The Vitamin D Binding Protein axis modifies disease severity in Lymphangioleiomyomatosis

Authors: 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].

Affiliations:

[1]Division of Respiratory Medicine, University of Nottingham and National Institute for Health Research, Respiratory Biomedical Research Centre, Nottingham, United Kingdom.

[2]John van Geest Cancer Research Centre, Nottingham Trent University, Nottingham, United Kingdom.

[3]Department of Health Sciences, College of Life Sciences, University of Leicester, Leicester, United Kingdom.

[4]Division of Pulmonary, Critical Care and Sleep Medicine, University of Cincinnati, Ohio, USA.

[5]National Institute for Health Research Leicester Respiratory Biomedical Research Centre, Glenfield Hospital, Leicester, United Kingdom.

[6]National Centre for LAM, Queen's Medical Centre, Nottingham, United Kingdom.

Data supplement

Supplementary 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, hence-forth termed the UK cohort. Ethical approval was obtained from the East Midlands Research Ethics Committee (13/EM/0264). For healthy controls, age and ethnicity were recorded. All subjects provided written informed consent. Clinical history, presence of TSC, angiomyolipoma, lymphatic disease, menopausal status and drug treatment were recorded. Lymphatic disease was defined as the presence of chylous collections in the chest or abdomen, lymphangioleiomyomas or lymphadenopathy due to LAM visible on CT scanning of the chest abdomen and pelvis. Disease duration was calculated as the time from first symptom attributable to LAM as previously described[7]. Blood samples were taken at enrolment and processed within one hour of phlebotomy. Whole blood collected in serum separator tubes were allowed to clot for 1 hour at room temperature and separated by centrifugation at 1000 g for 10 minutes. Blood and serum samples were stored at -80 °C until analysis.

A second cohort of 243 subjects recruited between 1998 and 2001 in the National Heart Lung and Blood Institute (NHLBI) LAM Registry to study the natural history of LAM was used as a replication cohort and to study long term survival. This cohort has been described in detail previously[18]. Serum and DNA samples at enrolment, available from 152 of these 243 subjects, along with clinical and prospective lung function data were obtained from the National Disease Research Interchange who now curate the resource. Outcome data, either all-cause mortality or the need for lung transplant in the period following baseline assessment, were obtained by querying the United States National Death Index and the United Network for Organ Sharing databases. 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. Suitable samples for serum protein measurement, genotyping or lung function over greater than one year's duration was not available in all subjects (Figure 1). The numbers included in individual analyses are stated in the results.

Lung function was measured at either Nottingham University Hospitals NHS Trust or the referring centre in the USA according to ERS/ATS standards. Prospective change in lung function was calculated as the difference between FEV_1 and DL_{CO} measured at recruitment and last follow up visit expressed in ml/year for FEV_1 (ΔFEV_1) and as ml/min/kPa/yr for DL_{CO} (ΔDL_{CO}). To reduce variation in measurement of disease progression, only values spanning one year or longer were used for this analysis[7]. Classification into stable or progressive disease at presentation was performed by calculating retrospective loss of FEV_1 until the time of enrolment between the first recorded FEV_1 value and the FEV_1 at study enrolment divided by the time interval and expressed in ml/year. Those with a retrospective ΔFEV_1 of less than -50 ml/yr were arbitrarily classified as more stable and greater than -50 ml/yr as more progressive.

Fifty women with LAM from the UK cohort who had not been treated with rapamycin formed the initial proteomic discovery population. A further 27 untreated patients from the UK were used as a replication cohort and 24 who were receiving rapamycin for treatment of LAM at recruitment were also studied (Table 1). The discovery cohort was subdivided into those with stable and more progressive disease based upon retrospective loss of lung function.

Proteomics

Serum was diluted and filtered at 2 µm to remove any particulates to a final 1 in 45 dilution in 100 mM Triethylammonium bicarbonate buffer. An alkylation and reduction step adding 2µL 0.5 mM Dithiothreitol (DTT) with 45 min shaking at 56 °C followed by 7.15 µL of 140 mM Iodoacetamide and 30 min incubation in darkness at room temperature. The reaction was then quenched using 1.95 µL of 0.5 mM DTT. Samples were digested with 1.95 µL of 1 µg/µL trypsin (T656720UG, Sigma, UK) over 17 hrs at 37°C while shaking/agitating after which samples were lyophilised in a speed vac and re-suspended at decreasing concentrations of acetonitrile (ACN) to a final mixture of 40 µL and 5% ACN 0.1% formic acid (FA). After high speed centrifugation, supernatants were transferred to appropriate tubes for mass spectrometric analysis. The Biognosis HRM retention time standard was added for downstream alignment. Samples were analysed on a SCIEX TripleTOF 6600 mass spectrometer hyphenated to an Eksigent nanoLC 425 system operating in micro flow (5 µL/min). The SCIEX SWATH mass spectrometric workflow[17] was utilised for relative protein quantitation, wherein data acquired from a quantitative data independent (DIA) SWATH, are assembled against libraries of protein identified using a data dependent acquisition. Chromatographic separation for protein identification (Information Dependent Acquisition/IDA) was over an 87 min gradient, 4 µL direct injection on a YMC 25 cm x 0.3mm Triart-C18 column (12nm, 3µm particle size) with a gradient of 3 % mobile phase B (2% acetonitrile, 5% DMSO in 0.1% FA) to 30 % over 38 min; to 40% B at 73 min, 80 % B at 75 min, held for 3 min then returned to 3 % over 1 min. Chromatographic separation for SWATH runs was conducted as above but on a 57 min gradient of 3 % mobile phase B (2% ACN, 5% DMSO in 0.1% FA) to 30 % over 38 min; to 40 % B at 43 min, 80 % B at 45 min held for 3 min then returned to 3 % over 1 min. The mass spectrometer set up and method settings consisted of a Duospray™ source (SCIEX) with a 50 µm electrode at +5500V (gas settings GS1 15; GS2 0; CUR 25; TEMP 0). IDA was carried out using parameters of Top 30 (TOFMS 250 ms accumulation time, production 60 ms, total cycle time 2.1 s); charge state 2 - 4 above a threshold of 200 cps; dynamic exclusion for 10 seconds using rolling collision energy (optimised for m/z of target ion). SWATH methods consisted of 100 variable windows optimised for serum and cell lysate proteins. MS/MS spectra were searched using ProteinPilot 5.0 (SCIEX) with the Swissprot human database (Jan 2015) at 1 % false discovery rate with an identification focus on biological modifications. SWATH data were aligned to the library files in PeakView (SCIEX) and uploaded to the SCIEX OneOmics platform for processing, compilation, assembly and annotation of SWATH data.

Supplementary results

Supplementary table E1. Serum proteins identified by proteomic screen in LAM and control serum.

Protein Name	UniProt ID	Full name
A1AG1	P02763	Alpha-1-acid glycoprotein 1
A1AT	P01009	Alpha-1-antitrypsin
A1BG	P04217	Alpha-1B-glycoprotein
A2GL	P02750	Leucine-rich alpha-2-glycoprotein
A2MG	P01023	Alpha-2-macroglobulin
ACTG	P63261	Actin, cytoplasmic 2
AFAM	P43652	Afamin
ALBU	P02768	Serum albumin
AMBP	P02760	Protein AMBP
ANGT	P01019	Angiotensinogen
ANT3	P01008	Antithrombin-III
APOA	P08519	Apolipoprotein(a)
APOA1	P02647	Apolipoprotein A-I
APOA2	P02652	Apolipoprotein A-II
APOA4	P06727	Apolipoprotein A-IV
APOB	P04114	Apolipoprotein B-100
APOC2	P02655	Apolipoprotein C-II
APOC3	P02656	Apolipoprotein C-III
APOD	P05090	Apolipoprotein D
APOE	P02649	Apolipoprotein E
APOF	Q13790	Apolipoprotein F
APOH	P02749	Beta-2-glycoprotein 1
APOL1	O14791	Apolipoprotein L1
APOM	O95445	Apolipoprotein M
C1QC	P02747	Complement C1q subcomponent subunit C
C1R	P00736	Complement C1r subcomponent
C1S	P09871	Complement C1s subcomponent
C4BPA	P04003	C4b-binding protein alpha chain
CAMP	P49913	Cathelicidin antimicrobial peptide
CBPN	P15169	Carboxypeptidase N catalytic chain
CD44	P16070	CD44 antigen
CD5L	O43866	CD5 antigen-like
CERU	P00450	Ceruloplasmin
CFAB	P00751	Complement factor B
CFAH	P08603	Complement factor H
CFAI	P05156	Complement factor I

CLUS	P10909	Clusterin
CO2	P06681	Complement C2
CO3	P01024	Complement C3
CO4B	P0C0L5	Complement C4-B
CO5	P01031	Complement C5
CO6	P13671	Complement component C6
CO8A	P07357	Complement component C8 alpha chain
CO8G	P07360	Complement component C8 gamma chain
CO9	P02748	Complement component C9
CXCL7	P02775	Platelet basic protein
FA12	P00748	Coagulation factor XII
FBLN1	P23142	Fibulin-1
FCG3A	P08637	Low affinity immunoglobulin gamma Fc region receptor III-A
FCN2	Q15485	Ficolin-2
FETUA	P02765	Alpha-2-HS-glycoprotein
FHR3	Q02985	Complement factor H-related protein 3
FHR5	Q9BXR6	Complement factor H-related protein 5
FIBA	P02671	Fibrinogen alpha chain
FINC	P02751	Fibronectin
FOXN3	O00409	Forkhead box protein N3
GELS	P06396	Gelsolin
H2AX	P16104	Histone H2AX
HBA	P69905	Hemoglobin subunit alpha
HBB	P68871	Hemoglobin subunit beta
HDAC1	Q13547	Histone deacetylase 1
HEMO	P02790	Hemopexin
HMMR	O75330	Hyaluronan mediated motility receptor
HPT	P00738	Haptoglobin
HPTR	P00739	Haptoglobin-related protein
HRG	P04196	Histidine-rich glycoprotein
HS12B	Q96MM6	Heat shock 70 kDa protein 12B
HS90B	P08238	Heat shock protein HSP 90-beta
HV101	P01742	Immunoglobulin heavy variable 1-69
HV304	P01765	Immunoglobulin heavy variable 3-23
HV305	P01766	Immunoglobulin heavy variable 3-13
HV306	P01767	Immunoglobulin heavy variable 3-53
HV311	P01772	Immunoglobulin heavy variable 3-33
IBP3	P17936	Insulin-like growth factor-binding protein 3
IC1	P05155	Plasma protease C1 inhibitor
IGHA1	P01876	Ig alpha-1 chain C region
IGHA2	P01877	Ig alpha-2 chain C region
IGHD	P01880	Ig delta chain C region
IGHG1	P01857	Ig gamma-1 chain C region
IGHG2	P01859	Ig gamma-2 chain C region
IGHG3	P01860	Ig gamma-3 chain C region
IGHM	P01871	Ig mu chain C region

IGJ	P01591	Immunoglobulin J chain
IGKC	P01834	Ig kappa chain C region
IGLL5	B9A064	Immunoglobulin lambda-like polypeptide 5
ITIH1	P19827	Inter-alpha-trypsin inhibitor heavy chain H1
ITIH2	P19823	Inter-alpha-trypsin inhibitor heavy chain H2
ITIH4	Q14624	Inter-alpha-trypsin inhibitor heavy chain H4
K1024	Q9UPX6	UPF0258 protein KIAA1024
K1C19	P08727	Keratin, type I cytoskeletal 19
KANK3	Q6NY19	KN motif and ankyrin repeat domain-containing protein 3
KCNC4	Q03721	Potassium voltage-gated channel subfamily C member 4
KI67	P46013	Proliferation marker protein Ki-67
KIFC2	Q96AC6	Kinesin-like protein KIFC2
KLKB1	P03952	Plasma kallikrein
KNG1	P01042	Kininogen-1
KV102	P01594	Immunoglobulin kappa variable 1-33
KV106	P01598	Immunoglobulin kappa variable 1-5
KV305	P01623	Immunoglobulin kappa variable 3-20
KV308	P04207	Immunoglobulin kappa variable 3-15
KV309	P04433	Immunoglobulin kappa variable 3-11
KV404	P06314	Immunoglobulin kappa variable 4-1
LAC2	P0CG05	Ig lambda-2 chain C regions
LG3BP	Q08380	Galectin-3-binding protein
LIPB2	Q8ND30	Liprin-beta-2
LV106	P04208	Immunoglobulin lambda variable 1-47
LV302	P80748	Immunoglobulin lambda variable 3-21
LV403	P01717	Immunoglobulin lambda variable 3-25
PEDF	P36955	Pigment epithelium-derived factor
PGRP2	Q96PD5	N-acetylmuramoyl-L-alanine amidase
PK3CG	P48736	Phosphatidylinositol 4,5-bisphosphate 3-kinase catalytic subunit gamma isoform
PLF4	P02776	Platelet factor 4
PLMN	P00747	Plasminogen
PON1	P27169	Serum paraoxonase/arylesterase 1
PROP	P27918	Properdin
PROS	P07225	Vitamin K-dependent protein S
RET4	P02753	Retinol-binding protein 4
SAA4	P35542	Serum amyloid A-4 protein
SHBG	P04278	Sex hormone-binding globulin
SHC1	P29353	SHC-transforming protein 1
THRB	P00734	Prothrombin
TRFE	P02787	Serotransferrin
TRPV2	Q9Y5S1	Transient receptor potential cation channel subfamily V member 2
TSP1	P07996	Thrombospondin-1
VTDB	P02774	Vitamin D-binding protein
VTNC	P04004	Vitronectin

Supplementary table E2. Comparison of protein expression between women with LAM and control women.

		LAM vs control	
Protein	UniProt ID	Fold change (Log2)	Confidence
VTDB	P02774	-2.6	0.65
ITIH4	Q14624	-0.5	0.34
HMMR	O75330	0.5	0.31
FETUA	P02765	-2.8	0.30
AMBP	P02760	-0.5	0.28
TRFE	P02787	-2.0	0.22
ALBU	P02768	0.5	0.21
FIBA	P02671	-0.4	0.17
ITIH2	P19823	0.5	0.16
IGHG3	P01860	3.1	0.16
SAA4	P35542	0.7	0.15
KI67	P46013	-1.3	0.15
HEMO	P02790	0.5	0.13
APOL1	O14791	-0.5	0.13
VTNC	P04004	-0.5	0.12
APOA1	P02647	0.8	0.12
CERU	P00450	-0.5	0.12
CXCL7	P02775	-0.5	0.12
GELS	P06396	2.1	0.12
LV106	P04208	-1.0	0.11
CFAB	P00751	-0.5	0.11
KIFC2	Q96AC6	1.7	0.10
A1AT	P01009	-0.6	0.09
CO2	P06681	-0.5	0.09
CFAH	P08603	-0.3	0.09
APOH	P02749	2.8	0.09
THRB	P00734	-0.3	0.08
CO3	P01024	-0.3	0.08
IGHM	P01871	0.6	0.08
CBPN	P15169	-0.4	0.08
IGHA1	P01876	-0.7	0.07
CLUS	P10909	0.3	0.07
C4BPA	P04003	0.5	0.07
KNG1	P01042	0.3	0.07
TRPV2	Q9Y5S1	-1.0	0.06
PK3CG	P48736	-1.2	0.06
CD5L	O43866	2.4	0.05
KV106	P01598	-1.4	0.05
IGJ	P01591	-0.5	0.05
APOA2	P02652	0.3	0.05

A2MG	P01023	4.0	0.05
HRG	P04196	0.4	0.05
AFAM	P43652	0.4	0.05
LAC2	P0CG05	-3.0	0.05
APOA4	P06727	0.4	0.05
LV302	P80748	0.5	0.05
A2GL	P02750	0.8	0.05
CO4B	P0C0L5	2.3	0.05
APOA	P08519	-1.1	0.05
APOD	P05090	-0.4	0.05
HBA	P69905	0.9	0.05
PLMN	P00747	-0.3	0.04
HV101	P01742	-2.0	0.04
KV309	P04433	0.5	0.04
KV102	P01594	0.4	0.04
IC1	P05155	0.5	0.04
CO8A	P07357	-0.3	0.04
PROS	P07225	0.4	0.04
IGKC	P01834	0.5	0.04
PON1	P27169	0.9	0.04
SHC1	P29353	0.3	0.04
HPT	P00738	1.1	0.04
IBP3	P17936	-0.5	0.04
RET4	P02753	-0.7	0.03
HDAC1	Q13547	-0.4	0.03
C1S	P09871	0.2	0.03
C1QC	P02747	0.4	0.03
KCNC4	Q03721	-0.9	0.03
CAMP	P49913	-0.5	0.03
LG3BP	Q08380	-0.7	0.03
ANT3	P01008	-0.3	0.03
A1AG1	P02763	0.8	0.03
H2AX	P16104	-1.2	0.03
FA12	P00748	-0.6	0.03
PROP	P27918	-0.5	0.03
CFAI	P05156	-0.3	0.03
A1BG	P04217	0.3	0.03
IGLL5	B9A064	0.5	0.03
IGHG1	P01857	-0.7	0.02
CO9	P02748	-0.3	0.02
ACTG	P63261	-0.8	0.02
HBB	P68871	0.9	0.02
APOB	P04114	-0.5	0.02
APOC3	P02656	-0.4	0.02
FINC	P02751	-0.3	0.02
IGHD	P01880	-1.0	0.02

FBLN1	P23142	-0.5	0.02
APOC2	P02655	0.5	0.02
PGRP2	Q96PD5	-0.4	0.02
APOF	Q13790	-0.7	0.02
ITIH1	P19827	0.5	0.02
FCN2	Q15485	0.5	0.02
FHR3	Q02985	0.6	0.01
KANK3	Q6NY19	-0.9	0.01
FCG3A	P08637	-0.5	0.01
ANGT	P01019	-0.4	0.01
HV311	P01772	-0.6	0.01
C1R	P00736	-0.7	0.01
APOE	P02649	0.4	0.01
CO6	P13671	-0.7	0.01
HS12B	Q96MM6	-0.3	0.01
APOM	O95445	0.2	0.01
IGHA2	P01877	0.3	0.01
TSP1	P07996	0.5	0.01
IGHG2	P01859	0.9	0.01
SHBG	P04278	-0.8	0.01
K1024	Q9UPX6	-0.4	0.01
LV403	P01717	-0.2	0.01
HPTR	P00739	-0.6	0.01
FOXN3	O00409	0.7	0.01
KLKB1	P03952	-0.9	0.01
CO5	P01031	-0.6	0.01
HS90B	P08238	-0.6	0.01
HV304	P01765	-0.6	0.01
HV305	P01766	0.6	0.01
PLF4	P02776	-0.2	0.00
KV305	P01623	0.7	0.00
PEDF	P36955	0.6	0.00

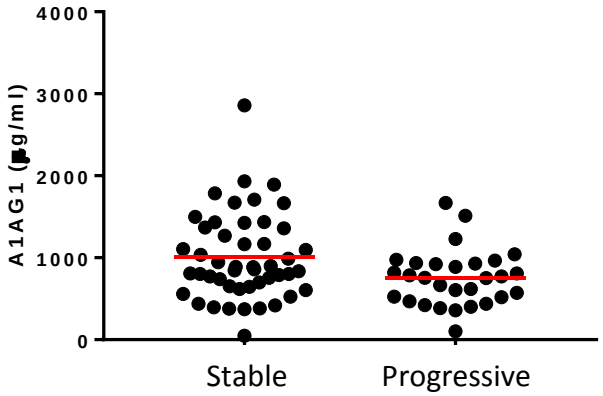
Protein differences are expressed as fold change ranked by significance.

Supplementary table E3. *GC* allele frequencies in control women, UK and NHLBI LAM cohorts

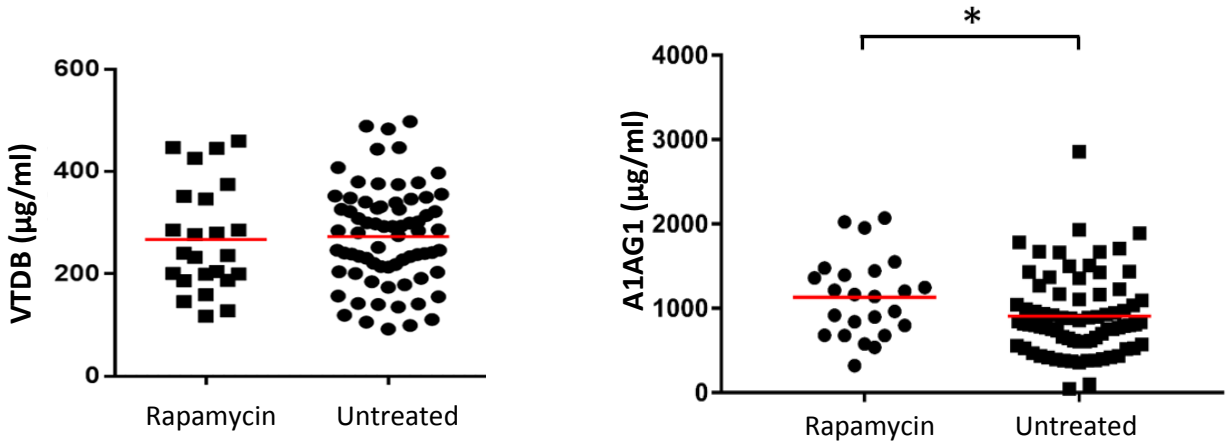
Genotype	Cohort				p value		
	Controls	UK LAM	NHLBI LAM		control vs. UK LAM	control vs. NHLBI LAM	UK LAM vs. NHLBI LAM
n	168141	65	rs7041	rs4588			
			145	146			
rs7041							
GG	31	22	37	-	0.076	0.29	0.075
GT	50	49	43				
TT	19	29	20				
rs4588							
AA	8	14	-	8	0.20	0.17	0.43
AC	42	34		34			
CC	50	52		58			

Percentage allele frequencies are shown for women of European ancestry in the three cohorts. Allele frequencies were compared using the Chi-squared test.

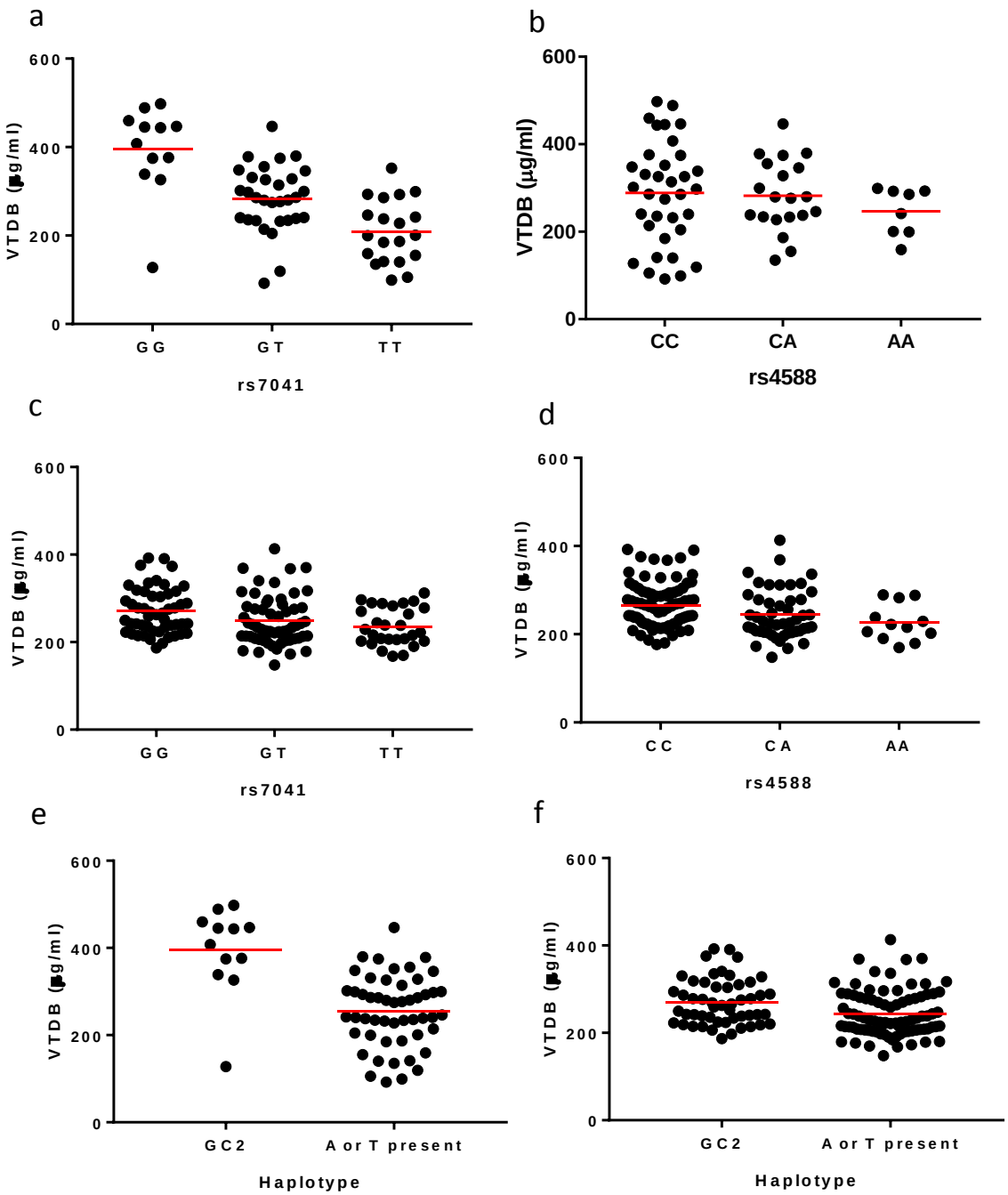
Supplementary figures



Supplementary figure E1. Relationship between serum Alpha1-acid glycoprotein (A1AG1) and disease activity. Serum A1AG1 of women with stable LAM is significantly higher than those with progressive disease (p=0.01).



Supplementary figure E2. Serum Vitamin D Binding Protein (VTDB) levels in patients with LAM untreated or treated with rapamycin. Serum A1AG1 levels in rapamycin treated LAM compared with untreated LAM, (*p=0.031).



Supplementary figure E3. Relationship between *GC* genotype and serum VTDB. In the UK LAM cohort, the presence of the T allele at rs7041 was dose dependently associated with lower serum VTDB levels (n=63, p<0.0001, panel a) although rs4588 was not associated with serum VTDB level (p=0.57, panel b). In the NHLBI Registry cohort, the T allele at rs7041 and the A allele at rs4588 were dose dependently associated with lower serum VTDB levels (n=139, p=0.010 and n=140, p=0.035 respectively, panels c and d). Haplotype analysis combining the allelic information at both SNPs showed the presence of the minor alleles at either rs7041 or rs4588 (T and A respectively) were associated with lower serum VTDB levels in both the UK and NHLBI cohorts, p<0.0001 and p=0.0018, respectively, panels e and f).