was intact.

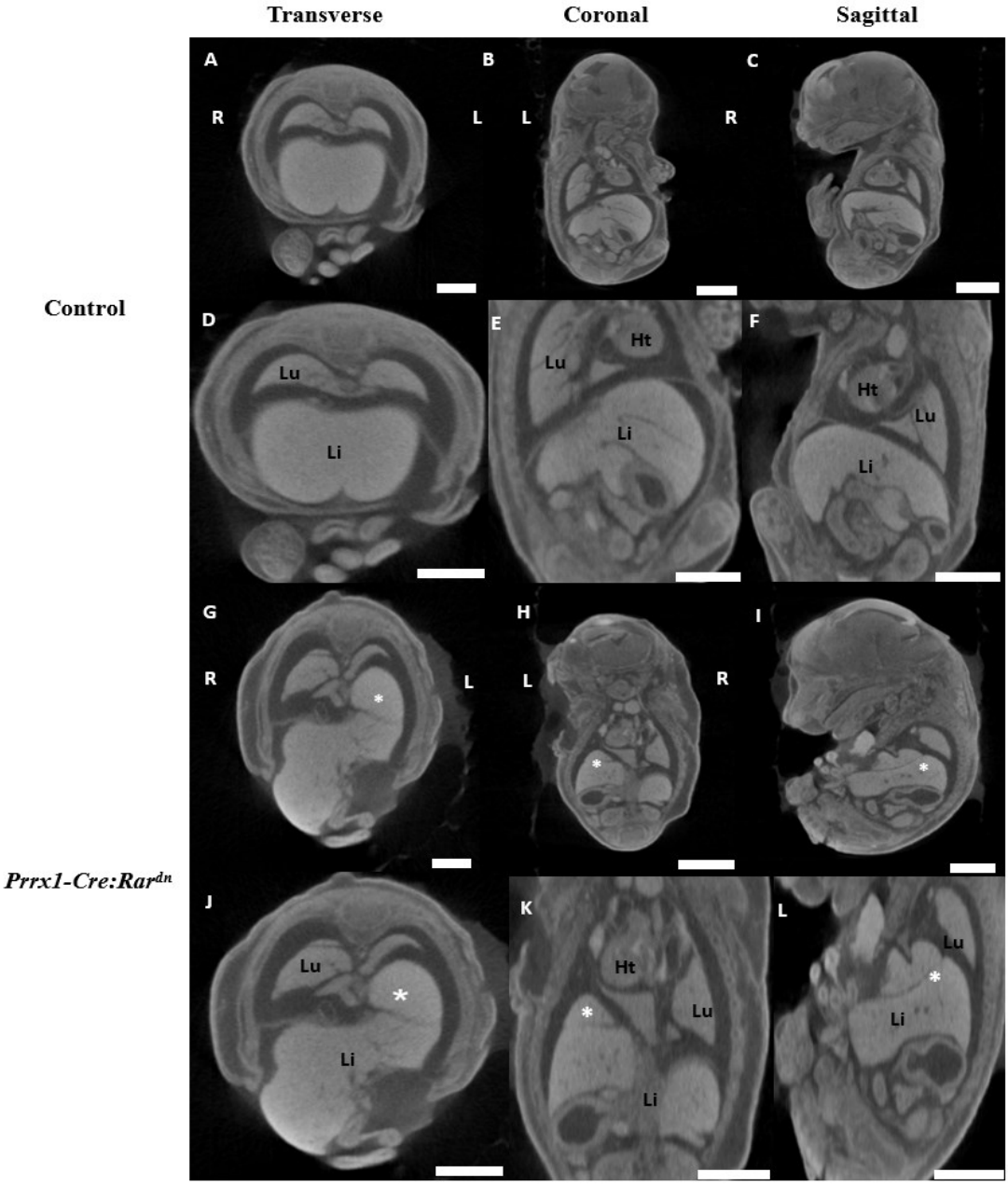


Figure 4.7: Micro-CT scans show CDH and lung hypoplasia in the different anatomical planes. A-C shows a control fetus in the transverse, coronal and sagittal sections respectively. D-E show enhanced versions of the images. G-I show a *Prrx1-Cre:Rar*^{dn} fetus in the transverse, coronal and sagittal planes, with J-L showing zoomed-in versions of the images. The herniation as well as the liver pushing into the thoracic cavity are represented by asterisks. Scale bars are 1 mm. For the abbreviations, Li = Livers, Lu = Lungs, He = Heart, L = Left, R = Right.

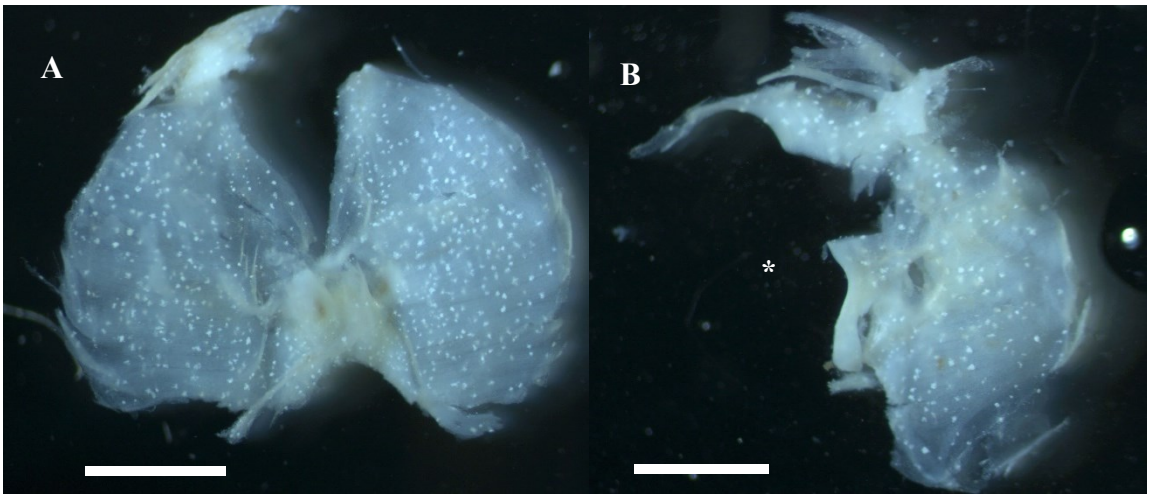


Figure 4.8: Diaphragm dissections from the fetuses confirm micro-CT analysis. The diaphragm from the control (A) as well as the *Prrx1-Cre:Rar*^{dn} (B) fetuses show the presence of a left side hernia in the *Prrx1-Cre:Rar*^{dn} fetus as seen in via micro-CT. CDH in B is shown with an asterisk. Scale bars are 0.5 mm. The white precipitates are due to HgCl staining.

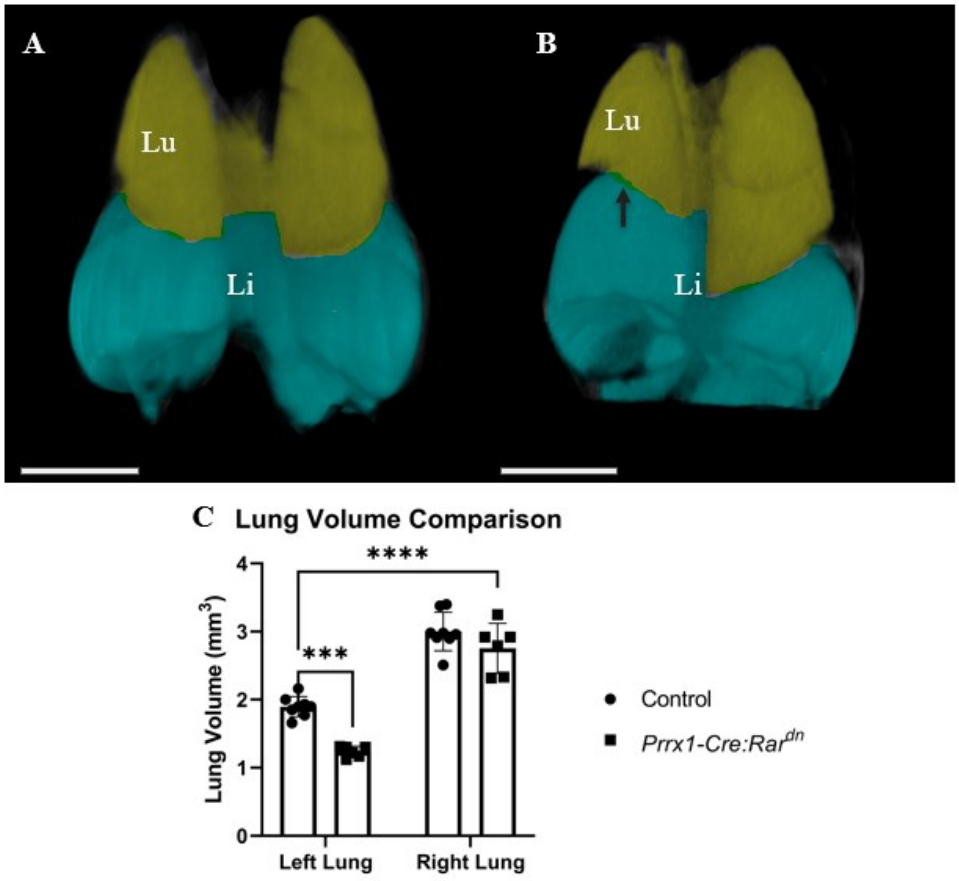


Figure 4.9: *Prrx1-Cre:Rar^{dn}* mice have lung hypoplasia. A & B show 3-D renders of a control and a *Prrx1-Cre:Rar^{dn}* fetus respectively. Scale bars are 0.5 mm. The elevated position of the liver in the *Prrx1-Cre:Rar^{dn}* fetus is shown with an arrow. C shows the difference in lung volumes between the WT and the *Prrx1-Cre:Rar^{dn}* fetuses. Data analyzed by student t-test, ****$p < 0.0001$, *** $p < 0.001$. For the annotations, Lu = lungs and Li = liver.

Visualization of both control and *Prrx1-Cre:Rar^{dn}* fetuses via micro-CT allowed us to analyze if the *Prrx1-Cre:Rar^{dn}* fetuses present lung hypoplasia, by measuring the difference in the lung volumes compared to the control fetuses. As seen in Figure 4.9C the left lungs of the *Prrx1-Cre:Rar^{dn}* fetuses present a decrease in volume of 34% compared to the controls. The average volume of the left lungs of the *Prrx1-Cre:Rar^{dn}* fetuses was 1.23 mm^3, whereas the average left lung volume of the control fetuses was 1.90 mm^3 (student T-test; $p < 0.001$). We were not able to demonstrate a statistically significant change in volume in the right side of the lungs between the *Prrx1-Cre:Rar^{dn}* and the control fetuses, as the average right lung volume in the *Prrx1-Cre:Rar^{dn}* was 2.75 mm^3, compared to 3.00 mm^3 in the controls.

4.4 Discussion

By generating a conditional mutant mouse model that blocks signaling of the RARs, we were able to target the mesenchymal component of the diaphragm and generate a *Prrx1-Cre:Rar^{dn}* mouse model that develops CDH with 100% penetrance. This shows that RA signaling in the mesenchymal component of the developing diaphragm is essential for proper diaphragm development, and that disruptions in the RA signaling cascade will cause CDH. As CDH was seen in all of the *Prrx1-Cre:Rar^{dn}* fetuses, this model might provide a powerful tool for studying the development of diaphragmatic defects that does not have the limitations that teratogenic/pharmacological models have, like off-target effects, timing of dose, and low penetrance (214,215). It is important to note that this model does not present any co-morbidities that are usually seen when disrupting the RA signaling cascade, like facial clefting, limb deformities, cranial malformation, etc. (216–218). The only comorbidity seen was oedema, a

common comorbidity associated with CDH, thought to be due to obstructions in venous return, increasing systemic volume pressure (219,220).

The *Prrx1-Cre:Rar^{dn}* fetuses develop CDH in the posterior left side of the diaphragm in the majority of the cases, correlating to what is seen in patients with Bochdalek herniation, the most common type of CDH (3). This relatedness with humans further highlights the importance that this mice model has as a tool for studying diaphragmatic development.

The micro-CT analysis was successful in showing diaphragmatic hernias in the *Prrx1-Cre:Rar^{dn}* fetuses. Although this technique is normally used to visualize mineralized tissue, like bones, we further demonstrate that HgCl staining of the fetuses prior to imaging can be useful when trying to visualize soft tissues, like the fetal lung, liver and diaphragm (221). However, one drawback of this type of staining is the presence of precipitates forming in the tissue, as seen in our diaphragm dissections. Through this technique, we were able to see CDH, liver herniation and lung hypoplasia. We were successful in demonstrating that the *Prrx1-Cre:Rar^{dn}* fetuses have lung hypoplasia compared to the control, albeit only in the left lungs. This might be because most of our herniations were on the left side. However, it is worth mentioning that there have been previous studies that show that lung hypoplasia can occur in the contralateral lung of the herniation, something which we were not able to demonstrate through this analysis (222–224). It is important to note that there is a statistically significant difference between the volumes of the left lungs compared to the right lung volumes, and this is seen in both the *Prrx1-Cre:Rar^{dn}* as well as the control fetuses, although this is expected, as the right lung is naturally larger than the left lung (225). Although unanswered questions remain, it is clear that micro-CT provides a powerful tool for analyzing the multifaceted effects of CDH in a fetus.

Our *Prrx1-Cre:Rar*^{dn} mouse model demonstrates the importance of RA signaling in the formation of CDH, further strengthening the retinoids hypothesis, which states that retinoids have a role in the incidence of CDH (31). Blocking of RAR signaling in our mouse model caused diaphragmatic defects, highlighting the importance that these receptors have in modulating the role of RA during development (18,19,46–48). Based on the scRNA-seq data, *Rarb* is the main isoform expressed in the developing diaphragm, and therefore might have the most important role in the formation of CDH. Transcriptomic analysis performed by other groups have also shown that *Rarb* is the main isoform expressed in both the PPF and the diaphragm (69). However, immunohistochemistry analysis of the PPF failed to show protein expression of this RAR isoform, as only RARA had protein expression (133). Furthermore, the low levels of expression of *Rarg* could explain why only the *Rara*$^{-/-}$:*Rarb*$^{-/-}$ compound knockout develop CDH, compared to the *Rara*$^{-/-}$:*Rarg*$^{-/-}$ and the *Rarb*$^{-/-}$:*Rarg*$^{-/-}$ mutants that were previously described (63). Finally, it is important to note that although our coexpression analysis show little coexpression of *Prrx1* and any of the RAR isoforms, expression of the genes was contained to the mesenchymal component of the developing diaphragm.

Although previous transgenic animal models have been used to disrupt the retinoid acid signaling cascade, there are significant barriers, mainly low penetrance (63,64) as well as embryonic lethality (211,226,227). Our animal model provides a powerful tool to analyze fetuses that are well into development, and that develop CDH with 100% penetrance. Although there have been some knockout studies in which ablation of a specific retinoid-associated gene does not lead to the formation of diaphragmatic defects (44,48), our study further demonstrates that improper function of some retinoid associated genes lead to CDH (63,64,85). Other groups have also determined that mutations in downstream target genes of the retinoid signaling cascade

cause CDH, demonstrating the importance that RA signaling has in the development of the diaphragm (42,103). Finally, further studying the retinoid signaling cascade, as well as which genes are of paramount importance for proper diaphragm development is needed to better understand the etiology of CDH.

4.5 Conclusion and Future Directions

In conclusion, our *Prxx1-Cre:Rar^{dn}* model provides a versatile as well as potent tool to study diaphragmatic hernias. All of the *Prrx1-Cre:Rar^{dn}* fetuses analyzed in this study present some form of CDH, providing a clear advantage to other techniques previously used in the field to induce CDH. Although we saw a wide array of herniations in our fetuses, most of the hernias were seen in the posterior left side of the diaphragm. Our micro-CT analysis was also successful in letting us visualize the diaphragmatic hernias in a non-invasive way. Furthermore, both lung volume determination and 3D rendering demonstrated the presence of lung hypoplasia in our *Prxx1-Cre:Rar^{dn}* animal model. Our histology analysis revealed that diaphragm development is disrupted early in development, supporting the idea that CDH might occur before full muscularization and development of the diaphragm (38,184).

This model can still be utilized to further study the etiology of CDH. First, treating the dams with RA will help us better understand the role of retinoids in proper diaphragm development. It is expected that RA intake by the dams will either decrease the penetrance and/or size of the hernias. Since the *Rar^{dn}* model affects all RAR isoforms, conditional knockouts of each individual isoform can be used to determine which of them is most important for proper diaphragm development. *Rarb* in specific seems of high importance, based on the single cell expression analysis, coupled with previous evidence that have linked mutations in this gene with rare human conditions that present CDH (137). Overall, our mouse model not only demonstrates

the importance of RA in the development of the diaphragm, but also provides a powerful tool to study diaphragm development in an animal model and can hopefully lead to a better understanding of the etiology of this complex condition.

Chapter 5: *Dhrs3*$^{-/-}$ fetuses present gross abnormalities but not CDH

5.1 Introduction

Ever since the retinoid hypothesis, which states that retinoid signaling is involved in the incidence of CDH, there has been some interest in characterizing the role that different retinoid genes play in the development of the diaphragm. One of the enzymes that plays a role in the retinoid signaling cascade, dehydrogenase/reductase (SDR family) member 3 *(Dhrs3)* is in charge of reducing retinaldehyde into retinol, effectively mediating the production of RA (228,229). A schematic of the vitamin A pathway was provided in Figure 1.2 of this thesis. *Dhrs3* has been previously studied to determine its role during development in multiple tissues as well as models (211,228,230). Previous work has demonstrated that loss-of-function of this enzyme can lead to excess RA, causing multiple abnormalities as well as late embryonic lethality (211,229,231). Furthermore, one group in particular determined that *Dhrs3*$^{-/-}$ mice had a defective coronary vasculature and myocardium, as well as modified protein expression in the epicardium (230).

As CDH often presents comorbidity with different heart defects (232,233), we were interested in determining if *Dhrs3*$^{-/-}$ fetuses presented any incidence of diaphragmatic hernias. Our hypothesis was that the null fetuses will develop CDH, due to alterations in the retinoid signaling cascade in the developing diaphragm.

5.2 Methods

5.2.1 Animal Care and Breeding

All experiments were approved by the University of Alberta Research Ethics Committee and were performed in accordance with the Canadian Council on Animal Care guidelines. *Dhrs3* heterozygous breeders were a generous gift from Dr. Alexander Moise (Northern Ontario School

of Medicine) which were previously described in 2018 by Wang *et al.* (230). The animals were housed in the University of Alberta barrier animal facility. To generate timed pregnancies, one male *Dhrs3*$^{+/-}$ mice and two female *Dhrs3*$^{+/-}$ mice were set in a cage after 16:00 hrs and left overnight. The following morning at 9:00 hrs the presence of a vaginal plug was examined to assess for copulation. Each cage was left for a maximum of two nights. After a vaginal plug was seen, noon of that day is considered gestational day (GD) 0.5.

5.2.2 Dam Dissection and Fetus Collection

Dams at GD 13.5-15.5 were euthanized via anesthesia using isoflurane inhalation immediately followed by cervical dislocation. An abdominal incision was made in the mice, and the entire uterus was removed. Both the number of fetuses and resorptions were recorded. The uterus was placed in a petri dish containing Phosphate Buffer Saline (PBS) (Sigma-Aldrich, Oakville, ON). To dissect the individual fetuses, stereomicroscopy was used (Stemi 508, Zeiss, Oberkochen, Germany). The Crown-Rump Length (CRL) of each fetus was recorded. To obtain the diaphragm, both the cranial and caudal section of the fetus were dissected, and the heart and lungs were removed. Afterwards, the diaphragm was gently collected from atop the liver to assess for the incidence of diaphragmatic defects. The diaphragms were stored in formalin for later analysis.

5.2.3 Genotyping

To genotype the *Dhrs3* fetuses, DNA was extracted from the fetal tails and PCR reactions were performed utilizing *Dhrs3* WT (Forward: 5'-CTG AGG GTA AAG GGA CTC TGG-3', Reverse: 5'-AAT AGC CAG CGA GAT ACC AAT C-3') (Invitrogen, Toronto, ON), as well as the null allele (Forward: 5'-GCA GCG CAT CGC CTT CTA TC-3', Reverse: 5'-AAT AGC

CAG CGA GAT ACC AAT C-3') (Invitrogen) primers. PCR reactions were set by adding 2 µL of DNA, 0.5 µL of each primer, as well as 21.5 µL of PCR super-mix (Thermo Fisher Scientific, Toronto). The ProFlex PCR system (Applied Biosystems, Toronto, ON) was used for the reactions. Gel electrophoresis was performed utilizing a 1.2% agarose gel, which was ran at 100 V for 90 minutes. The PCR products were visualized utilizing RedSafe dye (Intron Biotechnology, Seoul, KR), as well as the ChemiDoc Touch Imaging System (BioRad, Mississauga, ON). The WT allele PCR product had a size of 242 bp, and the null allele had a product of 308 bp.

5.3 Results

To determine the effect of *Dhrs3*KO in the developing diaphragm, multiple *Dhrs3*$^{+/+}$, *Dhrs3*$^{+/-}$, and *Dhrs3*$^{-/-}$ fetuses were analyzed. The distribution of each genotype, as well as general information for each litter can be seen in Table 5.1. Contrary to our initial hypothesis, *Dhrs3*$^{-/-}$ fetuses do not present diaphragmatic hernias. None of the null fetuses analyzed presented any sort of diaphragmatic defects, although they did present heart malformations like previously described by other groups (230). Representative images of the diaphragm and heart dissected from *Dhrs3*$^{+/+}$ and *Dhrs3*$^{-/-}$ fetuses can be seen in Figure 5.1.

			Fetal Genotype				
Litter	Gestational Day (GD)	Average CRL (mm)	+/+	+/-	-/-	Total	No. of Resorptions
1	13.5	10.5	0	5	3	8	0
2	15.5	14.5	2	2	0	4	2
3	15.5	14	2	1	2	5	3
4	14.5	12.5	4	1	3	8	1
5	14.5	12.5	1	4	2	6	1
6	14.5	14	1	3	2	6	1
		Fetal genotype distribution (n)	10	16	12	38	
		Fetal Genotype Distribution (%)	26	42	32	NA	
		No. of fetuses with CDH	0	0	0	0	

Table 5.1: General information for the *Dhrs3* fetuses. A total of 6 litters were used for our studies, with a total of 10 WT, 16 heterozygous, and 12 null fetuses analyzed. None of the fetuses analyzed had CDH.

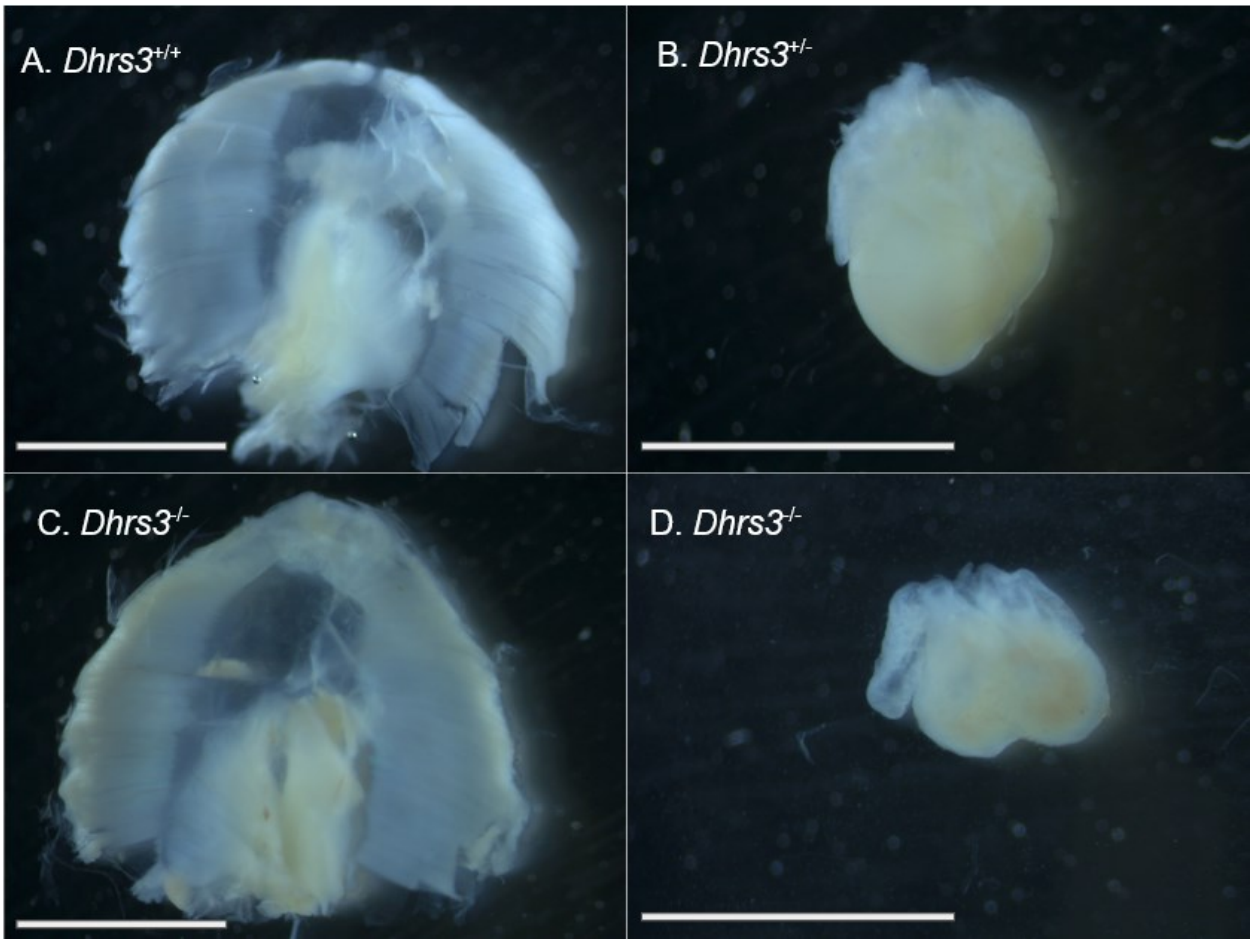


Figure 5.1: Dhrs3$^{-/-}$ mice do not develop CDH. Representative images of the diaphragm (A and C) and heart (B and D) from *Dhrs3*$^{+/+}$, *Dhrs3*$^{+/-}$ and *Dhrs3*$^{-/-}$ fetuses are shown, respectively. All scale bars are 0.5 mm.

5.4 Discussion

Dhrs3 is an important member of the retinoid signaling cascade during development and has been implicated with multiple congenital abnormalities (211,230); however, contrary to our initial hypothesis, loss-of-function of this gene does not cause CDH, since there is no difference in the diaphragms of WT fetuses compared to the *Dhrs3*$^{-/-}$ fetuses. It is important to note that we were able to see heart anomalies just like previous studies (230), demonstrating the importance of this gene during cardiac development. We did not observe any gross abnormalities in the diaphragm, as all fetuses analyzed had a fully muscularized diaphragm and no herniation, although is possible that there is an effect in the *Dhrs3*$^{-/-}$ fetuses that is not observable through microscopy. A key limitation to this study is that *Dhrs3*$^{-/-}$ animals present embryonic lethality after GD 15.5 (87,211,230), meaning that we were not able to dissect a fully muscularized diaphragm.

Single-cell transcriptomic studies could be used to test the differences between the diaphragms of WT and KO fetuses, and to assess if any specific cell population is affected in the *Dhrs3*$^{-/-}$ mice. Furthermore, even though based on our evidence this gene does not cause CDH, there are many different genes in the retinoid signaling cascade. Therefore, more transgenic studies can be done to test the effect that knocking out different retinoid-related genes might have on the proper development of the diaphragm. Another possible limitation would be that *Dhrs3* has a low penetrance, and the number of fetuses analyzed (n=12) was not sufficient to see diaphragmatic defects. However, this seems unlikely as other animal model experiments with low penetrance saw the formation of CDH in a similar sample number (63).

Chapter 6: Conclusion

6.1 Overall Conclusion

The relationship between vitamin A and CDH has been a topic of research for more than 70 years. Since the first studies that linked this micronutrient with diaphragmatic defects (234,235), to the formulation of the retinoid hypothesis in 2003 (31), it is clear that there is a link between retinoic acid (RA) signaling and CDH. Through this thesis, we hoped to answer two different questions to better understand the etiology of CDH; (i) if blocking the RA signaling cascade in the mesenchymal component of the diaphragm will cause CDH and (ii) which cellular populations of the developing diaphragm express both retinoid and CDH-associated genes. We were successful in answering both questions, by utilizing mutant mice studies for the first question, and single cell expression studies for the second question. Overall, we found that (i) the mesenchyme of the PPFs is highly important for proper organogenesis of the diaphragm, as disrupting the RA signaling cascade in this tissue leads to diaphragmatic hernias, as well as (ii) the fact that the non-muscular mesenchyme of the developing diaphragm has the highest percentage of expression of both retinoid and CDH-associated genes.

6.2 Overview and Discussion of Results

To determine which cellular populations of the developing diaphragm might be of higher importance for the incidence of CDH, we performed scRNA-seq on PPF tissue. Based on this expression analysis, we were able to identify 18 distinct clusters, with 10 unique cell populations. The majority of the cells were determined to be mesenchymal cells, and a total of five different mesenchymal clusters were identified. Furthermore, we determined the level of expression of both retinoid and CDH-associated genes through this analysis. In general, the expression of retinoid associated genes was mostly seen in the mesenchymal and mesothelial component of the diaphragm. For example, *Aldha1a2* is highest expressed in the ‘Mesothelial

Cells' cluster. For *Rarb*, the 'Mesenchymal Cells 1' cluster had the highest level of expression The same type of analysis was performed for CDH-associated genes. These genes were curated based on previous studies (21,199). Overall, the same pattern of expression was seen for CDH-associated genes, with the mesenchymal and mesothelial clusters having the highest number of genes expressed. For example, *Wt1,* a gene that is known to cause CDH when mutated (42,57,184) had highest expression in the 'Mesenchymal Cells 1', 'Mesenchymal Cells 3' and 'Mesothelial Cells' clusters. However, not all CDH-associated genes had the highest level of expression in the mesenchymal or mesothelial components. *Nr2f2*, for example, had the highest level of expression in the 'Endothelial Cells 1' cluster. Since this was the first time scRNA-seq analysis was performed in the PPF, we were not only able to determine that the mesenchymal and mesothelial component of the PPF show the highest level of expression of both retinoid and CDH-associated genes, but we were also able to better understand the composition of this tissue, and how it plays a role in the development of diaphragmatic defects.

To test if disrupting the RA signaling cascade will cause CDH, we generated a conditional mutant mouse model that express a truncated version of the RAR receptor (RAR^{dn}) in tissues that express *Prrx1*, by using a *Cre* recombinase system. As *Prrx1* is highly expressed in the mesenchyme, this system aims to test if blocking RA signaling in the mesenchyme of the diaphragm will cause CDH. Based on our studies, this *Prrx1-Cre:*Rar^{dn} system generates fetuses that develop CDH with 100% penetrance. The most common location of the hernias was on the left side of the diaphragm, and most of said hernias were type C, were more than 50% of one side of the diaphragm was missing. We also performed micro-CT analysis on these *Prrx1-Cre:*Rar^{dn} fetuses, and determined that the lung volume of the ipsilateral lung is lower in fetuses with CDH than it is in the control by around 30%. This decrease in volume was not seen in the

right lungs of the *Prrx1-Cre:Rar^{dn}* fetuses compared to the control. Overall, this mouse model confirmed our hypothesis that disrupting the RA signaling cascade in the mesenchymal component of the diaphragm causes CDH.

6.3 Limitations

There are several limitations to this study. Regarding our conditional mouse model, one key limitation is the formation of *Prrx1-Cre:Rar^{dn}* fetuses. A total of 61 fetuses were collected, but less than a third of them were *Prrx1-Cre:Rar^{dn}* fetuses, meaning that a high number of dams and fetuses were required for this study. Furthermore, a total of 6 *Prrx1-Cre:Rar^{dn}* fetuses and 6 control fetuses were analyzed via micro-CT. The low number of fetuses analyzed might explain why we were not able to see volume differences in the right lung of the *Prrx1-Cre:Rar^{dn}* fetuses compared to the control. Analyzing a higher number of fetuses would have given us higher statistical power, and we might have been able to see a significant difference in the right lung volumes. Regarding our scRNA-seq analysis, the biggest limitation is the fact expression data is based on RNA levels. Since RNA levels and protein expression is not fully correlated, it is possible that our analysis under or overrepresented some genes, and therefore we were not able to fully assess the expression pattern of both retinoid and CDH-associated genes.

6.4 Future Directions

Further studies can be done to further assess the important of RA signaling in the mesenchymal component of the diaphragm. To start, the *Rar^{dn}* system used in our conditional mutant mouse models sequesters RA, meaning the activity of all three RAR isoforms will be blocked. Utilizing mutant versions of each isoform individually might provide a better idea of which isoform is most important during development, and therefore most likely to cause CDH.

Based on both animal and human data, *Rarb* seems to be the most important isoform, as mutations in humans are known to cause CDH, and compound *Rara*$^{-/-}$:*Rarb*$^{-/-}$ mutant mice also develop CDH (63,64,137). Our scRNA-seq data also supports this statement, as *Rarb* shows the highest level of expression out of the three RAR isoforms, and it is mostly contained in the mesenchymal and mesothelial components of the PPF. Another plausible study with our *Prrx1-Cre:Rar*^{dn} model might be to supplement the diets of the dams with RA, to see the effect that this has in the *Prrx1-Cre:Rar*^{dn} fetuses. As low vitamin A intake by the dam has been implicated with a higher incidence of CDH (82), it is possible that RA supplementation will lead to a decrease in the incidence of CDH in the *Prrx1-Cre:Rar*^{dn} fetuses, or that this supplementation will cause changes in the size as well as the type of CDH seen in the *Prrx1-Cre:Rar*^{dn} fetuses.

Regarding our scRNA-seq studies, we could test for protein expression of key genes highlighted through this analysis, to determine if RNA levels correlate to expression in the PPF. Studies that assessed protein expression in the PPF have been previously performed, although there are some discrepancies in the results. For example, *Wt1* showed high levels of expression in our single cell expression analysis as well as in the protein expression studies (133). In contrast *Rarb* did not show protein expression, even though it is the main isoform expressed in the developing diaphragm based on our scRNA-seq analysis (133). Further protein expression analysis is therefore needed to solve some of these discrepancies. Single cell expression analysis can also be performed in a PPF sample with CDH, to determine how different it is from our control PPF sample. As the mesenchymal component of the PPF is necessary for proper diaphragm development, it is expected that scRNA-seq data of a PPF with CDH will have major differences in the mesenchymal and mesothelial components compared to our sample. We also

expect that the expression patterns of both retinoid and CDH-associated genes to be highly different, as the mesenchymal and mesothelial component will be different in a PPF with CDH.

6.5 Clinical Relevance

CDH is a disease with a high level of occurrence (2.3 births per 10000) that puts high strain in the health system. The average cost for the hospitalization and management of CDH is more than $350,000 in the US (11,236), and more than $50,000 in Canada (237). Furthermore, although the survivability of this condition is high in developed countries, the etiology of this syndrome is poorly understood. Through this thesis, we added knowledge into the already existing research trying to better understand the etiology and pathogenesis of CDH. We were able to further strengthen the link between RA signaling and CDH, as we showed that disrupting the RA signaling cascade in the mesenchymal component of the diaphragm causes diaphragmatic defects. Furthermore, we determined that retinoid related genes are mostly expressed in the mesenchymal and mesothelial components of the developing diaphragm. We were also able to better understand the PPF, the main source of CDH. Via our scRNA-seq analysis, we were able to show 10 distinct cell populations in the developing diaphragm, while also showing that the non-muscular mesenchyme shows the highest levels of expression of both retinoid and CDH-associated genes. Our analysis not only helps us better understand the basic physiology of the PPF, but also the relationship between RA signaling in the non-muscular mesenchyme and the incidence of CDH.

6.6 Final Remarks

Overall, this thesis further supports the retinoid hypothesis and the idea that the non-muscular mesenchyme is the main driver in the formation of diaphragmatic defects. Both of our hypotheses were validated as (i) disrupting the RA signaling cascade in the mesenchymal component of the diaphragm causes CDH and (ii) the mesenchymal and mesothelial components of the PPF are the ones with highest expression of retinoid and CDH-associated genes. Further studies based on this thesis can be performed to better understand the etiology of CDH and to better supplement the retinoid hypothesis.

Bibliography

1. McGivern, M. R. *et al.* Epidemiology of congenital diaphragmatic hernia in Europe: a register-based study. *Arch Dis Child Fetal Neonatal Ed* **100,** F137 (2015).

2. Shanmugam, H., Brunelli, L., Botto, L. D., Krikov, S. & Feldkamp, M. L. Epidemiology and Prognosis of Congenital Diaphragmatic Hernia: A Population-Based Cohort Study in Utah. *Birth Defects Res* **109,** 1451–1459 (2017).

3. Paoletti, M. *et al.* Prevalence and risk factors for congenital diaphragmatic hernia: A global view. *J Pediatr Surg* **55,** 2297–2307 (2020).

4. Wagner, R., Montalva, L., Zani, A. & Keijzer, R. Basic and translational science advances in congenital diaphragmatic hernia. *Semin Perinatol* **44,** 151170 (2020).

5. Zani, A. *et al.* Congenital diaphragmatic hernia. *Nat Rev Dis Primers* **8,** 37 (2022).

6. Wright, N. J. *et al.* Mortality from gastrointestinal congenital anomalies at 264 hospitals in 74 low-income, middle-income, and high-income countries: a multicentre, international, prospective cohort study. *The Lancet* **398,** 325–339 (2021).

7. Wynn, J. *et al.* Outcomes of Congenital Diaphragmatic Hernia in the Modern Era of Management. *J Pediatr* **163,** 114-119.e1 (2013).

8. Harting, M. T. & Lally, K. P. The Congenital Diaphragmatic Hernia Study Group registry update. *Semin Fetal Neonatal Med* **19,** 370–375 (2014).

9. Cameron, D. B. *et al.* Quantifying the Burden of Interhospital Cost Variation in Pediatric Surgery: Implications for the Prioritization of Comparative Effectiveness Research. *JAMA Pediatr* **171,** e163926–e163926 (2017).

10. Raval, M. v, Wang, X., Reynolds, M. & Fischer, A. C. Costs of congenital diaphragmatic hernia repair in the United States—extracorporeal membrane oxygenation foots the bill. *J Pediatr Surg* **46,** 617–624 (2011).

11. Lewit, R. A. & Jancelewicz, T. Sources of regional and center-level variability in survival and cost of care for congenital diaphragmatic hernia (CDH). *J Pediatr Surg* **56,** 130–135 (2021).

12. Wong, M. *et al.* Pulmonary hypertension in congenital diaphragmatic hernia patients: Prognostic markers and long-term outcomes. *J Pediatr Surg* **53,** 918–924 (2018).

13. Mills, J., Safavi, A. & Skarsgard, E. D. Chylothorax after congenital diaphragmatic hernia repair: a population-based study. *J Pediatr Surg* **47,** 842–846 (2012).

14. Nagata, K. *et al.* Risk Factors for the Recurrence of the Congenital Diaphragmatic Hernia—Report from the Long-Term Follow-Up Study of Japanese CDH Study Group. *European Journal of Pediatric Surgery* **25,** (2014).

15. Jancelewicz, T., Chiang, M., Oliveira, C. & Chiu, P. P. Late surgical outcomes among congenital diaphragmatic hernia (CDH) patients: Why long-term follow-up with surgeons is recommended. *J Pediatr Surg* **48,** 935–941 (2013).

16. Burns, N. G. & Kardon, G. in *Curr Top Dev Biol* (eds. Lipinski, R. J. & Krauss, R. S.) **152,** 115–138 (Academic Press, 2023).

17. Pober, B. Genetic aspects of human congenital diaphragmatic hernia. *Clin Genet* **74,** 1–15 (2008).

18. Peppa, M. *et al.* Congenital diaphragmatic hernia subtypes: Comparing birth prevalence, occurrence by maternal age, and mortality in a national birth cohort. *Paediatr Perinat Epidemiol* **37,** 143–153 (2023).

19. Ackerman, K. G. *et al.* Congenital Diaphragmatic Defects: Proposal for a New Classification Based on Observations in 234 Patients. *Pediatric and Developmental Pathology* **15,** 265–274 (2012).

20. Kardon, G. *et al.* Congenital diaphragmatic hernias: from genes to mechanisms to therapies. *Dis Model Mech* **10,** 955–970 (2017).

21. Dalmer, T. R. A. & Clugston, R. D. Gene ontology enrichment analysis of congenital diaphragmatic hernia-associated genes. *Pediatr Res* **85,** 13–19 (2019).

22. Major, D. *et al.* Retinol status of newborn infants with congenital diaphragmatic hernia. *Pediatr Surg Int* **13,** 547–549 (1998).

23. Russell, M. K. *et al.* Congenital diaphragmatic hernia candidate genes derived from embryonic transcriptomes. *Proc Natl Acad Sci U S A* **109,** 2978–2983 (2012).

24. Veenma, D. C. M., de Klein, A. & Tibboel, D. Developmental and genetic aspects of congenital diaphragmatic hernia. *Pediatr Pulmonol* **47,** 534–545 (2012).

25. Longoni, M. *et al.* Genome-wide enrichment of damaging de novo variants in patients with isolated and complex congenital diaphragmatic hernia. *Hum Genet* **136,** 679–691 (2017).

26. Yu, L. *et al.* Increased burden of de novo predicted deleterious variants in complex congenital diaphragmatic hernia. *Hum Mol Genet* **24,** 4764–4773 (2015).

27. McAteer, J. P., Hecht, A., De Roos, A. J. & Goldin, A. B. Maternal medical and behavioral risk factors for congenital diaphragmatic hernia. *J Pediatr Surg* **49,** 34–38 (2014).

28. Schulz, F. *et al.* Parental risk factors for congenital diaphragmatic hernia – a large German case-control study. *BMC Pediatr* **21,** 278 (2021).

29. García, A. M., Machicado, S., Gracia, G. & Zarante, I. M. Risk factors for congenital diaphragmatic hernia in the Bogota birth defects surveillance and follow-up program, Colombia. *Pediatr Surg Int* **32,** 227–234 (2016).

30. Felix, J. F. *et al.* Environmental factors in the etiology of esophageal atresia and congenital diaphragmatic hernia: Results of a case-control study. *Birth Defects Res A Clin Mol Teratol* **82,** 98–105 (2008).

31. Greer, J. J., Babiuk, R. P. & Thebaud, B. Etiology of Congenital Diaphragmatic Hernia: The Retinoid Hypothesis. *Pediatr Res* **53,** 726–730 (2003).

32. Giguère, V. & Evans, R. M. Chronicle of a discovery: the retinoic acid receptor. *J Mol Endocrinol* **69,** T1–T11 (2022).

33. Al Tanoury, Z., Piskunov, A. & Rochette-Egly, C. Vitamin A and retinoid signaling: genomic and nongenomic effects: Thematic Review Series: Fat-Soluble Vitamins: Vitamin A. *J Lipid Res* **54,** 1761–1775 (2013).

34. Duester, G. Retinoid signaling in control of progenitor cell differentiation during mouse development. *Semin Cell Dev Biol* **24,** 694–700 (2013).

35. Clugston, R. D. & Greer, J. J. Diaphragm development and congenital diaphragmatic hernia. *Semin Pediatr Surg* **16,** 94–100 (2007).

36. Merrell, A. & Kardon, G. Development of the diaphragm - A skeletal muscle essential for mammalian respiration. *FEBS J* **280,** (2013).

37. Sefton, E. M., Gallardo, M. & Kardon, G. Developmental origin and morphogenesis of the diaphragm, an essential mammalian muscle. *Dev Biol* **440,** 64–73 (2018).

38. Merrell, A. J. *et al.* Muscle connective tissue controls development of the diaphragm and is a source of congenital diaphragmatic hernias. *Nat Genet* **47,** 496–504 (2015).

39. Sefton, E. M. *et al.* Fibroblast-derived Hgf controls recruitment and expansion of muscle during morphogenesis of the mammalian diaphragm. *Elife* **11,** e74592 (2022).

40. Allan, D. W. & Greer, J. J. Pathogenesis of nitrofen-induced congenital diaphragmatic hernia in fetal rats. *J Appl Physiol* **83,** 338–347 (1997).

41. Babiuk, R. P. & Greer, J. J. Diaphragm defects occur in a CDH hernia model independently of myogenesis and lung formation. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **283,** L1310–L1314 (2002).

42. Carmona, R. *et al.* Conditional deletion of WT1 in the septum transversum mesenchyme causes congenital diaphragmatic hernia in mice. *Elife* **5,** e16009 (2016).

43. Paris, N. D., Coles, G. L. & Ackerman, K. G. Wt1 and β-catenin cooperatively regulate diaphragm development in the mouse. *Dev Biol* **407,** 40–56 (2015).

44. O'Byrne, S. M. *et al.* Retinoid Absorption and Storage Is Impaired in Mice Lacking Lecithin:Retinol Acyltransferase (LRAT)*. *Journal of Biological Chemistry* **280,** 35647–35657 (2005).

45. Steinhoff, J. S., Lass, A. & Schupp, M. Biological Functions of RBP4 and Its Relevance for Human Diseases. *Front Physiol* **12,** (2021).

46. Noy, N. in *The Biochemistry of Retinoid Signaling II: The Physiology of Vitamin A - Uptake, Transport, Metabolism and Signaling* (eds. Asson-Batres, M. A. & Rochette-Egly, C.) 77–93 (Springer Netherlands, 2016). doi:10.1007/978-94-024-0945-1_3

47. O'Connor, C., Varshosaz, P. & Moise, A. R. Mechanisms of Feedback Regulation of Vitamin A Metabolism. *Nutrients* **14,** (2022).

48. Quadro, L. *et al.* Impaired retinal function and vitamin A availability in mice lacking retinol-binding protein. *EMBO J* **18,** 4633–4644 (1999).

49. Li, E. & Norris, A. W. Structure/Function of Cytoplasmic Vitamin A-Binding Proteins. *Annu Rev Nutr* **16,** 205–234 (1996).

50. di Masi, A. *et al.* Retinoic acid receptors: From molecular mechanisms to cancer therapy. *Mol Aspects Med* **41,** 1–115 (2015).

51. Wolf, G. Is 9-Cis-Retinoic Acid the Endogenous Ligand for the Retinoic Acid-X Receptor? *Nutr Rev* **64,** 532–538 (2006).

52. Rühl, R. *et al.* 9-cis-13,14-Dihydroretinoic Acid Is an Endogenous Retinoid Acting as RXR Ligand in Mice. *PLoS Genet* **11,** e1005213- (2015).

53. Petkovich, M. & Chambon, P. Retinoic acid receptors at 35 years. *J Mol Endocrinol* **69,** T13–T24 (2022).

54. Schubert, M. & Gibert, Y. Retinoids in Embryonic Development. *Biomolecules* **10,** (2020).

55. Clagett-Dame, M. & Knutson, D. Vitamin A in Reproduction and Development. *Nutrients* **3,** 385–428 (2011).

56. Ghyselinck, N. B. & Duester, G. Retinoic acid signaling pathways. *Development* **146,** dev167502 (2019).

57. Cleal, L. *et al.* Resolving the heterogeneity of diaphragmatic mesenchyme: a novel mouse model of congenital diaphragmatic hernia. *Dis Model Mech* **14,** dmm046797 (2021).

58. Rosselot, C. *et al.* Non-cell-autonomous retinoid signaling is crucial for renal development. *Development* **137,** 283–292 (2010).

59. Anderson, D. H. @article{Anderson1949, author = {D.H. Anderson}, journal = {Am. J. Dis. Child.}, pages = {888-889}, title = {Incidence of congenital diaphragmatic hernia in the young of rats bred on a diet deficient in vitamin. *Am. J. Dis. Child.* **62,** 888–889 (1941).

60. Wilson, J. G., Roth, C. B., W. J. An analysis of the syndrome of malformations induced by maternal vitamin A deficiency. Effects of restoration of vitamin A at various times during gestation. *American Journal of Anatomy* 189–217 (1953).

61. Anderson, D. H. Effect of diet during pregnancy upon the incidence of congenital hereditary diaphragmatic hernia in the rat. *American Journal of Pathology* **25,** 163–185 (1949).

62. Anderson, D. H. Incidence of congenital diaphragmatic hernia in the young of rats bred on a diet deficient in vitamin. *The American Journal of Diseases of Children* **62,** 888–889 (1941).

63. Mendelsohn, C. *et al.* Function of the retinoic acid receptors (RARs) during development (II). Multiple abnormalities at various stages of organogenesis in RAR double mutants. *Development* **120,** 2749–2771 (1994).

64. Ghyselinck, N. B. *et al.* Role of the retinoic acid receptor beta (RARbeta) during mouse development. *Int J Dev Biol* **41,** 425–447 (1997).

65. Mey, J., Babiuk, R. P., Clugston, R., Zhang, W. & Greer, J. J. Retinal Dehydrogenase-2 Is Inhibited by Compounds that Induce Congenital Diaphragmatic Hernias in Rodents. *Am J Pathol* **162,** 673–679 (2003).

66. Chen, M. -h., MacGowan, A., Ward, S., Bavik, C. & Greer, J. J. The Activation of the Retinoic Acid Response Element Is Inhibited in an Animal Model of Congenital Diaphragmatic Hernia. *Neonatology* **83,** 157–161 (2003).

67. Thébaud, B. *et al.* Vitamin A decreases the incidence and severity of nitrofen-induced congenital diaphragmatic hernia in rats. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **277,** L423–L429 (1999).

68. Clugston, R. D., Zhang, W., Álvarez, S., de Lera, A. R. & Greer, J. J. Understanding Abnormal Retinoid Signaling as a Causative Mechanism in Congenital Diaphragmatic Hernia. *Am J Respir Cell Mol Biol* **42,** 276–285 (2010).

69. Russell, M. K. *et al.* Congenital diaphragmatic hernia candidate genes derived from embryonic transcriptomes. *Proceedings of the National Academy of Sciences* **109,** 2978–2983 (2012).

70. Edel, G. *et al.* Cellular Origin(s) of Congenital Diaphragmatic Hernia. *Front Pediatr* **9,** (2021).

71. Kluth, D. *et al.* Nitrofen-induced diaphragmatic hernias in rats: An animal model. *J Pediatr Surg* **25,** 850–854 (1990).

72. Tenbrinck, R. *et al.* Experimentally induced congenital diaphragmatic hernia in rats. *J Pediatr Surg* **25,** 426–429 (1990).

73. Chiu, P. P. L. New Insights into Congenital Diaphragmatic Hernia – A Surgeon's Introduction to CDH Animal Models. *Front Pediatr* **2,** (2014).

74. Noble, B. R. *et al.* Mechanisms of action of the congenital diaphragmatic hernia-inducing teratogen nitrofen. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **293,** L1079–L1087 (2007).

75. Kling, D. E. *et al.* Nitrofen induces apoptosis independently of retinaldehyde dehydrogenase (RALDH) inhibition. *Birth Defects Res B Dev Reprod Toxicol* **89,** 223–232 (2010).

76. Kutasy, B., Pes, L., Friedmacher, F., Paradisi, F. & Puri, P. Nitrofen increases total retinol levels in placenta during lung morphogenesis in the nitrofen model of congenital diaphragmatic hernia. *Pediatr Surg Int* **30,** 1017–1022 (2014).

77. Baptista, M. J. *et al.* Antenatal vitamin A administration attenuates lung hypoplasia by interfering with early instead of late determinants of lung underdevelopment in congenital diaphragmatic hernia. *J Pediatr Surg* **40,** 658–665 (2005).

78. Oshiro, T., Asato, Y., Sakanashi, M., Ohta, T. & Sugahara, K. Differential effects of vitamin A on fetal lung growth and diaphragmatic formation in nitrofen-induced rat model. *Pulm Pharmacol Ther* **18,** 155–164 (2005).

79. Babiuk, R. P., Thébaud, B. & Greer, J. J. Reductions in the incidence of nitrofen-induced diaphragmatic hernia by vitamin A and retinoic acid. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **286,** L970–L973 (2004).

80. Beurskens, L. W. J. E. *et al.* Dietary vitamin A intake below the recommended daily intake during pregnancy and the risk of congenital diaphragmatic hernia in the offspring. *Birth Defects Res A Clin Mol Teratol* **97,** 60–66 (2013).

81. Michikawa, T. *et al.* Maternal dietary intake of vitamin A during pregnancy was inversely associated with congenital diaphragmatic hernia: the Japan Environment and Children's Study. *British Journal of Nutrition* **122,** 1295–1302 (2019).

82. Rocke, A. W. *et al.* Low maternal vitamin A intake increases the incidence of teratogen induced congenital diaphragmatic hernia in mice. *Pediatr Res* **91,** 83–91 (2022).

83. Kool, H. M. *et al.* Inhibition of retinoic acid signaling induces aberrant pericyte coverage and differentiation resulting in vascular defects in congenital diaphragmatic hernia. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **317,** L317–L331 (2019).

84. Cipollone, D., Cozzi, D. A., Businaro, R. & Marino, B. Congenital diaphragmatic hernia after exposure to a triple retinoic acid antagonist during pregnancy. *Journal of Cardiovascular Medicine* **18,** (2017).

85. Mascrez, B. *et al.* The RXRα ligand-dependent activation function 2 (AF-2) is important for mouse development. *Development* **125,** 4691–4707 (1998).

86. Ghyselinck, N. B. *et al.* Cellular retinol-binding protein I is essential for vitamin A homeostasis. *EMBO J* **18,** 4903–4914 (1999).

87. Billings, S. E. *et al.* The retinaldehyde reductase DHRS3 is essential for preventing the formation of excess retinoic acid during embryonic development. *The FASEB Journal* **27,** 4877–4889 (2013).

88. Abu-Abed, S. *et al.* The retinoic acid-metabolizing enzyme, CYP26A1, is essential for normal hindbrain patterning, vertebral identity, and development of posterior structures. *Genes Dev* **15,** 226–240 (2001).

89. Niederreither, K., Subbarayan, V., Dollé, P. & Chambon, P. Embryonic retinoic acid synthesis is essential for early mouse post-implantation development. *Nat Genet* **21,** 444–448 (1999).

90. Yu, L., Hernan, R. R., Wynn, J. & Chung, W. K. The influence of genetics in congenital diaphragmatic hernia. *Semin Perinatol* **44,** 151169 (2020).

91. Wilm, B. & Muñoz-Chapuli, R. in *The Wilms' Tumor (WT1) Gene: Methods and Protocols* (ed. Hastie, N.) 23–39 (Springer New York, 2016). doi:10.1007/978-1-4939-4023-3_3

92. Clugston, R. D. *et al.* Teratogen-Induced, Dietary and Genetic Models of Congenital Diaphragmatic Hernia Share a Common Mechanism of Pathogenesis. *Am J Pathol* **169,** 1541–1549 (2006).

93. Kreidberg, J. A. *et al.* WT-1 is required for early kidney development. *Cell* **74,** 679–691 (1993).

94. Hastie, N. D. Wilms' tumour 1 (WT1) in development, homeostasis and disease. *Development* **144,** 2862–2872 (2017).

95. Guadix, J. A. *et al.* Wt1 controls retinoic acid signalling in embryonic epicardium through transcriptional activation of Raldh2. *Development* **138,** 1093–1097 (2011).

96. Balmer, J. E. & Blomhoff, R. Gene expression regulation by retinoic acid. *J Lipid Res* **43,** 1773–1808 (2002).

97. Lee, C. T. *et al.* The nuclear orphan receptor COUP-TFII is required for limb and skeletal muscle development. *Mol Cell Biol* **24,** 10835–10843 (2004).

98. Pereira, F. A., Qiu, Y., Zhou, G., Tsai, M.-J. & Tsai, S. Y. The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev* **13,** 1037–1049 (1999).

99. Doi, T., Sugimoto, K. & Puri, P. Prenatal retinoic acid up-regulates pulmonary gene expression of COUP-TFII, FOG2, and GATA4 in pulmonary hypoplasia. *J Pediatr Surg* **44,** 1933–1937 (2009).

100. Qiu, Y., Krishnan, V., Pereira, F. A., Tsai, S. Y. & Tsai, M. J. Chicken ovalbumin upstream promoter-transcription factors and their regulation. *J Steroid Biochem Mol Biol* **56,** 81–85 (1996).

101. Jonk, L. J. C. *et al.* Cloning and expression during development of three murine members of the COUP family of nuclear orphan receptors. *Mech Dev* **47,** 81–97 (1994).

102. Tsai, S. Y. & Tsai, M.-J. Chick Ovalbumin Upstream Promoter-Transcription Factors (COUP-TFs): Coming of Age*. *Endocr Rev* **18,** 229–240 (1997).

103. You, L.-R. *et al.* Mouse lacking COUP-TFII as an animal model of Bochdalek-type congenital diaphragmatic hernia. *Proc Natl Acad Sci U S A* **102,** 16351–16356 (2005).

104. Perrino, C. & Rockman, H. A. GATA4 and the Two Sides of Gene Expression Reprogramming. *Circ Res* **98,** 715–716 (2006).

105. Zhou, P., He, A. & Pu, W. T. in *Curr Top Dev Biol* (ed. Bruneau, B. G.) **100,** 143–169 (Academic Press, 2012).

106. Su, D. & Gudas, L. J. Retinoic acid receptor γ activates receptor tyrosine kinase Tie1 gene transcription through transcription factor GATA4 in F9 stem cells. *Exp Hematol* **36,** 624–641 (2008).

107. Jay, P. Y. *et al.* Impaired mesenchymal cell function in Gata4 mutant mice leads to diaphragmatic hernias and primary lung defects. *Dev Biol* **301,** 602–614 (2007).

108. Lu, J. R. *et al.* FOG-2, a heart- and brain-enriched cofactor for GATA transcription factors. *Mol Cell Biol* **19,** 4495–4502 (1999).

109. Ackerman, K. G. *et al.* Fog2 Is Required for Normal Diaphragm and Lung Development in Mice and Humans. *PLoS Genet* **1,** e10- (2005).

110. Arceci, R. J., King, A. A., Simon, M. C., Orkin, S. H. & Wilson, D. B. Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart. *Mol Cell Biol* **13,** 2235–2246 (1993).

111. Jiang, Y., Drysdale, T. A. & Evans, T. A Role for GATA-4/5/6 in the Regulation of Nkx2.5 Expression with Implications for Patterning of the Precardiac Field. *Dev Biol* **216,** 57–71 (1999).

112. Clabby, M. L., Robison, T. A., Quigley, H. F., Wilson, D. B. & Kelly, D. P. Retinoid X Receptor α Represses GATA-4-mediated Transcription via a Retinoid-dependent Interaction with the Cardiac-enriched Repressor FOG-2*. *Journal of Biological Chemistry* **278,** 5760–5767 (2003).

113. Huggins, G. S., Bacani, C. J., Boltax, J., Aikawa, R. & Leiden, J. M. Friend of GATA 2 Physically Interacts with Chicken Ovalbumin Upstream Promoter-TF2 (COUP-TF2) and COUP-TF3 and Represses COUP-TF2-dependent Activation of the Atrial Natriuretic Factor Promoter*. *Journal of Biological Chemistry* **276,** 28029–28036 (2001).

114. Gilbert, R. M. & Gleghorn, J. P. Connecting clinical, environmental, and genetic factors point to an essential role for vitamin A signaling in the pathogenesis of congenital diaphragmatic hernia. *American Journal of Physiology-Lung Cellular and Molecular Physiology* (2023). doi:10.1152/ajplung.00349.2022

115. Beurskens, L. W. J. E. *et al.* Retinol Status of Newborn Infants Is Associated With Congenital Diaphragmatic Hernia. *Pediatrics* **126,** 712–720 (2010).

116. Loo, C. K. C. *et al.* Lung and liver growth and retinoic acid status in human fetuses with congenital diaphragmatic hernia. *Early Hum Dev* **116,** 17–23 (2018).

117. Rocke, A. W. & Clugston, R. D. Comment on "Lung and Liver growth and retinoic acid status in human fetuses with congenital diaphragmatic hernia". *Early Hum Dev* **116,** 93 (2018).

118. Czuba, L. C. *et al.* Plasma Retinoid Concentrations Are Altered in Pregnant Women. *Nutrients* **14,** (2022).

119. Jeong, H. *et al.* Temporal changes in the systemic concentrations of retinoids in pregnant and postpartum women. *PLoS One* **18,** e0280424- (2023).

120. Bates, C. J. Vitamin A in pregnancy and lactation. *Proceedings of the Nutrition Society* **42,** 65–79 (1983).

121. Takahashi, Y. I., Smith, J. E. & Goodman, D. S. Vitamin A and retinol-binding protein metabolism during fetal development in the rat. *American Journal of Physiology-Endocrinology and Metabolism* **233,** E263 (1977).

122. Satre, M. A., Ugen, K. E. & Kochhar, D. M. Developmental changes in endogenous retinoids during pregnancy and embryogenesis in the mouse. *Biol Reprod* **46,** 802–810 (1992).

123. Brosens, E. *et al.* Unraveling the Genetics of Congenital Diaphragmatic Hernia: An Ongoing Challenge. *Front Pediatr* **9,** (2022).

124. Goumy, C. *et al.* Retinoid Pathway and Congenital Diaphragmatic Hernia: Hypothesis from the Analysis of Chromosomal Abnormalities. *Fetal Diagn Ther* **28,** 129–139 (2010).

125. Kawaguchi, R. *et al.* A Membrane Receptor for Retinol Binding Protein Mediates Cellular Uptake of Vitamin A. *Science (1979)* **315,** 820–825 (2007).

126. Andijani, A. A., Shajira, E. S., Abushaheen, A. & Al-Matary, A. Microphthalmia Syndrome 9: Case Report of a Newborn Baby with Pulmonary Hypoplasia, Diaphragmatic Eventration, Microphthalmia, Cardiac Defect and Severe Primary Pulmonary Hypertension. *American Journal of Case Reports* **20,** 354–360 (2019).

127. Pasutto, F., Flinter, F., Rauch, A. & Reis, A. Novel STRA6 null mutations in the original family described with Matthew–Wood syndrome. *Am J Med Genet A* **176,** 134–138 (2018).

128. Seller, M. J. *et al.* Two sibs with anophthalmia and pulmonary hypoplasia (the Matthew-Wood syndrome). *Am J Med Genet* **62,** 227–229 (1996).

129. Marcadier, J. L. *et al.* A novel mutation in two Hmong families broadens the range of STRA6-related malformations to include contractures and camptodactyly. *Am J Med Genet A* **170,** 11–18 (2016).

130. Gudas, L. J. Retinoid metabolism: new insights. *J Mol Endocrinol* **69,** T37–T49 (2022).

131. Beecroft, S. J. *et al.* Biallelic hypomorphic variants in ALDH1A2 cause a novel lethal human multiple congenital anomaly syndrome encompassing diaphragmatic, pulmonary, and cardiovascular defects. *Hum Mutat* **42,** 506–519 (2021).

132. Leon, E., Nde, C., Ray, R. S., Preciado, D. & Zohn, I. E. ALDH1A2-related disorder: A new genetic syndrome due to alteration of the retinoic acid pathway. *Am J Med Genet A* **191,** 90–99 (2023).

133. Clugston, R. D., Zhang, W. & Greer, J. J. Gene expression in the developing diaphragm: significance for congenital diaphragmatic hernia. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **294,** L665–L675 (2008).

134. Dollé, P. Developmental expression of retinoic acid receptors (RARs). *Nucl Recept Signal* **7,** nrs.07006 (2009).

135. Srour, M. *et al.* Recessive and Dominant Mutations in Retinoic Acid Receptor Beta in Cases with Microphthalmia and Diaphragmatic Hernia. *The American Journal of Human Genetics* **93,** 765–772 (2013).

136. Nobile, S., Pisaneschi, E., Novelli, A. & Carnielli, V. P. A rare mutation of retinoic acid receptor-β associated with lethal neonatal Matthew-Wood syndrome. *Clin Dysmorphol* **28,** (2019).

137. Srour, M. *et al.* Gain-of-Function Mutations in RARB Cause Intellectual Disability with Progressive Motor Impairment. *Hum Mutat* **37,** 786–793 (2016).

138. Holder, A. M. *et al.* Genetic Factors in Congenital Diaphragmatic Hernia. *The American Journal of Human Genetics* **80,** 825–845 (2007).

139. Miller-Hodges, E. & Hohenstein, P. WT1 in disease: shifting the epithelial–mesenchymal balance. *J Pathol* **226,** 229–240 (2012).

140. Klaassens, M. *et al.* Congenital Diaphragmatic Hernia and Chromosome 15q26: Determination of a Candidate Region by Use of Fluorescent In Situ Hybridization and Array-Based Comparative Genomic Hybridization. *The American Journal of Human Genetics* **76,** 877–882 (2005).

141. High, F. A. *et al.* De novo frameshift mutation in COUP-TFII (NR2F2) in human congenital diaphragmatic hernia. *Am J Med Genet A* **170,** 2457–2461 (2016).

142. Yu, L. *et al.* Variants in GATA4 are a rare cause of familial and sporadic congenital diaphragmatic hernia. *Hum Genet* **132,** 285–292 (2013).

143. Longoni, M. *et al.* Prevalence and penetrance of ZFPM2 mutations and deletions causing congenital diaphragmatic hernia. *Clin Genet* **87,** 362–367 (2015).

144. Kammoun, M. *et al.* Genetic profile of isolated congenital diaphragmatic hernia revealed by targeted next-generation sequencing. *Prenat Diagn* **38,** 654–663 (2018).

145. Francisco, T., Gonçalves, R. M., Borges, C. & Neto, M. T. Multiple haemangiomas, diaphragmatic eventration and Beckwith-Wiedemann syndrome: an unusual association. *BMJ Case Rep* **2013,** bcr2013010077 (2013).

146. Cohen Jr, M. M., Gorlin, R. J., Feingold, M. & ten Bensel, R. W. The Beckwith-Wiedemann syndrome. Seven new cases. *Am J Dis Child* **122,** 515–519 (1971).

147. Huang, N. *et al.* Two distinct nuclear receptor interaction domains in NSD1, a novel SET protein that exhibits characteristics of both corepressors and coactivators. *EMBO J* **17,** 3398–3412 (1998).

148. Hong, S.-H., David, G., Wong, C.-W., Dejean, A. & Privalsky, M. L. SMRT corepressor interacts with PLZF and with the PML-retinoic acid receptor α (RARα) and PLZF-RARα oncoproteins associated with acute promyelocytic leukemia. *Proceedings of the National Academy of Sciences* **94,** 9028–9033 (1997).

149. Longoni, M. *et al.* Molecular pathogenesis of congenital diaphragmatic hernia revealed by exome sequencing, developmental data, and bioinformatics. *Proceedings of the National Academy of Sciences* **111,** 12450–12455 (2014).

150. Coles, G. L. & Ackerman, K. G. Kif7 is required for the patterning and differentiation of the diaphragm in a model of syndromic congenital diaphragmatic hernia. *Proceedings of the National Academy of Sciences* **110,** E1898–E1905 (2013).

151. Gupta, N. *et al.* Nasopharyngeal teratoma, congenital diaphragmatic hernia and Dandy–Walker malformation – a yet uncharacterized syndrome. *Clin Genet* **90,** 470–471 (2016).

152. Yang, W. *et al.* Nutrient intakes in women and congenital diaphragmatic hernia in their offspring. *Birth Defects Res A Clin Mol Teratol* **82,** 131–138 (2008).

153. Carmichael, S. L. *et al.* Congenital diaphragmatic hernia and maternal dietary nutrient pathways and diet quality. *Birth Defects Res* **112,** 1475–1483 (2020).

154. Bailey, R. L., Pac, S. G., Fulgoni 3rd, V. L., Reidy, K. C. & Catalano, P. M. Estimation of Total Usual Dietary Intakes of Pregnant Women in the United States. *JAMA Netw Open* **2,** e195967 (2019).

155. Jankowska, A. *et al.* Determinants of the Essential Elements and Vitamins Intake and Status during Pregnancy: A Descriptive Study in Polish Mother and Child Cohort. *Nutrients* **13,** (2021).

156. Hanson, C., Lyden, E., Abresch, C. & Anderson-Berry, A. Serum Retinol Concentrations, Race, and Socioeconomic Status in of Women of Childbearing Age in the United States. *Nutrients* **8,** (2016).

157. Caspers, K. M. *et al.* Maternal periconceptional exposure to cigarette smoking and alcohol consumption and congenital diaphragmatic hernia. *Birth Defects Res A Clin Mol Teratol* **88,** 1040–1049 (2010).

158. Molotkov, A. & Duester, G. Retinol/Ethanol Drug Interaction during Acute Alcohol Intoxication in Mice Involves Inhibition of Retinol Metabolism to Retinoic Acid by Alcohol Dehydrogenase*. *Journal of Biological Chemistry* **277,** 22553–22557 (2002).

159. Yılmaz, G. *et al.* The effect of passive smoking and breast feeding on serum antioxidant vitamin (A, C, E) levels in infants. *Acta Paediatr* **98,** 531–536 (2009).

160. Fernandes-Silva, H., Araújo-Silva, H., Correia-Pinto, J. & Moura, R. S. Retinoic Acid: A Key Regulator of Lung Development. *Biomolecules* **10,** (2020).

161. Marquez, H. A. & Chen, F. in *The Biochemistry of Retinoid Signaling III: Vitamin A and Retinoic Acid in Embryonic Development* (eds. Asson-Batres, M. A. & Rochette-Egly, C.) 151–174 (Springer International Publishing, 2020). doi:10.1007/978-3-030-42282-0_6

162. Malpel, S., Mendelsohn, C. & Cardoso, W. V. Regulation of retinoic acid signaling during lung morphogenesis. *Development* **127,** 3057–3067 (2000).

163. Keijzer, R., Liu, J., Deimling, J., Tibboel, D. & Post, M. Dual-Hit Hypothesis Explains Pulmonary Hypoplasia in the Nitrofen Model of Congenital Diaphragmatic Hernia. *Am J Pathol* **156,** 1299–1306 (2000).

164. Montalva, L. & Zani, A. Assessment of the nitrofen model of congenital diaphragmatic hernia and of the dysregulated factors involved in pulmonary hypoplasia. *Pediatr Surg Int* **35,** 41–61 (2019).

165. Gallot, D. *et al.* Congenital diaphragmatic hernia: A retinoid-signaling pathway disruption during lung development? *Birth Defects Res A Clin Mol Teratol* **73,** 523–531 (2005).

166. Coste, K. *et al.* Metabolic disturbances of the vitamin A pathway in human diaphragmatic hernia. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **308,** L147–L157 (2014).

167. Verla, M. A., Style, C. C. & Olutoye, O. O. Prenatal intervention for the management of congenital diaphragmatic hernia. *Pediatr Surg Int* **34,** 579–587 (2018).

168. Delabaere, A. *et al.* Retinoic acid and tracheal occlusion for diaphragmatic hernia treatment in rabbit fetuses. *Prenat Diagn* **38,** 482–492 (2018).

169. Kirby, E. & Keijzer, R. Congenital diaphragmatic hernia: current management strategies from antenatal diagnosis to long-term follow-up. *Pediatr Surg Int* **36,** 415–429 (2020).

170. Qiao, L. *et al.* Rare and de novo variants in 827 congenital diaphragmatic hernia probands implicate LONP1 as candidate risk gene. *The American Journal of Human Genetics* **108,** 1964–1980 (2021).

171. Beurskens, L. W. J. E., de Jonge, R., Schoonderwaldt, E. M., Tibboel, D. & Steegers-Theunissen, R. P. M. Biomarkers of the one-carbon pathway in association with congenital diaphragmatic hernia. *Birth Defects Res A Clin Mol Teratol* **94,** 557–560 (2012).

172. Adrien, N. *et al.* Early pregnancy vitamin D status and risk of select congenital anomalies in the National Birth Defects Prevention Study. *Birth Defects Res* **115,** 290–301 (2023).

173. Turkmen, G. G. *et al.* Levels of serum vitamin D and calcium in pregnancies complicated with fetal congenital diaphragmatic hernia and normal pregnancies. *The Journal of Maternal-Fetal & Neonatal Medicine* **30,** 990–994 (2017).

174. Antounians, L. *et al.* Fetal lung underdevelopment is rescued by administration of amniotic fluid stem cell extracellular vesicles in rodents. *Sci Transl Med* **13,** eaax5941 (2021).

175. Bogenschutz, E. L., Sefton, E. M. & Kardon, G. Cell culture system to assay candidate genes and molecular pathways implicated in congenital diaphragmatic hernias. *Dev Biol* **467,** 30–38 (2020).

176. Goumy, C. *et al.* Fetal skin fibroblasts: A cell model for studying the retinoid pathway in congenital diaphragmatic hernia. *Birth Defects Res A Clin Mol Teratol* **88,** 195–200 (2010).

177. Perry, S. F., Similowski, T., Klein, W. & Codd, J. R. The evolutionary origin of the mammalian diaphragm. *Respir Physiol Neurobiol* **171,** 1–16 (2010).

178. Iritani, I. Experimental study on embryogenesis of congenital diaphragmatic hernia. *Anat Embryol (Berl)* **169,** 133–139 (1984).

179. Babiuk, R. P., Zhang, W., Clugston, R., Allan, D. W. & Greer, J. J. Embryological origins and development of the rat diaphragm. *Journal of Comparative Neurology* **455,** 477–487 (2003).

180. Allan, D. W. & Greer, J. J. Embryogenesis of the phrenic nerve and diaphragm in the fetal rat. *Journal of Comparative Neurology* **382,** 459–468 (1997).

181. Kardon, G. *et al.* Congenital diaphragmatic hernias: from genes to mechanisms to therapies. *Dis Model Mech* **10,** 955–970 (2017).

182. Mesas Burgos, C., Ehrén, H., Conner, P. & Frenckner, B. Maternal Risk Factors and Perinatal Characteristics in Congenital Diaphragmatic Hernia: A Nationwide Population-Based Study. *Fetal Diagn Ther* **46,** 385–391 (2019).

183. Mey, J., Babiuk, R. P., Clugston, R., Zhang, W. & Greer, J. J. Retinal Dehydrogenase-2 Is Inhibited by Compounds that Induce Congenital Diaphragmatic Hernias in Rodents. *Am J Pathol* **162,** 673–679 (2003).

184. Clugston, R. D. *et al.* Teratogen-Induced, Dietary and Genetic Models of Congenital Diaphragmatic Hernia Share a Common Mechanism of Pathogenesis. *Am J Pathol* **169,** 1541–1549 (2006).

185. Gosemann, J.-H. *et al.* Pax3 gene expression is not altered during diaphragmatic development in nitrofen-induced congenital diaphragmatic hernia. *J Pediatr Surg* **47,** 1067–1071 (2012).

186. Nie, X. *et al.* Clustering ensemble in scRNA-seq data analysis: Methods, applications and challenges. *Comput Biol Med* **159,** 106939 (2023).

187. Hao, Y. *et al.* Integrated analysis of multimodal single-cell data. *Cell* **184,** 3573-3587.e29 (2021).

188. Franzén, O., Gan, L.-M. & Björkegren, J. L. M. PanglaoDB: a web server for exploration of mouse and human single-cell RNA sequencing data. *Database* **2019,** baz046 (2019).

189. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat Genet* **25,** 25–29 (2000).

190. Consortium, T. G. O. *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224,** iyad031 (2023).

191. Cao, J. *et al.* The single-cell transcriptional landscape of mammalian organogenesis. *Nature* **566,** 496–502 (2019).

192. Peng, L. *et al.* Single-cell RNA-seq clustering: datasets, models, and algorithms. *RNA Biol* **17,** 765–783 (2020).

193. Yu, L., Qiu, C. & Chen, R. A narrative review of research advances in the study of molecular markers of airway smooth muscle cells. *Ann Transl Med* **10,** 375 (2022).

194. Taniguchi, T. *et al.* Regulation of Mesothelial Cell Fate during Development and Human Diseases. *Int J Mol Sci* **23,** (2022).

195. Kanamori-Katayama, M. *et al.* LRRN4 and UPK3B Are Markers of Primary Mesothelial Cells. *PLoS One* **6,** e25391- (2011).

196. Yan, J. *et al.* A series of robust genetic indicators for definitive identification of cardiomyocytes. *J Mol Cell Cardiol* **97,** 278–285 (2016).

197. Pusztaszeri, M. P., Seelentag, W. & Bosman, F. T. Immunohistochemical Expression of Endothelial Markers CD31, CD34, von Willebrand Factor, and Fli-1 in Normal Human Tissues. *Journal of Histochemistry & Cytochemistry* **54,** 385–395 (2006).

198. Jain, V. *et al.* Single Cell RNA-Seq Analysis of Human Red Cells. *Front Physiol* **13,** (2022).

199. Callaway Ian M.; Stover Samantha R.; Hernandez-Garcia Andres; Jhangiani Shalini N.; Punetha Jaya; Paine Ingrid S.; Posey Jennifer E.; Muzny Donna; Lally Kevin P.; Lupski James R.; Shaw Chad A.; Fernandes Caraciolo J.; Scott Daryl A., D. A. ; C. Prioritization of Candidate Genes for Congenital Diaphragmatic Hernia in a Critical Region on Chromosome 4p16 using a Machine-Learning Algorithm. *J Pediatr Genet* **07,** 164–173 (2018).

200. Thompson, B. *et al.* Genetics and functions of the retinoic acid pathway, with special emphasis on the eye. *Hum Genomics* **13,** 61 (2019).

201. Duester, G. Retinoic Acid Synthesis and Signaling during Early Organogenesis. *Cell* **134,** 921–931 (2008).

202. Allan, D. W. & Greer, J. J. Pathogenesis of nitrofen-induced congenital diaphragmatic hernia in fetal rats. *J Appl Physiol* **83,** 338–347 (1997).

203. Wagner, R., Montalva, L., Zani, A. & Keijzer, R. Basic and translational science advances in congenital diaphragmatic hernia. *Semin Perinatol* **44,** 151170 (2020).

204. Wynn, J. *et al.* Outcomes of Congenital Diaphragmatic Hernia in the Modern Era of Management. *J Pediatr* **163,** 114-119.e1 (2013).

205. Mehollin-Ray, A. R. Congenital diaphragmatic hernia. *Pediatr Radiol* **50,** 1855–1871 (2020).

206. Babic, I. *et al.* Prenatal Diagnosis and Management of a Rare Central Tendon Defect Type of Congenital Diaphragmatic Hernia with a Massive Pericardial Effusion. *Case Rep Obstet Gynecol* **2020,** 6798253 (2020).

207. Zani, A. *et al.* Congenital diaphragmatic hernia. *Nat Rev Dis Primers* **8,** 37 (2022).

208. Antonius, T. *et al.* Denys–Drash syndrome and congenital diaphragmatic hernia: Another case with the 1097G > A(Arg366His) mutation. *Am J Med Genet A* **146A,** 496–499 (2008).

209. Coste, K. *et al.* Metabolic disturbances of the vitamin A pathway in human diaphragmatic hernia. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **308,** L147–L157 (2014).

210. Clugston, R. D., Zhang, W. & Greer, J. J. Early development of the primordial mammalian diaphragm and cellular mechanisms of nitrofen-induced congenital diaphragmatic hernia. *Birth Defects Res A Clin Mol Teratol* **88,** 15–24 (2010).

211. Billings, S. E. *et al.* The retinaldehyde reductase DHRS3 is essential for preventing the formation of excess retinoic acid during embryonic development. *FASEB J* **27,** 4877–4889 (2013).

212. Tsao, K. & Lally, K. P. The Congenital Diaphragmatic Hernia Study Group: a voluntary international registry. *Semin Pediatr Surg* **17,** 90–97 (2008).

213. Babicki, S. *et al.* Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res* **44,** W147–W153 (2016).

214. Montalva, L. & Zani, A. Assessment of the nitrofen model of congenital diaphragmatic hernia and of the dysregulated factors involved in pulmonary hypoplasia. *Pediatr Surg Int* **35,** 41–61 (2019).

215. Costlow, R. D. & Manson, J. M. The heart and diaphragm: Target organs in the neonatal death induced by nitrofen (2,4-dichlorophenyl-p-nitrophenyl ether). *Toxicology* **20,** 209–227 (1981).

216. Wang, W. *et al.* All-Trans Retinoic Acid-Induced Craniofacial Malformation Model: A Prenatal and Postnatal Morphological Analysis. *The Cleft Palate Craniofacial Journal* **54,** 391–399 (2017).

217. Hong, Q., Li, X.-D., Xie, P. & Du, S.-X. All-trans-retinoic acid suppresses rat embryo hindlimb bud mesenchymal chondrogenesis by modulating HoxD9 expression. *Bioengineered* **12,** 3900–3911 (2021).

218. Abbott, B. D. & Birnbaum, L. S. Retinoic acid-induced alterations in the expression of growth factors in embryonic mouse palatal shelves. *Teratology* **42,** 597–610 (1990).

219. Sydorak, R. M. *et al.* Congenital diaphragmatic hernia and hydrops: A lethal association? *J Pediatr Surg* **37,** 1678–1680 (2002).

220. Pelizzo, G. *et al.* Liver pathological alterations in fetal rabbit model of congenital diaphragmatic hernia. *Congenit Anom (Kyoto)* **62,** 105–112 (2022).

221. PAUWELS, E., VAN LOO, D., CORNILLIE, P., BRABANT, L. & VAN HOOREBEKE, L. An exploratory study of contrast agents for soft tissue visualization by means of high resolution X-ray computed tomography imaging. *J Microsc* **250,** 21–31 (2013).

222. Phithakwatchara, N. *et al.* Differential patterns of prenatal ipsilateral and contralateral lung growth in cases of isolated left-sided congenital diaphragmatic hernia. *Prenat Diagn* **35,** 769–776 (2015).

223. Weis, M. *et al.* Region of interest-based versus whole-lung segmentation-based approach for MR lung perfusion quantification in 2-year-old children after congenital diaphragmatic hernia repair. *Eur Radiol* **26,** 4231–4238 (2016).

224. Bargy, F., Beaudoin, S. & Barbet, P. Fetal Lung Growth in Congenital Diaphragmatic Hernia. *Fetal Diagn Ther* **21,** 39–44 (2005).

225. Fleming, J. *et al.* Determination of regional lung air volume distribution at mid-tidal breathing from computed tomography: a retrospective study of normal variability and reproducibility. *BMC Med Imaging* **14,** 25 (2014).

226. Rhinn, M. & Dollé, P. Retinoic acid signalling during development. *Development* **139,** 843–858 (2012).

227. Niederreither, K., Vermot, J., Fraulob, V., Chambon, P. & Dollé, P. Retinaldehyde dehydrogenase 2 (RALDH2)- independent patterns of retinoic acid synthesis in the mouse embryo. *Proceedings of the National Academy of Sciences* **99,** 16111–16116 (2002).

228. Feng, L., Hernandez, R. E., Waxman, J. S., Yelon, D. & Moens, C. B. Dhrs3a regulates retinoic acid biosynthesis through a feedback inhibition mechanism. *Dev Biol* **338,** 1–14 (2009).

229. Adams, M. K., Belyaeva, O. V, Wu, L. & Kedishvili, N. Y. The retinaldehyde reductase activity of DHRS3 is reciprocally activated by retinol dehydrogenase 10 to control retinoid homeostasis. *J Biol Chem* **289,** 14868–14880 (2014).

230. Wang, S. *et al.* Alterations in retinoic acid signaling affect the development of the mouse coronary vasculature. *Dev Dyn* **247,** 976–991 (2018).

231. Shannon, S. R., Moise, A. R. & Trainor, P. A. New insights and changing paradigms in the regulation of vitamin A metabolism in development. *Wiley Interdiscip Rev Dev Biol* **6,** (2017).

232. Gowda Neil, S. H. ; P. "Heart of the Matter": Cardiac Dysfunction in Congenital Diaphragmatic Hernia. *Am J Perinatol* (2023). doi:10.1055/a-2067-7925

233. Tydén, K. Ö., Nordenstam, F., Frenckner, B. & Burgos, C. M. Hidden cardiovascular morbidity in children and young adults born with congenital diaphragmatic hernia: A population-based study. *J Pediatr Surg* **57,** 510–515 (2022).

234. Andersen, D. H. Effect of diet during pregnancy upon the incidence of congenital hereditary diaphragmatic hernia in the rat; failure to produce cystic fibrosis of the pancreas by maternal vitamin A deficiency. *Am J Pathol* **25,** 163–185 (1949).

235. Andersen, D. H. Incidence of congenital diaphragmatic hernia in the young of rats bred on a diet deficient in vitamin A. *Am J Dis Child* **62,** 888–889 (1941).

236. Snyder, A. N., Cheng, T. & Burjonrappa, S. A nationwide database analysis of demographics and outcomes related to Extracorporeal Membrane Oxygenation (ECMO) in congenital diaphragmatic hernia. *Pediatr Surg Int* **37,** 1505–1513 (2021).

237. Lam, J. C., Claydon, J., Mitton, C. R. & Skarsgard, E. D. A risk-adjusted study of outcome and resource utilization for congenital diaphragmatic hernia. *J Pediatr Surg* **41,** 883–887 (2006).