The discovery of novel mechanisms for lymphangioleiomyomatosis pathogenesis through GWAS: a rarity in rare respiratory disorders

Arnold S. Kristof[1] and Victor E. Ortega[2]

Affiliations: [1]Meakins-Christie Laboratories and Translational Research in Respiratory Diseases Program, Research Institute of the McGill University Health Centre, Depts of Medicine and Critical Care, Montreal, QC, Canada. [2]Dept of Internal Medicine, Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, NC, USA.

Correspondence: Victor E. Ortega, Center for Precision Medicine, Wake Forest School of Medicine, Medical Center Boulevard, Winston-Salem, NC 27157, USA. E-mail: vortega@wakehealth.edu

@ERSpublications
This is the largest genome-wide association study of risk for sporadic LAM, a rare genetic disorder, from an international effort which identified novel variants for LAM pathogenesis in a plausible gene, independent of variation in *TSC1* and *TSC2* http://bit.ly/2X0SeZ1

Cite this article as: Kristof AS, Ortega VE. The discovery of novel mechanisms for lymphangioleiomyomatosis pathogenesis through GWAS: a rarity in rare respiratory disorders. *Eur Respir J* 2019; 53: 1900863 [https://doi.org/10.1183/13993003.00863-2019].

Lymphangioleiomyomatosis (LAM) is a rare respiratory disorder primarily of women that can occur sporadically (S-LAM), or as a manifestation of tuberous sclerosis complex (TSC) [1]. TSC is a syndrome of neurodevelopmental disorders and tumours in multiple organs, caused by heterozygous germline variants in the *TSC1* or *TSC2* genes, which can be inherited in autosomal dominant fashion or which can occur *de novo* [2, 3]. The *TSC1* and *TSC2* gene products, hamartin (TSC1) and tuberin (TSC2), constitute a heterodimeric suppressor of the kinase "mechanistic target of rapamycin complex 1" (mTORC1). In both TSC- and S-LAM, tumours arise *via* somatic loss of the functional "wild-type" allele (*i.e.* loss of heterozygosity) at the *TSC1* or *TSC2* locus, resulting in excessive mTOR activity, uninhibited cell growth, and unchecked neoplastic behaviour. In individuals with LAM, microscopic pulmonary interstitial tumours consisting of *TSC2*-deficient "LAM cells" progressively destroy the lung, culminating in death or lung transplantation.

As potent inhibitors of the mTORC1 signalling pathway, the drug rapamycin and its analogues can be used in patients with LAM to reduce the growth of tumours and prevent deterioration in respiratory function [4, 5]. Curative or preventive approaches for LAM, however, await a better understanding of the molecular pathways underlying disease susceptibility and progression. One major gap in the field arises from a lack of clarity regarding the genomic and epigenomic factors that contribute to S-LAM tumour initiation and progression, especially since these somatic tumours arise in the absence of germline *TSC2* mutations. It is possible that germline genetic variation independent of *TSC2* drives other distinctive phenotypic features of S-LAM tumours, including neural crest differentiation, sensitivity to female sex hormones, excessive angio- and lymphangiogenesis, or predisposition to loss of *TSC2* function.

Received: April 30 2019 | Accepted after revision: May 15 2019

In this issue of the *European Respiratory Journal*, Kim *et al.* [6] report the results of a genome-wide association study (GWAS) which tested over 5.4 million single nucleotide polymorphisms (SNPs) for the identification of LAM risk loci in 426 white females with sporadic LAM from Europe and the USA, and 852 matched controls. This GWAS identified a signal on chromosome 15q26.2 that reached genome-wide significance adjacent to the gene encoding *NR2F2*, which was successfully replicated in an independent cohort. Fine mapping of the genomic regions represented by the top SNP associations identified a possible causative locus associated with LAM while gene-based SKAT-O testing, which covers lower frequency variants, only found significant associations in *NR2F2*.

This is the largest GWAS of sporadic LAM cases and unaffected controls assembled to date from an international effort which identified novel gene variation for disease risk, independent of variation in *TSC1* and *TSC2*. These investigators also leveraged multi-omic databases for epidemiological and *in silico* analyses which demonstrated that 1) the associated SNPs were within a genomic region containing regulatory elements for *NR2F2* and none of the other neighbouring genes, 2) the minor allele frequencies of these *NR2F2* variants in these US and European sporadic LAM cohorts were not only significantly lower than the controls from COPDGene (the basis of the GWAS case–control design) but also lower compared to seven additional general populations, and 3) NR2F2 protein expression was higher in LAM-affected tissues compared to other cancer and normal tissue. In addition, immunohistochemistry showed strong nuclear expression of *NR2F2* in LAM lung and renal angiomyolipoma cells. The ability to replicate their novel GWAS associations in combination with the use of bioinformatic resources and molecular phenotyping approaches on LAM-affected tissues suggests that *NR2F2* is a novel candidate gene for LAM pathogenesis.

The biological relevance of *NR2F2* as a candidate gene for LAM susceptibility provides further supportive evidence. *NR2F2* encodes an orphan receptor in the nuclear receptor superfamily of ligand-activated receptors [7]. The known biological activities of *NR2F2* fit with several pathogenic features of LAM. Consistent with a role in tumour progression and invasiveness, *NR2F2* is a prognostic indicator in a variety of cancers, especially those that are sex hormone-responsive [7, 8]. *NR2F2* is also an important modifier of tumour angio- and lymphangiogenesis; in fact, *NR2F2* knockout mice die *in utero* from aberrant vascular development [8]. Finally, consistent with the expression of other neural crest lineage markers in LAM cells, *NR2F2* is a chromatin-binding modifier of neural crest gene expression and lineage specification [9, 10].

The significant genetic association between *NR2F2* variation and LAM susceptibility, its expression in LAM tumours, and its known biological properties, altogether make a compelling case for *NR2F2* variation as risk loci for sporadic LAM, independent of *TSC1* and *TSC2* variation. These genetic studies by Kim *et al.* [6] nonetheless raise several questions for future investigation. First, this study used genotyping chip complemented by imputed genotype data which could have missed *TSC1* and *TSC2* pathogenic variation. Second, this GWAS did not identify pathogenic variation in these genes and was not able to evaluate gene–gene interactions with *NR2F2* variation. Third, this GWAS compared allele frequencies between LAM cases and unaffected controls to identify risk loci, but genetic studies limited to LAM cases will be necessary to determine if this *NR2F2* variant determines disease progression and severity. Fourth, the study was not designed to compare *NR2F2* mRNA expression in tumours *versus* normal tissues from the same LAM cases, and levels of expression were not tested for correlation with the risk genotypes found by GWAS. Finally, NR2F2 protein was equally detected in many cell types within the tumours, raising questions regarding the non-cell autonomous effects by which NR2F2 dysregulation might affect the development of LAM tumours.

Future studies aimed at developing new mechanistic hypotheses might evaluate the co-localisation of *NR2F2* with known LAM cell markers (*e.g.* HMB45), sex hormone receptors, and regulators of angio- and lymphangiogenesis in tumour cells, and compare its expression in lung structures peripheral to tumour micronodules. While such studies are limited by availability of appropriate tissues, more detailed genotype–phenotype correlations could provide insights into the specific mechanisms which link *NR2F2* variants to TSC tumorigenesis. In addition, recent developments in preclinical models of LAM (*i.e.* induced pluripotent stem cell technology [11], emerging murine models) might be exploited to investigate the mechanism(s) by which *NR2F2* interacts with other oncogenic mechanisms, including the loss of *TSC2*, mTOR hyperactivity, sex hormone responsiveness, and neural crest behaviour, that drive disease risk and progression.

Over the past decade, GWAS have been increasing in size and number for common, genetically complex traits and respiratory diseases such as asthma. These studies have provided novel insights into biological pathways that drive disease risk and severity and, in a minority of cases, have identified therapeutic targets, such as *TSLP* for severe asthma [12, 13]. The requirement for large sample sizes to detect novel genetic

signals is the primary reason why there are few GWAS for rare genetic diseases caused by rare pathogenic variants with strong effects (*i.e*. *TSC1* and *TSC2* variation). As clearly demonstrated by these investigators, productive, international collaborative efforts that include expert physicians, scientists, patients and funding agencies can result in improved power and the potential for novel variant discovery. In this case, such efforts resulted in the identification of a novel, highly plausible gene locus for sporadic LAM risk achieving both genome-wide significance and replication, a rarity in rare respiratory disorders.

Support statement: This work was supported by Canadian Institutes of Health Research Grant PJT 155971 and United States National Institutes of Health grant R01 HL142992. Funding information for this article has been deposited with the Crossref Funder Registry.

Conflict of interest: None declared.

References

1 Moss J, Avila NA, Barnes PM, *et al.* Prevalence and clinical characteristics of lymphangioleiomyomatosis (LAM) in patients with tuberous sclerosis complex. *Am J Respir Crit Care Med* 2001; 164: 669–671.
2 Lam HC, Nijmeh J, Henske EP. New developments in the genetics and pathogenesis of tumours in tuberous sclerosis complex. *J Pathol* 2017; 241: 219–225.
3 Peron A, Au KS, Northrup H. Genetics, genomics, and genotype-phenotype correlations of TSC: Insights for clinical practice. *Am J Med Genet C Semin Med Genet* 2018; 178: 281–290.
4 McCormack FX, Inoue Y, Moss J, *et al.* Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *N Engl J Med* 2011; 364: 1595–1606.
5 Sasongko TH, Ismail NF, Zabidi-Hussin Z. Rapamycin and rapalogs for tuberous sclerosis complex. *Cochrane Database Syst Rev* 2016; 7: CD011272.
6 Kim W, Giannikou K, Dreier JR, *et al.* A genome-wide association study implicates *NR2F2* in lymphangioleiomyomatosis pathogenesis. *Eur Respir J* 2019; 53: 1900329.
7 Safe S, Jin UH, Hedrick E, *et al.* Minireview: role of orphan nuclear receptors in cancer and potential as drug targets. *Mol Endocrinol* 2014; 28: 157–172.
8 Qin J, Tsai SY, Tsai MJ. The critical roles of COUP-TFII in tumor progression and metastasis. *Cell Biosci* 2014; 4: 58.
9 Rada-Iglesias A, Bajpai R, Prescott S, *et al.* Epigenomic annotation of enhancers predicts transcriptional regulators of human neural crest. *Cell Stem Cell* 2012; 11: 633–648.
10 Rada-Iglesias A, Prescott SL, Wysocka J. Human genetic variation within neural crest enhancers: molecular and phenotypic implications. *Philos Trans R Soc Lond B Biol Sci* 2013; 368: 20120360.
11 Julian LM, Delaney SP, Wang Y, *et al.* Human pluripotent stem cell-derived TSC2-haploinsufficient smooth muscle cells recapitulate features of lymphangioleiomyomatosis. *Cancer Res* 2017; 77: 5491–5502.
12 Torgerson DG, Ampleford EJ, Chiu GY, *et al.* Meta-analysis of genome-wide association studies of asthma in ethnically diverse North American populations. *Nat Genet* 2011; 43: 887–892.
13 Corren J, Parnes J, Wang L, *et al.* Tezepelumab in adults with uncontrolled asthma. *N Engl J Med* 2017; 377: 936–946.