ORIGINAL ARTICLE
RARE LUNG DISEASES

The vitamin D binding protein axis modifies disease severity in lymphangioleiomyomatosis

Suzanne Miller[1], Clare Coveney[2], Janice Johnson[1], Aliki-Eleni Farmaki[3], Nishant Gupta[4], Martin D. Tobin[3,5], Louise V. Wain[3,5], Francis X. McCormack[4], David J. Boocock[2] and Simon R. Johnson[1,6,7]

Affiliations: [1]Division of Respiratory Medicine, University of Nottingham and National Institute for Health Research, Respiratory Biomedical Research Centre, Nottingham, UK. [2]John van Geest Cancer Research Centre, Nottingham Trent University, Nottingham, UK. [3]Dept of Health Sciences, College of Life Sciences, University of Leicester, Leicester, UK. [4]Division of Pulmonary, Critical Care and Sleep Medicine, University of Cincinnati, Cincinnati, OH, USA. [5]National Institute for Health Research Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, UK. [6]National Centre for LAM, Queen's Medical Centre, Nottingham, UK. [7]Nottingham Molecular Pathology Node, Nottingham, UK.

Correspondence: Simon R. Johnson, Division of Respiratory Medicine, University of Nottingham, Nottingham, NG7 2UH, UK. E-mail: simon.johnson@nottingham.ac.uk

@ERSpublications
The vitamin D binding protein and *GC* genotype are associated with lung function and survival in women with LAM http://ow.ly/UacI30leLzr

Cite this article as: Miller S, Coveney C, Johnson J, *et al*. The vitamin D binding protein axis modifies disease severity in lymphangioleiomyomatosis. *Eur Respir J* 2018; 52: 1800951 [https://doi.org/10.1183/13993003.00951-2018].

ABSTRACT Lymphangioleiomyomatosis (LAM) is a rare disease of women. Decline in lung function is variable, making appropriate targeting of therapy difficult. We used unbiased serum proteomics to identify markers associated with outcome in LAM.

101 women with LAM and 22 healthy controls were recruited from the National Centre for LAM in the UK. 152 DNA and serum samples with linked lung function and outcome data were obtained from patients in the National Heart, Lung and Blood Institute LAM Registry in the USA. Proteomic analysis was performed on a discovery cohort of 50 LAM and 20 control serum samples using a SCIEX SWATH mass spectrometric workflow. Protein levels were quantitated by ELISA and single nucleotide polymorphisms in *GC* (group-specific component) encoding vitamin D binding protein (VTDB) were genotyped.

Proteomic analysis showed VTDB was 2.6-fold lower in LAM than controls. Serum VTDB was lower in progressive compared with stable LAM (p=0.001) and correlated with diffusing capacity of the lung for carbon monoxide (p=0.01). Median time to death or lung transplant was reduced by 46 months in those with *CC* genotypes at rs4588 and 38 months in those with non-A-containing haplotypes at rs7041/4588 (p=0.014 and 0.008, respectively).

The VTDB axis is associated with disease severity and outcome, and *GC* genotype could help predict transplant-free survival in LAM.

This article has supplementary material available from erj.ersjournals.com

Received: May 21 2018 | Accepted after revision: July 29 2018

Introduction

Lymphangioleiomyomatosis (LAM) is a rare multisystem disease characterised by lung cysts and lymphatic abnormalities. The disease is almost exclusively restricted to women, of whom it affects around nine per million, and can occur both sporadically and in those with tuberous sclerosis complex (TSC) [1, 2]. In LAM, cysts progressively replace the lung parenchyma leading to recurrent pneumothorax and often respiratory failure over a variable period of years [3]. Lymphatic obstruction leads to chyloptysis, chylous effusions and ascites. Around half of patients with sporadic LAM and most with TSC-LAM also have angiomyolipomas, a benign tumour, generally occurring in the kidneys [2]. The lungs and lymphatics of patients are infiltrated by LAM cells: a clonal, metastatic cell with a combined smooth muscle and melanocyte phenotype characteristic of perivascular epithelioid cell neoplasms [4]. LAM cells have biallelic TSC mutations [5] which lead to hyperactivation of the mechanistic target of rapamycin (mTOR), a component of two multiprotein complexes, controlling proliferation, migration, autophagy and metabolism [6].

Most women with LAM lose lung function at an accelerated rate with forced expiratory volume in 1 s (FEV_1) declining by 70–140 mL per year [7, 8]; however, some progress rapidly while others can remain stable for many years [3, 9]. Treatment with mTOR inhibitors prevents loss of lung function in most with progressive disease [8–10]. Recognising progressive disease in individuals with mild lung function impairment is important, although generally requires multiple measurements over a prolonged period [7]. Markers of disease activity are therefore required to predict those at risk of loss of lung function to allow treatment before this occurs. Furthermore, stratification of patients with active disease could reduce the size, duration, cost and feasibility of phase II studies of new therapies.

A number of clinical and serum prognostic factors have been identified. Elevated serum vascular endothelial growth factor (VEGF)-D is associated with both the presence of LAM [11] and more rapid loss of lung function. Presentation with dyspnoea rather than pneumothorax and a response to bronchodilators have been associated with worse outcomes [12–14], whereas post-menopausal status is associated with slower lung function loss [7, 15]. Despite this, it is not possible to accurately predict prognosis within individuals. Here, we used serum proteomics to identify proteins associated with the presence and severity of LAM, and identified that changes in vitamin D binding protein (VTDB) and its gene, *GC* (group-specific component), are associated with disease severity and survival in LAM.

Materials and methods

Patients and sample collection

101 women with LAM and 22 healthy control women were recruited between 2011 and 2016 from the National Centre for LAM (Nottingham, UK) (figure 1). Ethical approval was obtained from the East Midlands Research Ethics Committee (13/EM/0264). All subjects provided written informed consent. A second cohort of 152 women with LAM recruited between 1998 and 2001 in the National Heart, Lung and Blood Institute (NHLBI) LAM Registry (USA) was used for replication and to study long-term survival [16] (figure 1). Baseline chest and abdominal computed tomography, serial lung function, serum, and DNA at recruitment were obtained for all subjects. Clinical assessment, lung function and sample analysis for both cohorts are described in the supplementary material. Due to duration of follow-up, all-cause mortality or the need for lung transplant was only available for the NHLBI LAM Registry cohort and was obtained by querying the US National Death Index (www.cdc.gov/nchs/ndi/index.htm) and the United Network for Organ Sharing databases (https://unos.org/data). As data on the use of rapamycin was not available for this cohort, outcome data were censored at 2010 before rapamycin was widely used for the treatment of LAM in the USA.

Proteomics

70 serum samples (50 LAM and 20 controls) were analysed on a SCIEX (Warrington, UK) TripleTOF 6600 mass spectrometer hyphenated to an Eksigent nanoLC 425 system using the SCIEX SWATH mass spectrometric workflow [17]. Tandem mass spectrometry (MS/MS) spectra were searched using ProteinPilot version 5.0 (SCIEX) with the Swiss-Prot human database (www.uniprot.org; January 2015) at 1% false discovery rate with an identification focus on biological modifications. SWATH data were aligned to library files in PeakView (SCIEX), uploaded and processed using the SCIEX OneOmics platform [18]. Full details are given in the supplementary material.

Serum protein quantification

Serum VTDB, α_1-acid glycoprotein 1 (A1AG1) and VEGF-D were determined in the UK cohort using Quantikine ELISA kits DVDBP0, DAGP00 and DVED00, respectively (R&D Systems, Abingdon, UK). VTDB in the NHLBI LAM Registry was measured using Quantikine ELISA kit DVDBP0B (R&D Systems).

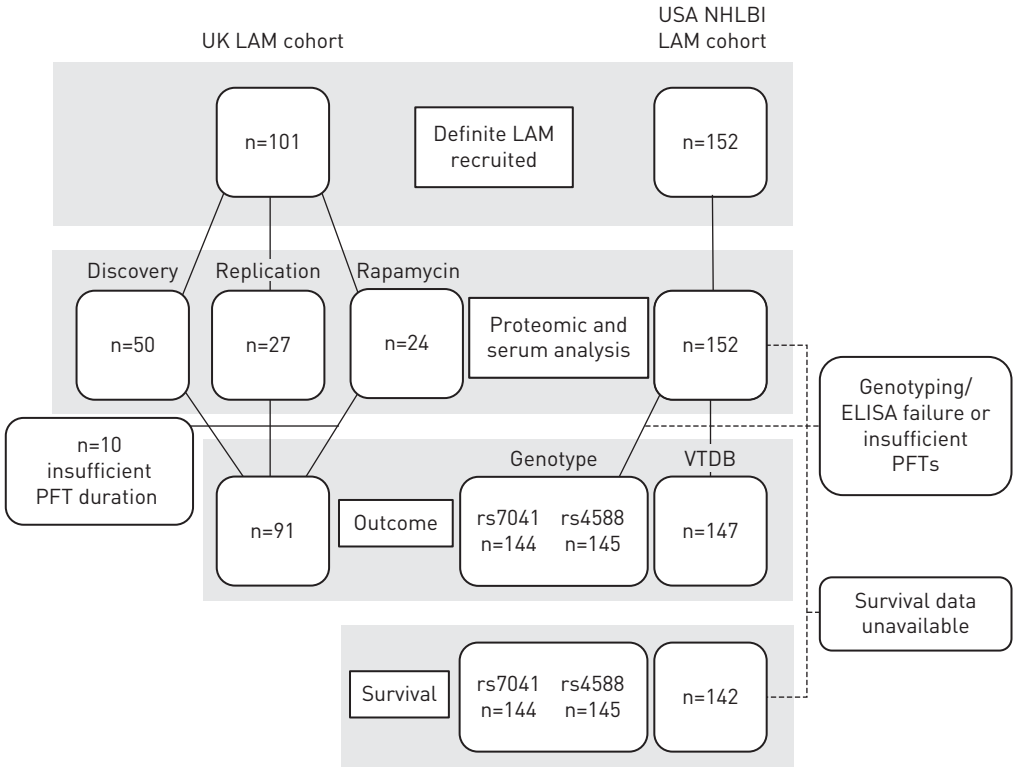

FIGURE 1 Enrolment and samples tested: recruitment and access to samples and lung function data in the UK and the USA National Heart, Lung and Blood Institute (NHLBI) lymphangioleiomyomatosis (LAM) cohorts. PFT: pulmonary function test; VTDB: vitamin D binding protein. The UK discovery cohort comprised 50 serum samples from individuals with LAM, the UK replication cohort comprised 27 LAM serum samples and the USA NHLBI LAM Registry cohort comprised 152 serum samples.

Genotyping

DNA was extracted from whole blood using the Wizard Genomic DNA Purification Kit (Promega, Southampton, UK). As *GC* genotype varies across populations, genetic analysis was confined to those of European ancestry. 65 UK LAM samples and 168 141 unrelated control women of European ancestry from the UK Biobank (www.ukbiobank.ac.uk) were genotyped using the Axiom UK Biobank array (Affymetrix,

TABLE 1 Clinical data for cohorts studied

	UK discovery cohort			UK replication cohort	UK rapamycin-treated group	USA NHLBI cohort	Healthy controls
	All	Stable	Progressive				
Subjects n	50	26	24	27	24	152	22
Age years	50.6±10.9	50.9±11.8	50.3±10.0	49.4±13.9	46.4±9.7	45.4±9.0	35.0±11.7
Disease duration years	13.9±11.1	14.2±11.4	13.5±11.1	9.1±9.5	13.1±9.5	4.6±4.3	NA
Angiomyolipoma	72	77	67	55	54	NT	NA
Lymphatic disease	16	15	17	23	25	NT	NA
TSC	14	19	8	15	21	NT	NA
Pneumothorax	48	50	46	40	46	NT	NA
Post-menopause	34	42	25	30	25	48	NA
FEV_1 % pred	68.9±20.6	76.4±18.9	60.8±19.5	77.4±23.4	46.7±14.8	74.1±27.5	NA
D_{LCO} % pred	59.8±15.8	68.9±12.7	50.0±12.9	62.9±17.1	43.3±12.3	55.7±25.6	NA
VEGF-D pg·mL^{-1}	1327±1187	985±833	1698±1405	1275±1527	1082±1257	NT	397±125

Data are presented as mean±SD (at recruitment) or % (present at any time in disease course), unless otherwise stated. NHLBI: National Heart, Lung and Blood Institute; TSC: tuberous sclerosis complex; FEV_1: forced expiratory volume in 1 s; D_{LCO}: diffusing capacity of the lung for carbon monoxide; VEGF: vascular endothelial growth factor; NA: not applicable; NT: not available for testing. Disease duration in the UK lymphangioleiomyomatosis cohort was from first symptom to enrolment, while in the NHLBI cohort disease duration was from diagnosis to enrolment. In the NHLBI cohort menopause was assumed if ⩾50 years of age.

High Wycombe, UK). Ancestry was determined from *k*-means clustering of the first two principal components from the genome-wide single nucleotide polymorphism (SNP) data [19]. Women from the NHLBI LAM Registry cohort were genotyped using KASP PCR genotyping (LGC Genomics. Hoddesdon, UK) with ancestry obtained by questionnaire.

Statistical analysis

Proteins identified by proteomics were considered differentially expressed if they were $\leqslant -2$ or $\geqslant 2$ $\log_2$ fold different between groups with a confidence $\geqslant 0.7$ as described by LAMBERT *et al.* [18]. Welch's t-test or the Mann–Whitney U-test were used for categorical data; linear regression and Spearman's correlation were used for continuous data. *GC* allele frequencies for women with LAM and UK Biobank controls were compared using Chi-squared tests [20]. Survival analyses were performed using Kaplan–Meier plots with differences analysed by the Mantel–Cox log-rank test. Analyses were performed using Prism version 7 (GraphPad, La Jolla, CA, USA) and SPSS version 24 (IBM, Armonk, NY, USA).

Results

Discovery cohort and serum proteomics

The first 50 UK women with LAM enrolled who were not treated with an mTOR inhibitor and 20 healthy control women formed the discovery cohort. The cohort was divided into more progressive and stable disease based upon a retrospective loss of FEV1 of >50 mL per year over a mean±SD period of observation of 11±4 years. Those with progressive disease had lower FEV1, diffusing capacity of the lung for carbon monoxide (*D*LCO) and higher serum VEGF-D values, but were of similar age and disease duration as those with stable disease (table 1).

MS of the 70 serum samples identified 126 proteins, including the serum proteins albumin, haemopexin, acid glycoprotein, immunoglobulins, complement components, clotting factors, proteases and protease inhibitors (supplementary table E1). VTDB levels were 2.6-fold lower (confidence 0.65) in LAM than healthy control women (supplementary table E2). To identify markers of severity we compared the proteomic profiles of those with stable and progressive disease. A1AG1 levels were 3.6-fold higher (confidence 0.70) in those with progressive compared with stable disease. Comparison of pre- and post-menopausal women with LAM did not identify differentially expressed proteins at the pre-specified confidence level.

Serum protein quantification

MS findings were validated using ELISAs for VTDB and A1AG1. Consistent with the proteomic findings, serum VTDB was lower in 50 women with LAM in the UK discovery cohort and 27 women with LAM in the UK replication cohort than in controls (p=0.007 and p=0.002, respectively). For the 77 women in the UK discovery (n=50) and replication cohorts (n=27) combined, VTDB was 273±96 $\mu g \cdot mL^{-1}$ in LAM and 347±92 $\mu g \cdot mL^{-1}$ in control women (p=0.002) (figure 2a). When assessed by ELISA, A1AG1 was higher in women with LAM in the discovery and replication cohorts than control women (p=0.04 and p=0.0001, respectively). A1AG1 was 910±478 $\mu g \cdot mL^{-1}$ for all women with LAM and 619±270 $\mu g \cdot mL^{-1}$ in control women (p=0.004) (figure 2b).

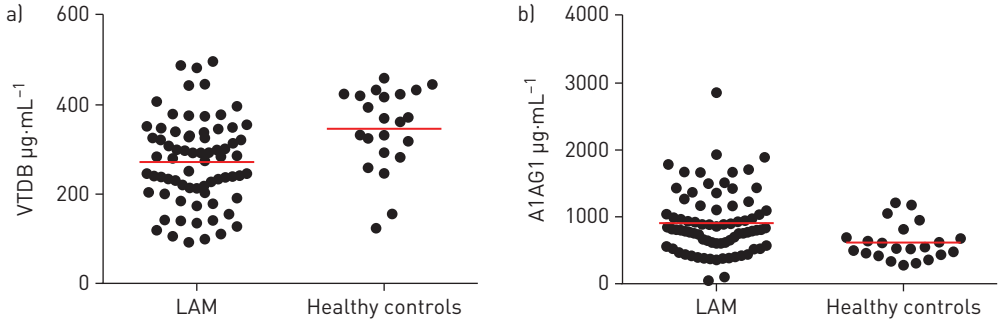

FIGURE 2 Serum vitamin D binding protein (VTDB) and α_1-acid glycoprotein 1 (A1AG1) in lymphangioleiomyomatosis (LAM) and healthy controls. a) Women with LAM had lower levels of serum VTDB compared with healthy control women (p=0.002). b) Women with LAM had higher levels of serum A1AG1 compared with healthy control women (p=0.004).

VTDB is associated with disease severity

VTDB was significantly lower in those with more progressive compared with more stable lung disease at recruitment (progressive 221±89 μg·mL^{-1} *versus* stable 299±90 μg·mL^{-1}; p=0.001) (figure 3a). VTDB level was positively associated with percent predicted *D*LCO (p=0.01) but not forced vital capacity (p=0.09) or FEV1 (p=0.23) (figure 3b–d). A1AG1 was higher in those with stable compared with progressive disease (stable 1004±525 μg·mL^{-1} *versus* progressive 753±341 μg·mL^{-1}; p=0.01) (supplementary figure E1), but was not related to lung function. Levels of VTDB were not associated with age, age at diagnosis, menopausal status, nature of presenting symptom, presence of tuberous sclerosis, angiomyolipomas, lymphatic disease or serum VEGF-D level (data not shown). The distribution of VTDB was similar in the 77 untreated women and 24 women receiving treatment with rapamycin for LAM, whereas A1AG1 was higher in the rapamycin-treated group (treated 1132±474 μg·mL^{-1} *versus* untreated 910±478 μg·mL^{-1}; p=0.031) (supplementary figure E2).

Association of GC genotypes with LAM and serum VTDB

As *GC* genotype varies according to ancestry, genetic analyses were restricted to the 65 individuals in the UK and 145 individuals in the NHLBI LAM Registry cohorts of European origin. Two SNPs within *GC* at rs7041 and rs4588 define the major *GC* haplotypes: 1) *GC1F* where rs7041 (G) and rs4588 (A), 2) *GC1S* where rs7041 (T) and rs4588 (A), and 3) *GC2* where rs7041 (G) and rs4588 (C). The allele frequencies at these SNPs in the UK and NHLBI LAM Registry cohorts did not differ from control women in the UK Biobank or each other (supplementary table E3). In both LAM cohorts, as in the general population, serum VTDB was dependent on *GC* genotype (supplementary figure E3) [21].

Association of VTDB protein and genotype with outcome

From the UK cohort, 91 women with LAM had lung function measured over >1 year after enrolment (64 untreated and 27 receiving rapamycin for LAM). The mean period of observation was 19 months, corresponding to 144 patient-years of observation. Within the NHLBI LAM Registry cohort, 136 women with untreated LAM had lung function measured over >1 year after enrolment with a mean period of observation of 40 months, corresponding to 500 patient-years of observation. Serum VTDB was not associated with prospective change in lung function in either cohort (table 2).

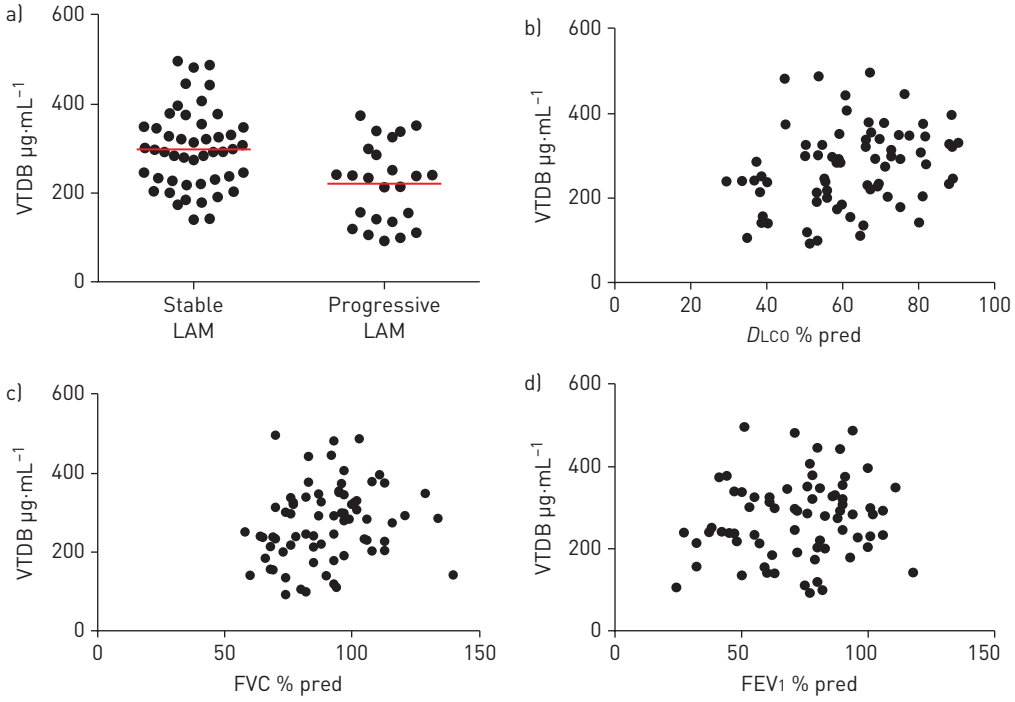

FIGURE 3 Vitamin D binding protein (VTDB) is associated with disease severity. LAM: lymphangioleiomyomatosis; *D*LCO: diffusing capacity of the lung for carbon monoxide; FVC: forced vital capacity; FEV1: forced expiratory volume in 1 s. a) Lower levels of serum VTDB are associated with progressive LAM compared with stable LAM (p=0.001). b) VTDB level is positively correlated with *D*LCO % pred (p=0.01). c) VTDB is not associated with FVC % pred (p=0.09). d) VTDB is not associated with FEV1 % pred (p=0.23).

TABLE 2 Prospective change in forced expiratory volume in 1 s (FEV_1) and diffusing capacity of the lung for carbon monoxide (D_{LCO}) and relationship to vitamin D binding protein (VTDB)

	UK cohort				NHLBI cohort	
	Untreated	p-value#	Rapamycin	p-value#	Untreated	p-value#
Subjects n	64		27		136	
ΔFEV_1 mL per year	−32.6±111.2	NS	24.3±141.4	NS	−94.7±96.2	NS
ΔD_{LCO} mmol·min^{-1}·kPa^{-1} per year	−0.2±0.40	NS	−0.17±0.23	NS	−0.23±0.31	NS
VTDB µg·mL^{-1}	273±96		281±105		255±53.4	

Data are presented as mean±SD, unless otherwise stated. NHLBI: National Heart, Lung and Blood Institute. #: p-value for Spearman's correlation with serum VTDB. NS: nonsignificant.

Within the NHLBI LAM Registry cohort, those with low serum VTDB, the AA genotype at rs4588 and TT at rs7041 had the highest rates of loss of FEV_1 and D_{LCO}, although not significantly so (table 3). We then examined the relationship of the VTDB axis with time to death or lung transplant in the NHLBI LAM Registry cohort. Although time to death or transplant was not associated with serum VTDB level (p=0.76) (figure 4a) or rs7041 genotype, there was an association with rs4588 genotype. Median time to death or transplant for the AA or AC genotype at rs4588 was 150 months compared with 104 months for CC (p=0.014) (figure 4b). Median time to death or transplant for all haplotypes with an A allele at rs4588 (including *GC1F* and *GC1S* haplotypes) was 150 months compared with 112 months for haplotypes with no A allele present (including *GC2*; p=0.008) (figure 4c).

Discussion

We have shown for the first time that the VTDB axis is associated with both severity and outcome in women with LAM. VTDB levels were associated with D_{LCO} and disease activity at assessment. Those with progressive disease, defined by a loss of FEV_1 of >50 mL per year, tended to have lower levels of VTDB than those with more stable disease with a loss of FEV_1 of <50 mL per year, despite being matched for age and other clinical manifestations. Haplotypes of *GC* were associated with the time to death or lung transplant. As such, *GC* genotype is the first genetic host factor found to influence transplant-free survival in LAM.

VTDB is a glycosylated α-globulin produced by the liver, kidneys, adipose tissue and neutrophils. Coded for by the *GC* gene on chromosome 4q, two SNPs in exon 11, *i.e.* rs7041 (Glu416Asp) and rs4588 (Thr420Lys), define the three major haplotypes of VTDB: *GC1F* (416Asp/420Thr), *GC1S* (416Glu/420Lys) and *GC2* (416Asp/420Lys), with serum VTDB level related to these SNPs [21]. VTDB binds 25(OH)-vitamin D and $1,25(OH)_2$-vitamin D, although vitamin D levels are far exceeded by the transport capacity of VTDB. Serum levels of VTDB and vitamin D are unrelated in many diseases studied, including chronic obstructive pulmonary disease (COPD) [22]. The *GC* variants have differing affinities for vitamin D; the complexities of the VTDB isoforms, vitamin D and their impact on lung disease are not yet clear [23].

TABLE 3 Relationship of vitamin D binding protein (VTDB) genotype with clinical features, serum VTDB and change in lung function in the National Heart, Lung and Blood Institute Lymphangioleiomyomatosis (LAM) Registry cohort

	rs4588 SNP				rs7041 SNP			
	AA	CA	CC	p-value	TT	GT	GG	p-value
Subjects n	11	46	74		25	57	48	
Age at diagnosis years	37.4±6.7	42.1±9.9	40.6±9.2	NS	39.9±7.6	41.8±9.7	40.5±9.6	NS
Age at recruitment years	40.9±6.4	47.2±9.4	45.3±8.9		43.8±7.5	46.4±9.4	45.4±9.3	
FEV_1 % pred	88.0±21.0	78.8±25.2	72.9±29.4	NS	79.9±26.8	79.0±30.2	72.0±24.0	NS
D_{LCO} % pred	58.3±17.5	59.5±22.5	57.0±29.3	NS	56.3±18.2	58.1±30.8	58.9±23.2	NS
VTDB µg·mL^{-1}	220±36	245±57	266±52	0.022	233±43	250±57	270±53	0.026
ΔFEV_1 mL per year	−125±142	−78±81	−99±97	NS	−135±126	−80±84	−94±94	NS
ΔD_{LCO} mmol·min^{-1}·kPa^{-1} per year	−0.35±0.23	−0.21±0.36	−0.22±0.3	NS	−0.26±0.27	−0.20±0.35	−0.26±0.27	NS

Data are presented as mean±SD for women with LAM of European ancestry, unless otherwise stated (data for forced expiratory volume in 1 s (FEV_1) % pred, diffusing capacity of the lung for carbon monoxide (D_{LCO}) % pred, VTDB and age at recruitment were all at entry to the study; ΔFEV_1 and ΔD_{LCO} are prospective changes from recruitment). Linear regression was used to model the relationship between genotype, clinical factors and VTDB. NS: nonsignificant.

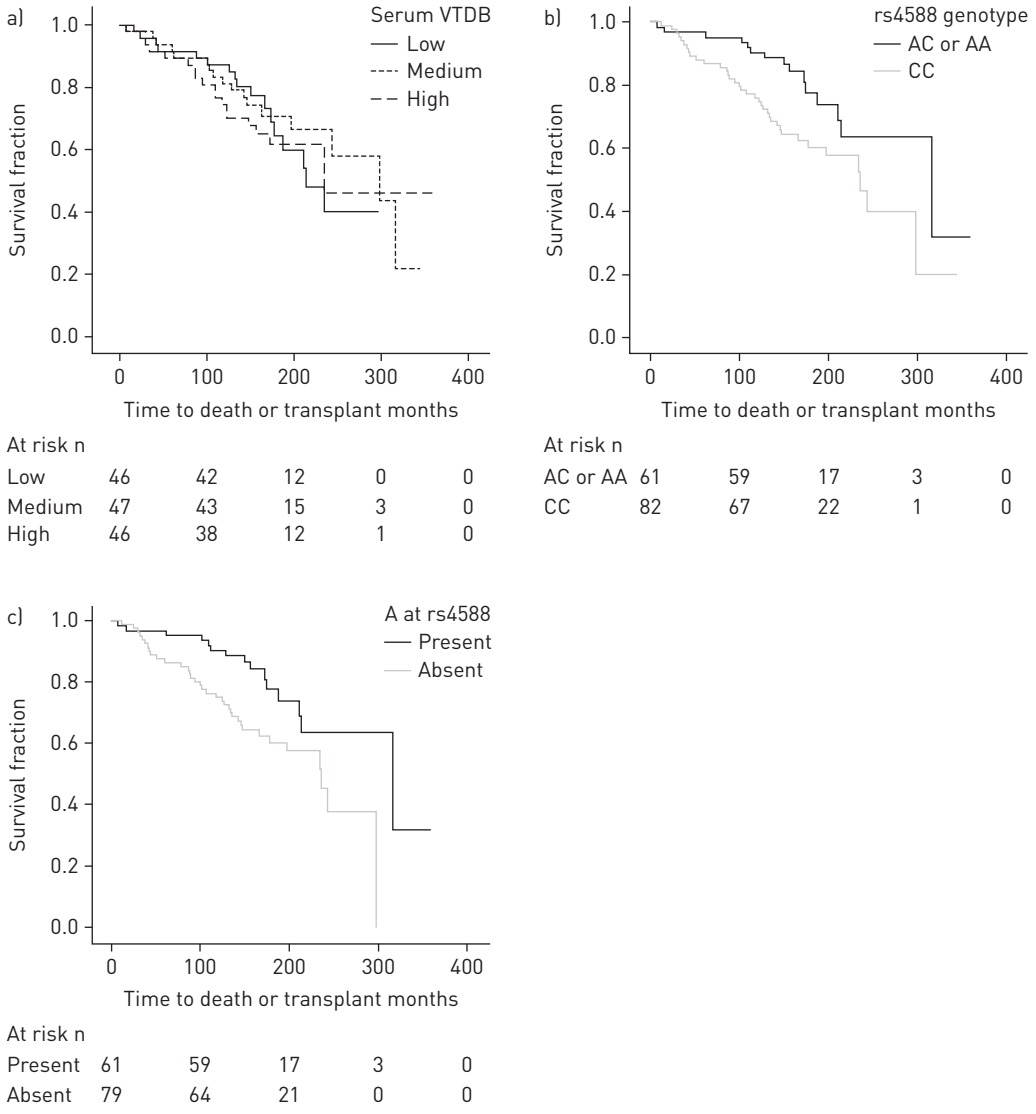

FIGURE 4 Survival analysis for vitamin D binding protein (VTDB) level and *GC* (group-specific component) genotype in the National Heart, Lung and Blood Institute Lymphangioleiomyomatosis Registry cohort. a) Overall time to death or transplant did not differ with serum VTDB level (low (147–221 $\mu g \cdot mL^{-1}$), medium (222–275 $\mu g \cdot mL^{-1}$) and high (276–413 $\mu g \cdot mL^{-1}$)) (p=0.76). b) Individuals with the AA or AC genotype at rs4588 had greater time to death or transplant than those with the CC genotype (p=0.014). c) Haplotypes with an A allele at rs4588 (*GC1F* and *GC1S*) were associated with longer time to death or transplant (p=0.008).

The mechanism relating *GC* genotype and serum VTDB is also unknown. rs7041 and rs4588 are intronic SNPs; neither are in linkage disequilibrium with known promotor or enhancer SNPs, nor are they known to affect protein stability. Factors other than *GC* genotype, including epigenetics, may also influence serum VTDB levels, as although serum VTDB is lower in women with LAM than controls, *GC* genotype in our study was not different.

Our findings reflect the complexity of both the VTDB axis and LAM. We observed that lower serum VTDB was associated with lower lung function and more active lung disease at presentation. As VTDB is not associated with other aspects of the LAM phenotype, including the presence of angiomyolipoma or lymphatic disease, it is likely that the VTDB axis is not related to LAM *per se*, but as in other lung diseases may alter the tissue response to disease. Importantly, *GC* genotype was associated with time to death or lung transplantation. The strongest effect was for the *GC1F* and *GC1S* haplotypes, which were associated with an increase in median survival of >3 years. Interestingly, these and other *GC* variants associated with improved survival were not those associated with the lowest serum VTDB levels. VTDB is a multifunctional protein which may impact upon the response to lung damage in a number of ways. *GC1F* and *GC1S* are associated with increased macrophage activation over *GC2* [22], and increased macrophage activation may be protective

in LAM, either by enhancing protective neutrophil responses or enhancing the chemotactic effect of complement-derived C5a [24, 25]. VTDB also acts as an actin scavenging protein and therefore has the potential to influence disease by different mechanisms, including altered innate immunity and tissue repair. Different *GC* haplotypes are already associated with susceptibility to lung disease, with *GC1F* being associated with an enhanced risk of COPD over *GC1S* and *GC2* [26].

These observations underscore the multiple potential functions of VTDB, and how these functions may be related to genotype and the complex relationship with lung disease. The complexity of LAM, a multisystem disease, is also likely to be important. For example, VTDB protein is associated with D_{LCO} but not FEV_1, forced vital capacity or event-free survival. While FEV_1 is generally used to study the natural history of LAM, D_{LCO} is usually impaired before FEV_1 and may better reflect early parenchymal damage in LAM, with loss of FEV_1 occurring later due to loss of elastic recoil and premature airway closure brought about by parenchymal damage. Pulmonary vascular disease, host defence, peripheral muscle function and other processes potentially affected by VTDB function may also contribute to survival.

One of the strengths of our study was the use of an unbiased proteomic method that identified VTDB as a protein of interest in LAM. The involvement of the vitamin D axis in other diseases associated with tissue remodelling make our findings biologically plausible [27]. However, our study also has limitations, including the low number of control samples, technical limitations and those inherent in studying rare diseases. First, VTDB was one of only two proteins differentially expressed in the serum of women with LAM and the proteomic methodology used did not identify other LAM markers such as VEGF-D. VEGF-D is expressed at picomolar levels [28], whereas VTDB is present a micromolar levels, suggesting that only relatively abundant serum proteins with robust differences between women with LAM and healthy controls could be detected using this proteomic strategy. It is therefore likely that other potentially useful biomarkers remain undiscovered. Consistent with this, A1AG1, also known as orosomucoid, the other protein linked to the presence of LAM in our proteomic screen, is another relatively abundant plasma α-globulin, comprising 1–3% of plasma proteins. As A1AG1 is an acute-phase protein, already recognised as a biomarker of overall survival in many populations, we did not study it further [29]. As LAM is very rare, studying the disease relies upon cohorts accumulated over longer periods of time. Although both cohorts studied used protocol-driven assessments to capture key data including lung function, there are some differences in the data available for these groups. Although the two cohorts used were similar in terms of age and lung function, prospective change in lung function differed, probably due to the use of rapamycin in the UK cohort resulting in reduced loss of FEV_1. Conversely, due to time of recruitment, long-term survival prior to rapamycin use can now only be studied in the NHLBI LAM Registry cohort. Current individuals with progressive disease, including those in the UK cohort studied here, tend to be treated with rapamycin [10] and longer periods of observation are needed to study the effect of the VTDB protein or genotype on survival in women with LAM treated with rapamycin.

In conclusion, low levels of VTDB are associated with poor lung function in LAM and *GC* genotypes are associated with long-term outcome. Our findings suggest that the VTDB axis is a host factor that may protect against lung damage in LAM and could be of prognostic significance. Further studies are required to validate our findings and understand how the VTDB isoforms modulate lung damage in LAM and other diseases.

Acknowledgements: We are grateful to the investigators who contributed to the National Heart, Lung and Blood Institute LAM Registry, and to the women with LAM who contributed to the research. This research used the ALICE High Performance Computing Facility at the University of Leicester (Leicester, UK). L.V. Wain holds a GSK/British Lung Foundation Chair in Respiratory Research. M.D. Tobin holds a Wellcome Trust Investigator Award (WT 202849/Z/16/Z). This research has been performed using the UK Biobank Resource under application 648.

Author contributions: S.R. Johnson conceived and designed the study. S. Miller, S.R. Johnson, J. Johnson, N. Gupta and F.X. McCormack collected clinical information and samples. S. Miller, S.R. Johnson, C. Coveney, A-E. Farmaki, M.D. Tobin, L.V. Wain and D.J. Boocock analysed and interpreted the data. S. Miller and S.R. Johnson wrote the manuscript. All authors critically reviewed and approved the final manuscript.

Conflict of interest: S. Miller has nothing to disclose. C. Coveney has nothing to disclose. J. Johnson has nothing to disclose. A-E. Farmaki has nothing to disclose. N. Gupta has nothing to disclose. M.D. Tobin reports grants from GSK and Pfizer, outside the submitted work. L.V. Wain reports grants from GSK and Pfizer, outside the submitted work. F.X. McCormack reports nonfinancial support (consultancy) for LAM Therapeutics, nonfinancial support (data and safety monitoring board) for Takeda and Promedior, nonfinancial support (speaking) for Sanofi-Aventis US, and personal fees for speaking, consultancy and received travel support from Boehringer Ingelheim International, F. Hoffmann-La Roche and Novartis, and honoraria for an adjudication committee from Gilead Sciences, outside the submitted work; in addition, F.X. McCormack has a patent "Use of VEGF-D in the diagnosis of lymphangioleiomyomatosis" issued (with royalties paid to the University of Cincinnati), and F.X. McCormack's spouse owns stocks, stock options and other ownership interests in Sanofi-Aventis, US. D.J. Boocock has nothing to disclose. S. R. Johnson reports grants from NIHR and LAM Foundation, during the conduct of the study.

Support statement: This study was funded by the National Institute for Health Research Rare Disease Translational Research Collaboration, a LAM Foundation Biomarker Innovation Grant and a LAM Foundation project grant. The funders had no role in the study design, analysis, interpretation, manuscript writing or decision to submit. The authors have not been paid to write this article. S. Miller and S.R. Johnson were involved in all stages of study development and delivery, had full access to all data in the study, and had final responsibility for the decision to submit the publication. This article presents independent research funded partially by the UK National Institute for Health Research (NIHR). The views expressed are the authors' own and not necessarily those of the NHS or the NIHR. Funding information for this article has been deposited with the Crossref Funder Registry.

References

1 Harknett EC, Chang WYC, Byrnes S, *et al.* Regional and national variability suggests underestimation of prevalence of lymphangioleiomyomatosis. *Q J Med* 2011; 104: 971–979.
2 Johnson SR. Lymphangioleiomyomatosis. *Eur Respir J* 2006; 27: 1056–1065.
3 Johnson SR, Whale CI, Hubbard RB, *et al.* Survival and disease progression in UK patients with lymphangioleiomyomatosis. *Thorax* 2004; 59: 800–803.
4 Pea M, Martignoni G, Bonetti F, *et al.* Tumors characterized by the presence of HMB45-positive perivascular epithelioid cell (PEC) – a novel entity in surgical pathology. *Electron J Pathol Histol* 1997; 3: 28–40.
5 Carsillo T, Astrinidis A, Henske EP. Mutations in the tuberous sclerosis complex gene *TSC2* are a cause of sporadic pulmonary lymphangioleiomyomatosis. *Proc Natl Acad Sci USA* 2000; 97: 6085–6090.
6 Saxton RA, Sabatini DM. mTOR signaling in growth, metabolism, and disease. *Cell* 2017; 169: 361–371.
7 Johnson SR, Tattersfield AE. Decline in lung function in lymphangioleiomyomatosis: relation to menopause and progesterone treatment. *Am J Respir Crit Care Med* 1999; 160: 628–633.
8 McCormack FX, Inoue Y, Moss J, *et al.* Efficacy and safety of sirolimus in lymphangioleiomyomatosis. *N Engl J Med* 2011; 364: 1595–1606.
9 Bee J, Bhatt R, McCafferty I, *et al.* Audit, research and guideline update: a 4-year prospective evaluation of protocols to improve clinical outcomes for patients with lymphangioleiomyomatosis in a national clinical centre. *Thorax* 2015; 70: 1202–1204.
10 Bee J, Fuller S, Miller S, *et al.* Lung function response and side effects to rapamycin for lymphangioleiomyomatosis: a prospective national cohort study. *Thorax* 2018; 73: 369–375.
11 Young LR, VanDyke R, Gulleman PM, *et al.* Serum vascular endothelial growth factor-D prospectively distinguishes lymphangioleiomyomatosis from other diseases. *Chest* 2010; 138: 674–681.
12 Oprescu N, McCormack FX, Byrnes S, *et al.* Clinical predictors of mortality and cause of death in lymphangioleiomyomatosis: a population-based registry. *Lung* 2013; 191: 35–42.
13 Cohen MM, Pollock-BarZiv S, Johnson SR. Emerging clinical picture of lymphangioleiomyomatosis. *Thorax* 2005; 60: 875–879.
14 Taveira-DaSilva AM, Steagall WK, Rabel A, *et al.* Reversible airflow obstruction in lymphangioleiomyomatosis. *Chest* 2009; 136: 1596–1603.
15 Taveira-DaSilva AM, Stylianou MP, Hedin CJ, *et al.* Decline in lung function in patients with lymphangioleiomyomatosis treated with or without progesterone. *Chest* 2004; 126: 1867–1874.
16 Ryu JH, Moss J, Beck GJ, *et al.* The NHLBI Lymphangioleiomyomatosis Registry: characteristics of 230 patients at enrollment. *Am J Respir Crit Care Med* 2006; 173: 105–111.
17 Sajic T, Liu Y, Aebersold R. Using data-independent, high-resolution mass spectrometry in protein biomarker research: perspectives and clinical applications. *Proteomics Clin Appl* 2015; 9: 307–321.
18 Lambert J-P, Ivosev G, Couzens AL, *et al.* Mapping differential interactomes by affinity purification coupled with data-independent mass spectrometry acquisition. *Nat Methods* 2013; 10: 1239–1245.
19 Shrine N, Guyatt AL, Erzurumluoglu AM, *et al.* New genetic signals for lung function highlight pathways and pleiotropy, and chronic obstructive pulmonary disease associations across multiple ancestries. *bioRxiv* 2018; preprint [https://doi.org/10.1101/343293].
20 Wain LV, Shrine N, Miller S, *et al.* Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med* 2015; 3: 769–781.
21 Moy KA, Mondul AM, Zhang H, *et al.* Genome-wide association study of circulating vitamin D-binding protein. *Am J Clin Nutr* 2014; 99: 1424–1431.
22 Wood AM, Bassford C, Webster D, *et al.* Vitamin D-binding protein contributes to COPD by activation of alveolar macrophages. *Thorax* 2011; 66: 205–210.
23 Arnaud J, Constans J. Affinity differences for vitamin D metabolites associated with the genetic isoforms of the human serum carrier protein (DBP). *Hum Genet* 1993; 92: 183–188.
24 Zhang J, Kew RR. Identification of a region in the vitamin D-binding protein that mediates its C5a chemotactic cofactor function. *J Biol Chem* 2004; 279: 53282–53287.
25 Yamamoto N, Homma S. Vitamin D3 binding protein (group-specific component) is a precursor for the macrophage-activating signal factor from lysophosphatidylcholine-treated lymphocytes. *Proc Natl Acad Sci USA* 1991; 88: 8539–8543.
26 Horita N, Miyazawa N, Tomaru K, *et al.* Vitamin D binding protein genotype variants and risk of chronic obstructive pulmonary disease: a meta-analysis. *Respirology* 2015; 20: 219–225.
27 Chishimba L, Thickett DR, Stockley RA, *et al.* The vitamin D axis in the lung: a key role for vitamin D-binding protein. *Thorax* 2010; 65: 456–462.
28 Young LR, Lee H-S, Inoue Y, *et al.* Serum VEGF-D concentration as a biomarker of lymphangioleiomyomatosis severity and treatment response: a prospective analysis of the Multicenter International Lymphangioleiomyomatosis Efficacy of Sirolimus (MILES) trial. *Lancet Respir Med* 2013; 1: 445–452.
29 Fischer K, Kettunen J, Würtz P, *et al.* Biomarker profiling by nuclear magnetic resonance spectroscopy for the prediction of all-cause mortality: an observational study of 17 345 persons. *PLoS Med* 2014; 11: e1001606.