Remarks prepared by Dr. Pawel Stankiewicz, a Professor of human genetics at Baylor College of Medicine in Houston. For the last 13 years, his research laboratory has been working on Alveolar Capillary Dysplasia (ACD), a lethal lung developmental disorders in neonates.

The current state and summary of developments in genetic research on alveolar capillary dysplasia

Throughout embryonic and fetal life, human lungs undergo significant morphological chnages that are necessary to maintain respiration at birth. These complex developmental processes involve coordinated action of multiple signaling pathways. Disruption of intrauterine lung development can result in terrible lethal lung developmental disorders, one of which is called Alveolar capillary dysplasia with misalignment of pulmonary veins (ACDMPV).

ACDMPV is a rarely diagnosed condition (~ 1/100,000 births) disrupting the blood-air barrier between alveoli and capillaries - that facilitates exchange of oxygen and carbon dioxide. This irreversible lung malformation in the vast majority of cases is fatal in the newborn period and currently no cure other than lung transplant can be offered.

Newborns affected with ACDMPV usually are born at term with normal birth weight, but within the first 24-48 hours of life develop progressive hypoxemic respiratory failure and severe pulmonary hypertension and almost uniformly pass away within the first month of life. ACDMPV is often associated with additional malformations of the cardiovascular, gastrointestinal, or genitourinary systems. Very rarely, some of the ACDMPV neonates can live a few months and even years – they are so called long survivors with milder or late onset ACD. We think we now know why some of the ACDMPV infants can survive longer.

In 2009, we identified that loss-of-function due to point mutations (single nucleotide variants, SNVs) or genomic deletions (also referred to as copy-number variant, CNV deletions) involving the *FOXF1* gene locus on chromosome 16q24.1 in 80-90% of patients with ACDMPV. *FOXF1* is primarily expressed in mesoderm-derived tissues where it mediates signaling from epithelial cells of the developing airways - so called epithelial-mesenchymal transition.

Supporting these findings, mouse model of *Foxf1* deficiency developed in 2001 exhibited lung immaturity resembling ACDMPV.

Since our first study in 2009, here at Baylor College of Medicine in Texas Medical Center in Houston, we have collected samples from almost 200 ACDMPV families. To date, greater than 70 point mutations and 60 deletions involving the *FOXF1* region have been reported.

Approximately 50% of neonates with ACDMPV have a point mutations in *FOXF1* whereas the other half have deletions in the *FOXF1* locus. Half of these deletions, so 25% of all cases, involve *FOXF1*, but the other half leave *FOXF1* intact but encompass its regulatory region - so called enhancer that is located far away 300 kb upstream to *FOXF1*.

Interestingly, unlike point mutations, the causative genomic deletions involving *FOXF1* or its upstream regulatory enhancer regions arise almost exclusively *de novo* (absent in parents) on the maternal chromosome 16, suggesting that the *FOXF1* locus may be genomically imprinted. CNV deletions on the paternal chromosome are thought to be prenatally lethal due to congenital anomalies of other organs.

We were able to narrow the *FOXF1* enhancer region from ~60 kb to ~15 kb, and finally to ~5 kb core interval, loss of which results in full phenotype of lethal ACDMPV. Recently, in this narrowed ~5 kb interval, using whole genome sequencing in ACDMPV families with an unexpectedly mitigated ACDMPV phenotype, we identified rare non-coding single nucleotide polymorphisms located on the other chromosome 16 (in *trans*) to the causative point mutations or genomic deletions involving *FOXF1* that we think might have acted on the *FOXF1* enhancer as so called hypermorphs, likely preventing the ACDMPV lethality.

Importantly, in a few ACD families, two or even more children were born with ACDMPV. We found that in such cases, one of the parent was a carrier of a somatic mosaic mutation in the *FOXF1* gene. In some families, this mosaicism was so low that it was not detectable by routine diagnostic methods. We developed and applied much more sensitive techniques to detect this low-level somatic mosaicism.

The *FOXF1* gene locus harbors also long noncoding RNAs (lncRNAs) that can control biological functions by regulating expression of many protein-coding genes. One of these long non-coding RNAs is called *FENDRR* - adjacent and divergent to *FOXF1*. Interestingly, mice with deficient in the *Fendrr* gene and intact *Foxf1* gene, developed hypoplastic lungs similar to those found in *Foxf1* deficient mice and infants with ACDMPV. Most recently, we found that lung-specific *FOXF1* enhancer is also essential for the expression of *FENDRR*, suggesting that this lncRNA may play a role in etiology of ACDMPV.

Historically, ACDMPV diagnosis was based on the histopathological appearance at lung biopsy or autopsy. Today, genetic study is important part complementing ACDMPV diagnostic process. The optimal material for molecular testing of newborn with suspected ACD is DNA extracted from fresh blood sample. However, DNA can also be isolated from tissue obtained at lung biopsy or autopsy. Post mortem specimens should be collected in the shortest possible time to provide best DNA quality, ensuring the most accurate genetic results. Lung tissue could be fresh frozen or formalin-fixed and paraffin-embedded. However, due to the significant degradation of

DNA or RNA, fresh frozen tissue provides better quality of nucleic acids, which is very important in gene expression studies.

If sequence analysis of the coding portion of *FOXF1* is negative, chromosomal microarray analysis using for example array comparative genomic hybridization targeting the *FOXF1* locus and specifically its two exons should be performed to search for CNV deletions involving *FOXF1* and/or its enhancer. If both of these molecular tests are negative, exome sequencing (ES) or whole genome sequencing (WGS) should be considered to investigate the possible involvement of other coding and/or non-coding variants.

With the constantly decreasing costs of next generation sequencing technologies, WGS is thought to replace other methods currently used in routine diagnostics, enabling identification of all clinically relevant protein-coding as well as non-coding variants in lung developmental genes or lung-specific genes enhancer regions.