


Proteomics discovery of pulmonary hypertension biomarkers: Insulin-like growth factor binding proteins are associated with disease severity

Melanie K. Nies¹ | Jun Yang¹  | Megan Griffiths^{1,2} | Rachel Damico³ | Jie Zhu¹ | Dhananjay Vaydia^{3,4} | Zongming Fu⁵ | Stephanie Brandal¹ | Eric D. Austin⁶ | Dunbar D. Ivy⁷ | Paul M. Hassoun³ | Jennifer E. Van Eyk^{8,9} | Allen D. Everett¹

¹Department of Pediatrics, Division of Cardiology, Johns Hopkins University, Baltimore, Maryland, USA

²Department of Pediatrics, Division of Pediatric Cardiology, Columbia University, New York, New York, USA

³Department of Medicine, Division of Pulmonary and Critical Care Medicine, Johns Hopkins University, Baltimore, Maryland, USA

⁴Department of Epidemiology, School of Public Health, Johns Hopkins University, Baltimore, Maryland, USA

⁵Department of Pediatrics, Division of Hematology, Johns Hopkins University, Baltimore, Maryland, USA

⁶Department of Pediatrics, Division of Allergy, Immunology, and Pulmonary Medicine, Vanderbilt University, Nashville, Tennessee, USA

⁷Department of Pediatric Cardiology, Children's Hospital Colorado, Aurora, Colorado, USA

⁸Department of Internal Medicine, Division of Cardiology, Johns Hopkins University, Baltimore, Maryland, USA

⁹Advanced Clinical Biosystems Research Institute, Cedars-Sinai Medical Center, Los Angeles, California, USA

Correspondence

Allen D. Everett, Department of Pediatrics, Blalock-Taussig-Thomas Division of Cardiology, Johns Hopkins University, 1800 Orleans St, Room M2321, Baltimore, MD 21287, USA.
Email: aeveret3@jhmi.edu

Funding information

National Institute of Child Health and Human Development, Grant/Award Number: K12-HD000850; National Heart, Lung, and Blood Institute, Grant/Award Numbers: NHLBI-HV-10-05 (2), R01HL135114, R01HL150070, R03HL110830, R24HL123767; Matthew and Michael Wojciechowski Pulmonary Hypertension Pediatric Proof-of-Concept Grant (Dr. Robyn J. Barst Pediatric PH Research and Mentoring Fund Grant);

Abstract

Pulmonary arterial hypertension (PAH) is a progressive disease characterized by sustained elevations of pulmonary artery pressure. To date, we lack circulating, diagnostic, and prognostic markers that correlate to clinical and functional parameters. In this study, we performed mass spectrometry-based proteomics analysis to identify circulating biomarkers of PAH. Plasma samples from patients with idiopathic pulmonary arterial hypertension (IPAH, $N = 9$) and matched normal controls ($N = 9$) were digested with trypsin and analyzed using data-dependent acquisition on an Orbitrap mass spectrometer. A total of 826 (false discovery rate [FDR] 0.047) and 461 (FDR 0.087) proteins were identified across all plasma samples obtained from IPAH and control subjects, respectively. Of these, 153 proteins showed >2 folds change ($p < 0.05$) between groups. Circulating levels of carbonic anhydrase 2 (CA2), plasma kallikrein (KLKB1), and the insulin-like growth factor binding proteins (IGFBP1-7) were quantified by immunoassay in an independent verification cohort ($N = 36$

Melanie K. Nies and Jun Yang contributed equally to this study.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2022 The Authors. *Pulmonary Circulation* published by John Wiley & Sons Ltd on behalf of Pulmonary Vascular Research Institute.

Cardiovascular Medical Research and Education Fund (CMREF)

PAH and $N = 35$ controls). CA2 and KLKB1 were significantly different in PAH versus control but were not associated with any functional or hemodynamic measurements. Whereas, IGFBP1 and 2 were associated with higher pulmonary vascular resistance, IGFBP2, 4, and 7 with decreased 6-min walk distance (6MWD), and IGFBP1, 2, 4, and 7 with worse survival. This plasma proteomic discovery analysis suggests the IGF axis may serve as important new biomarkers for PAH and play an important role in PAH pathogenesis.

KEYWORDS

biomarker, IGF binding protein, proteomics, pulmonary hypertension

INTRODUCTION

Pulmonary arterial hypertension (PAH) in children or adults is a heterogeneous disease defined by elevated mean pulmonary arterial pressure (pulmonary vascular resistance [PVR]), without a known cure. PAH is progressive, but the inciting event in individual patients for progression to fatal right ventricular (RV) decompensation, as well as the time course, is unknown.¹ While cardiac catheterization for hemodynamic assessment remains the gold standard for diagnosis and evaluation of pulmonary hypertension, clinicians often rely on noninvasive means for screening, therapeutic and prognostic monitoring. However, our current noninvasive diagnostic/prognostic state of the art, including echocardiography, 6-min walk distance (6MWD) and amino-terminal pro-brain natriuretic peptide (NT-proBNP) levels, have limitations and are incomplete.^{2,3} Echocardiography can be limited by poor acoustic windows and inability to obtain reliable spectral doppler signal of the tricuspid valve regurgitant jet necessary for estimating RV pressure. In addition, RV function can be difficult to ascertain by echocardiography given the complex geometry of the RV, and 6MWD can't be reliably obtained in young patients. The current state of the art clinical blood biomarker NT-proBNP is not lung/vascular specific, confounded by systemic diseases,² and levels differing significantly by age.³ A noninvasive, reliable, and more lung/vascular specific measure of PAH could improve diagnostic and prognostic monitoring. We hypothesized that proteomic evaluation of plasma obtained from patients with PAH would reveal quantitative protein changes leading to the identification of new PAH specific circulating, diagnostic and prognostic markers.

In this study, we developed an agnostic, in-depth plasma proteomics approach coupled with functionally phenotyped enrollees in the Pulmonary Hypertension Breakout Initiative (PHBI) before lung transplantation, identified and verified a number of circulating proteins with significant changes in PAH cases as potential circulating markers.

METHODS

Study cohorts

All study cohorts were approved by Institutional Review Board, and informed consent was obtained from all subjects by the PHBI (<https://www.ipahresearch.org>).

Discovery cohort

Ethylendiaminetetraacetic acid plasma analysis was performed on samples from PHBI participants ($N = 9$) with BMP2R mutation negative IPAH. The PHBI is a highly phenotyped cohort of patients with severe PAH, funded by the Cardiovascular Medical Research and Education Fund (CMREF). Blood samples were obtained from PHBI participants before lung transplant in accordance with the PHBI study protocol. Age and sex matched non-PAH controls were purchased from Bioreclamation (BioIVT). See Table 1 and supplemental methods.

Verification disease cohort

Proteomic discovery results were verified using an independent cohort of adult patients cared by the Johns Hopkins Pulmonary Hypertension Program (JHPH). The serum samples of the patients were collected at the time of enrollment. The inclusion, exclusion criteria, and clinical assessments and therapy were published previously.⁴ Briefly, participants 18 years old or greater diagnosed with PAH by right heart catheterization and evaluated between January 1, 2007 and December 31, 2012 were included and classified into etiologic groups based on current guidelines.¹ This cohort was used for verification, clinical correlations, and survival analysis.

TABLE 1 PHBI and control cohorts demographics and characteristics

	PHBI	Control
<i>n</i>	9	9
Age (years)	35.2 ± 11.0	34.8 ± 10.6
Female, <i>n</i> (%)	8 (89%)	9 (100%)
Race- EA, AA, other <i>n</i> (%)	6/1/2 (67/11/22%)	1/5/3 (11/55/33%)
IPAH, <i>n</i> (%)	9 (100%)	NA
6MWD (m)	319 (range 161–515)	NA
NYHA-FC, <i>n</i>		
III	6	-
IV	2	-
No data	1	-
Laboratory chemistries		
BNP, pg/ml (<i>n</i>)	361 (8), range 147–631	-
Hemodynamics		
RAP (mmHg)	13.9 ± 4.9	-
mPAP (mmHg)	52.8 ± 10.9	-
PCWP (mmHg)	10.0 ± 5.0	-
PVR (WU)	11.5 ± 7.0	-
CO (L min ⁻¹)	4.0 ± 1.6	-
CI (L min ⁻¹ m ⁻²)	2.8 ± 1.3	-

Note: Data expressed as mean and standard deviation (SD), number (*n*), percentage (%), or range as indicated.

Abbreviations: AA, African American; BNP, brain natriuretic peptide; CI, cardiac index; CO, cardiac output; EA, European American; mPAP, mean pulmonary artery pressure; NYHA-FC, New York Heart Association-functional class; PCWP, pulmonary capillary wedge pressure; PHBI, Pulmonary Hypertension Breakout Initiative; PVR, pulmonary vascular resistance; RAP, right atrial pressure; 6MWD, 6-min walk distance.

Verification control cohort

Control serum samples for ELISA verification were adult volunteers without PAH enrolled at Johns Hopkins University (*n* = 35).

Mass spectrometry data acquisition and analysis

Sample preparation

Samples underwent abundant protein depletion using a Seppro IgY14 LC10 column (Sigma Aldrich) per manufacturer's instruction using the Beckman Coulter

ProteomeLab PF2D. Briefly, a 170 µl plasma sample was diluted with 950 µl 1× dilution buffer and filtered by centrifugation (Spin-X, 2.0 ml Polypropylene Tube, COSTAR). A 900 µl filtered sample was injected on to the depletion column at a flow rate of 2 ml/min, with the unbound flow through collected at 10–30 min intervals. To further decrease sample complexity, intact depleted protein samples (150 µg) were concentrated and fractionated using reverse phase high-performance liquid chromatography (HPLC). The depleted protein was diluted with column eluent acetonitrile: trifluoroacetic acid (50:2.5 V/V), to a total 5 ml solution, centrifugation at 3400×G, and loaded onto the HPLC system (Beckman coulter HPRP module, 5 ml loop, coupled to a 96-well plate auto collector and UV detector at 214 nm, Column: Jupiter C18). Forty fractionations were collected in 96-well plates with 1 M Na bicarbonate to neutralize pulmonary hypertension (PH). The adjacent samples were combined and dried by speed vacuum. Each of the 20 samples were rehydrated with 70 µl 0.01 µg/µl trypsin (sequencing grade modified, Promega). The trypsin:protein ratio was maintained at 1:5, at pH 8, and incubated at 37°C overnight for full digestion. Aliquots were tested by SDS/PAGE to confirm complete digestion before proceeding with mass spectrometry. Digestion was stopped with 4% phosphoric acid.^{5,6} Finally, the tryptic peptides were desalted by HLB plate (Waters Oasis desalting column).

MS analysis of the digested peptides

The digested peptides were analyzed using label-free quantification by reverse-phase liquid chromatography with tandem mass spectrometry (RPLC-MS/MS). An Orbitrap Elite mass spectrometer (Thermo Scientific) coupled to an Easy-nLC 1000 source (Thermo Scientific) was used. The dried peptides were reconstituted in 20 µl of 0.1% formic acid and 5% of acetonitrile in mass spectrometry grade water. Ten microliters of the peptides were concentrated on a C18 trap column (Acclaim PepMap 100) in 0.1% aqueous formic acid and then subsequently separated on an EASY-Spray™ ES801A C18 (Thermo Scientific) analytical column using a linear solvent gradient from 5% to 35% acetonitrile over 65 min with a flow rate of 250nl/min. The analysis was operated in a data-dependent mode with full scan MS spectra acquired in the Orbitrap analyzer at a 60,000× resolution, followed by MS2 acquisition of the top 10 ions in the ion trap.

Database searching and processing

All raw data from the Orbitrap Elite were converted to mzXML format using Msconvert from ProteoWizard for

peaklist generation. All MS/MS samples were analyzed using Sorcerer 2-SEQUENT algorithm (Sage-N Research) searched against the concatenated target and decoy Human Uniprot 2016 database. The search settings included trypsin as the digestion enzyme, parent ion tolerance of 20 PPM and fragment ion tolerance of 0.6 Da, carbamidomethyl of cysteine as a fixed modification, and oxidation of methionine as a variable modification. The post-search analysis was performed using Scaffold 4 (Proteome Software, Inc.) with protein and peptide probability thresholds set to 95% and 90%, respectively, and a minimum of two peptides required for identification. Relative protein quantification was obtained from MS data using spectral counting.

Immunoassay verification

Human CA2 (LS-F21799, Lifespan Bioscience Inc.) and Human plasma kallikrein (KLKB1) sandwich ELISA kits (EKC35033, Biomatik Corp.) were used to measure circulating CA2 and KLKB1 levels, following the manufacturer's instructions.

Based on sample dilution characteristic, two multiplex ELISAs for insulin-like growth factor binding proteins (IGFBP) (1, 4), and IGFBP (2, 3, 6) were developed on the MSD platform (Meso Scale Diagnostics) using Duoset reagents from R&D Systems. The capture antibodies for IGFBP1 and IGFBP4 (DY871, DY804, R&D Systems) were commercially robotically spotted on MSD plates, calibrators and biotin labeled detection antibodies from the same Duosets were used to optimize the performance of individual ELISAs. The quality of the performance of the ELISAs were equal or better than the ones that Duosets described by manufacturer (R&D Systems). The same approach was used to build a 3-plex ELISA for IGFBP2, IGFBP3, and IGFBP6 (DY674, DY675, DY876, R&D Systems, MN). IGFBP5 and IGFBP7 were measured using commercial ELISA kits (ELH-IGFBP5-1, ELH-IGFBPRP1-1, Raybiotech), following the manufacture's instruction.

Statistical analysis

Metaboanalyst (<https://www.metaboanalyst.ca/>)^{7–10} was used to compare spectral counts for identified plasma proteins between PAH and control groups. Using unpaired *t*-tests we compared the fold change (FC) in protein concentrations in PAH and controls; this was visualized by volcano plot with a FC threshold of ≥ 2 , $p < 0.05$ and an false discovery rate (FDR) < 0.008 . Spectral counts for identified proteins were log transformed for normality,

centered, and scaled. Dimension reduction was done using principal component analysis (PCA) to better understand the important features in the data. Relationships between clinical variables (hemodynamics and functional measures) were explored by spearman's rank correlation. The area under the curve (AUC) of the receiver operating characteristics (ROC) curve was calculated to evaluate the performance of biomarkers as a discriminator of the presence of PAH. Kaplan–Meier analysis and Cox proportional hazard ratios adjusted for age were performed to compare survival distributions. Statistical analysis was performed using MedCalc Statistical Software version 20.009 (MedCalc Software) or R statistical packages (version 4.1.0) when appropriate.

RESULTS

IPAH discovery cohort

Plasma samples from nine PHBI IPAH and nine healthy adults were matched for age and sex. As shown in Table 1, the IPAH patients had severe PAH with a functional class of III–IV, 6MWD mean of 319 meters, mean PA pressure of 52.8 ± 10.9 mmHg and PVR of 11.5 ± 7 WU. Compared to the IPAH group, the controls were similar for age and sex.

IPAH and control protein discovery

The total number of proteins quantified were 826 (FDR 0.047) and 461 (0.087 FDR) for IPAH and control respectively. Four hundred twenty-three proteins were unique to the IPAH and 58 unique to the control cohorts (Figure 1). PCA generated five principal

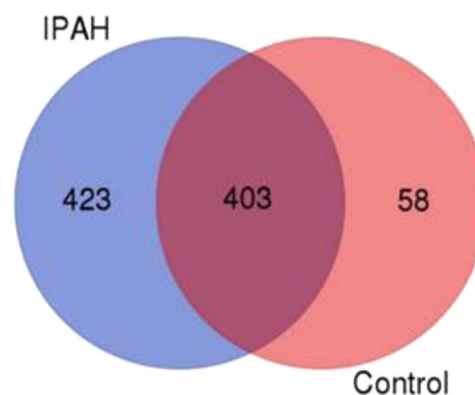


FIGURE 1 Venn diagram of protein identification overlap between idiopathic pulmonary arterial hypertension (IPAH) and control cohorts

components with loads >5%, which together described 60.6% of variability in the data. A scores plot (Figure 2) demonstrated significant separation of principal components between the PAH group (red) and the control (green). Of the top 15 differentially expressed proteins, 9 proteins were increased and 6 proteins were decreased in IPAH compared to controls (Figure 3,

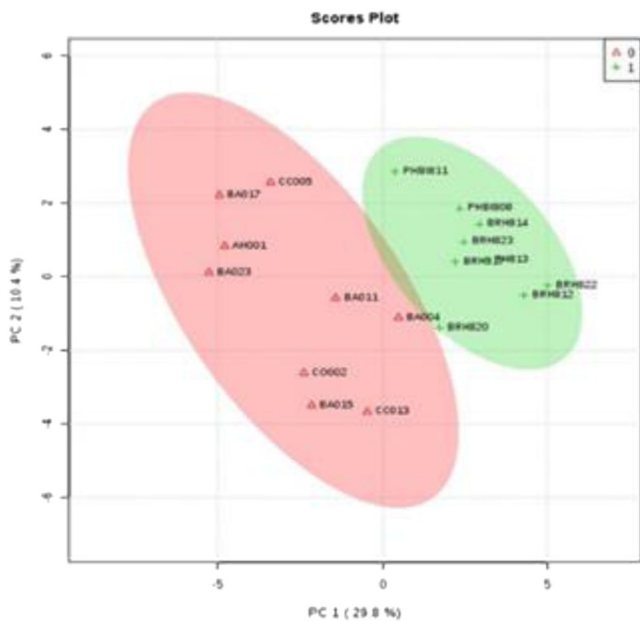


FIGURE 2 Principle component analysis scores plot of overlap of idiopathic pulmonary arterial hypertension and control protein datasets

Table 2). Findings were confirmed with significance analysis of microarray analysis to account for multiple testing (Table S1). These proteins included enzyme regulator activity (carboxypeptidase N subunit 2), smooth muscle cell proteins (vinculin, alpha-actinin, transgelin-2), protein synthesis (elongation factor 1-alpha), metabolism (carbonic anhydrase 2 [CA2], fructose-bisphosphate aldolase A, IGF binding protein 2 [IGFBP2]), cell cycle (14-3-3 protein sigma), serine-type endopeptidase activity (kallikrein B1 [KLKB1]), calcium ion binding (serum paraoxonase/arylesterase 1) and redox (superoxide dismutase [Cu-Zn]). Based on the availability of high-quality ELISA reagents, CA2 and KLKB1 were moved forward for verification.

We have previously shown the association between the proteins, IGF1, IGF2, IGFBP2, and PAH in a larger cohort,⁴ but little information regarding the other IGF binding proteins in PAH is available. Based on biologic plausibility and that the IGF binding protein family has seven members, we extracted spectral counts data from MS raw files, and found that all seven binding proteins were detectable in most of these plasma samples. As shown in Table 3, using Mann-Whitney test to compare the control and PAH patients, the spectral counts for additional IGF binding proteins other than IGFBP2 were altered in IPAH. As shown, IGFBP1, IGFBP2, and IGFBP7 were significantly different ($p < 0.05$). To verify the accuracy and efficiency of our proteomic strategy, based on their potential biological relevance to PAH, all the IGF binding proteins were chosen as additional verification targets in this study.

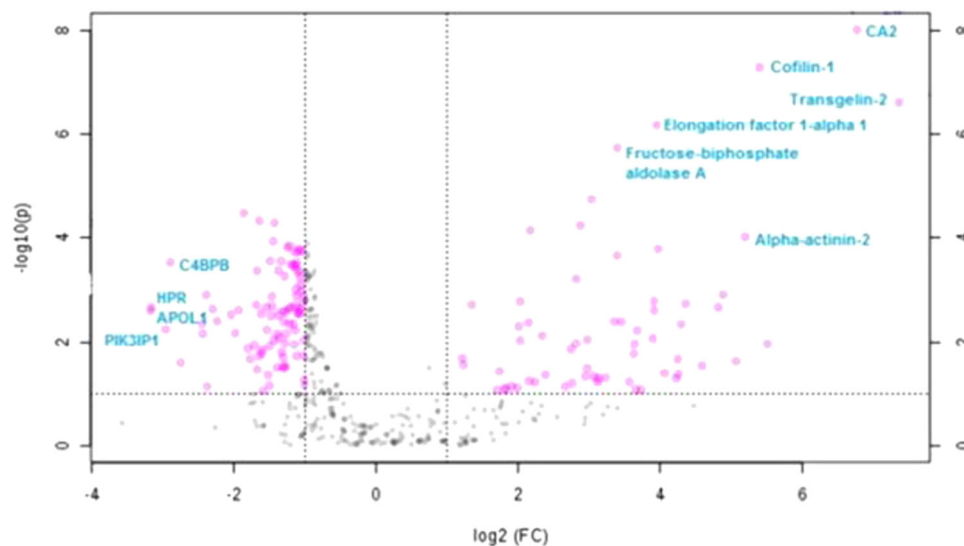


FIGURE 3 Important features selected by volcano plot with fold change (FC) threshold (\times) 2 and t -tests threshold (y) 0.1. The red circles represent features above the threshold. Note both FCs and p values are \log -transformed. The further its position away from the (0,0), the more significant the feature is. All features in Table S1

Protein name	Gene name	FC	log ₂ (FC)	p-value
Carbonic anhydrase 2	CA2	108.87	6.7665	9.61E-09
Cofilin-1	CFL1	42.256	5.4011	5.04E-08
Transgelin-2, isoform 2	TAGLN2	164.33	7.3604	2.44E-07
Elongation factor 1-alpha 1	EEF1A1	15.713	3.9739	6.55E-07
Fructose-bisphosphate aldolase A	ALDOA	10.525	3.3958	1.81E-06
Superoxide dismutase (Cu-Zn)	SOD1	8.1928	3.0344	1.79E-05
Plasma kallikrein	KLKB1	0.27738	-1.8501	3.32E-05
Serum paraoxonase/arylesterase 1	PON1	0.32125	-1.6382	4.69E-05
Carboxypeptidase N subunit 2	CPN2	0.37298	-1.4228	5.15E-05
14-3-3 protein sigma	SFN	7.3566	2.879	5.74E-05
Insulin-like growth factor-binding protein 2	IGFBP2	4.5149	2.1747	7.14E-05
Alpha-actinin, isoform 2	ACTN2	36.7	5.1977	9.61E-05
Coagulation factor IX	F9	0.36834	-1.4409	0.000116
Biotinidase	BTD	0.43334	-1.2064	0.000141
Carboxypeptidase B2	CPB2	0.42306	-1.2411	0.000152

Abbreviation: FC, fold change.

TABLE 3 IGF binding protein family spectral counts change in PAH and controls

	Spectral counts (mean, range)		Fold change	p-value ^a
	PAH	Control		
IGFBP1	1.7 (0-3)	0.1 (0-1)	17.0	0.002
IGFBP2	47.2 (12-104)	5.4 (0-15)	8.8	<0.0001
IGFBP3	34.9 (9-51)	50 (28-84)	0.7	0.123
IGFBP4	11.9 (7-19)	9.7 (3-21)	1.2	0.171
IGFBP5	3.5 (1-8)	4.3 (1-8)	0.8	0.565
IGFBP6	4.6 (1-7)	4.3 (1-7)	1.1	0.759
IGFBP7	8.7 (3-21)	2.9 (0-5)	3.0	0.002

Abbreviation: PAH, pulmonary arterial hypertension.

^aMann-Whitney test p-value.

CA2 and KLKB1 verification

A verification disease cohort of JHPH ($N = 36$) and control ($N = 35$) serum samples were obtained at a single time point (enrollment) and used for ELISA verification analysis (Table 4). The median age of patients at the time of enrollment was 62 (interquartile range: 53-69), the JHPH cohort was composed of both IPAH (44%) and associated pulmonary arterial hypertension (APAH, 56%) cases, with moderate to severe PAH with a median PA

TABLE 2 Volcano plot top 15 features

pressure of 40 mmHg, and PVR of 6.3 WU. Most of the patients were categorized into functional class II and III (44% and 33%, respectively), with 5.6% in functional class IV. The median follow-up was 4.6 years, with maximum 9.5 years. Mortality in the PAH group was 61% during follow-up period.

As shown in Table 5, both median serum CA2 and KLKB1 concentrations were significantly increased in JHPH vs control (241.9 ng/ml vs. 188.1 ng/ml, $p < 0.0001$; $4.7 \mu\text{g/ml}$ vs. $4.0 \mu\text{g/ml}$, $p = 0.01$), respectively. In ROC analysis, the quantity of circulating CA2 (AUC 0.845, $p < 0.0001$) and KLKB1 (AUC 0.636, $p = 0.031$) were sufficient to distinguish PAH from controls (Table 6, Figure 4a,b). Interestingly, neither CA2 nor KLKB1 correlated (spearman's rank correlation) with any hemodynamic, functional measures of PH severity (Table 7) or mortality.

IGF binding proteins and PAH

Using two custom built multiplex ELISA assays and two commercial immunoassays, we evaluated all seven IGF binding proteins in the PAH and control cohorts (Table 5). As shown, the direction of changes of the serum IGF binding protein concentrations by the immunoassays were largely consistent with the MS results with IGFBP1, 2, 4, 7 significantly increased in PAH, IGFBP3 significantly decreased in PAH, and IGFBP6 unchanged.

TABLE 4 Demographic of JHPH and control cohorts

Variable	N	PAH (N = 36)	Healthy control (N = 35)	p-value
Sex (male/female)	-	9/27	7/28	0.78
Age (years)	-	62 (53–69)	45 (36–56)	<0.001
Race (EA/AA/unknown)	-	10/1/25	27/5/3	0.44
Death	-	22	0	<0.001
PAH subtype	36	-	-	-
IPAH	-	16 (44%)	-	-
APAH	-	20 (56%)	-	-
Hemodynamics				
mRAP (mmHg)	36	7.0 (4.0–10.2)	-	-
mPAP (mmHg)	36	40 (32–48)	-	-
mPCWP (mmHg)	36	9.5 (7.0–11.2)	-	-
RVSP (mmHg)	28	62 (52–79)	-	-
PVR (WU)	36	6.3 (4.6–9.7)	-	-
PVRI (WU/m ²)	34	12.0 (8.2–17.7)	-	-
CO (L/min)	36	4.60 (3.80–5.58)	-	-
CI (L/min/m ²)	34	2.40 (2.20–2.98)	-	-
Functional measures				
6-min walk distance (m)	32	399 (310–458)	-	-
NYHA functional class				
1	-	6 (17%)	-	-
2	-	16 (44%)	-	-
3	-	12 (33%)	-	-
4	-	2 (5.6%)	-	-
Laboratory	25	-	-	-
Clinical NT-proBNP (pg/ml)	-	943 (289–2768)	-	-

Note: Data expressed as median (interquartile range), number, and percentage when appropriate. Abbreviations: APAH, associated pulmonary arterial hypertension; CI, cardiac index; CO, cardiac output; IPAH, idiopathic pulmonary arterial hypertension; JHPH, Johns Hopkins Pulmonary Hypertension Program; mPAP, mean pulmonary artery pressure; mPCWP, pulmonary capillary wedge pressure; mRAP, mean right atrial pressure; NT-proBNP, amino-terminal pro-brain natriuretic peptide; NYHA, New York Heart Association; PAH, pulmonary arterial hypertension; PVR, pulmonary vascular resistance; PVRI, pulmonary vascular resistance index; RVSP, right ventricular systolic pressure

IGF binding proteins discriminate PAH from healthy control

IGF binding protein values from the verification PAH and control cohorts were used to generate ROC curves. As shown in Table 6 and Figure 4c, IGFBP1, 2, 4, 5, 7 discriminated PAH from healthy control with an AUC > 0.7. IGFBP5 had the largest AUC at 0.945, $p < 0.0001$. Youden analysis established a threshold of IGFBP5 value at 60.6 ng/ml to distinguish PAH from

control; this threshold value had a sensitivity and specificity for PAH of 83.3% and 85.8%, respectively.

IGF binding proteins and PAH severity

As shown in Table 7, IGFBP 2, 4, and 7 were inversely correlated with 6MWD, with a higher protein concentration correlating with a shorter 6MWD. Higher concentrations of IGFBP 1 and 2 were correlated with

	PAH (N = 36) Median (IQR)	Healthy control (N = 35) Median (IQR)	p-value
CA2 (ng/ml)	241.9 (225.2–258.4)	188.1 (163.4–217.1)	<0.001*
KLKB1 (U/ml)	4.7 (3.5–5.5)	4.0 (3.2–4.6)	0.01*
IGFBP1 (ng/mL)	2579 (1301–5225)	1074 (416–1703)	0.001*
IGFBP2 (ng/mL)	220 (148–414)	127 (93–167)	<0.001*
IGFBP3 (ng/mL)	2442 (1924–3138)	2925 (2489–3485)	0.05*
IGFBP4 (ng/mL)	792 (531–978)	462 (405–511)	<0.001*
IGFBP5 (ng/mL)	76 (69–98)	0 (0–0)	<0.001*
IGFBP6 (ng/mL)	246 (147–361)	248 (173–297)	0.8
IGFBP7 (ng/mL)	124 (79–223)	63 (55–93)	<0.001*

Abbreviations: IQR, interquartile range; PAH, pulmonary arterial hypertension.

* $p < 0.05$.

TABLE 6 AUC analysis of biomarkers

	AUC	SE	95% CI
CA2	0.845	0.049	0.736–0.921
KLKB1	0.636	0.063	0.520–0.742
IGFBP1	0.721	0.062	0.601–0.821
IGFBP2	0.775	0.0582	0.660–0.865
IGFBP3	0.636	0.067	0.513–0.747
IGFBP4	0.839	0.0497	0.733–0.916
IGFBP5	0.945	0.0273	0.864–0.985
IGFBP6	0.515	0.0704	0.393–0.635
IGFBP7	0.753	0.0591	0.637–0.848

Abbreviations: AUC, area under the curve; CI, confidence interval.

higher PVR, higher IGFBP4 with higher pulmonary capillary wedge pressure, and higher IGFBP1 with lower cardiac output (Table 7). IGFBP3, 5, and 6 were not associated with clinical severity.

IGF binding proteins association with PAH mortality

Kaplan–Meier curves were generated to assess the relationship between IGF binding protein levels (dichotomized at the median) and mortality. In unadjusted Kaplan–Meier survival analysis, above the median blood concentrations of IGFBP1, 2, 4, and 7 were associated with significantly worse survival, with a hazard ratio of 3.75 (95% confidence interval [CI]: 1.56–8.99, $p = 0.003$); 3.43 (95% CI: 1.41–8.32, $p = 0.007$); 3.27 (95% CI: 1.35–7.89, $p = 0.009$); 3.28 (95% CI: 1.3336–7.91,

TABLE 5 Biomarkers in PAH and control cohorts

$p = 0.008$), respectively (Figure 5). Multivariable Cox proportional hazards models were constructed to analyze the association between biomarkers and mortality, proportionality tests were performed to confirm the assumption was met. After adjusting for age, IGFBP1, 2, 4, and 7 were independently associated with survival (Table 8).

DISCUSSION

Pulmonary hypertension is a chronic severe disease of the pulmonary vasculature where a paucity of pulmonary vascular specific markers makes noninvasive monitoring difficult. Using unbiased high-resolution mass spectrometry, we demonstrated that a number of proteins were altered in IPAH versus control. We identified and then verified that the circulating concentration of CA2, KLKB1, and multiple IGFBPs were abnormally altered in PAH and many could sensitively discriminate PAH from control. Finally, we show that as a protein functional class many of the IGFBPs are also associated with PAH severity.

CA2 is a member of a large family of 13 carbonic anhydrase (CA) isoenzymes, which catalyze reversible hydration of carbon dioxide. CA2 is a cytoplasmic isoform of CA and expressed in lung Type 1 and Type 2 alveolar epithelial cells and capillary endothelial cells.^{11,12} Acetazolamide, a sulfonamide diuretic, is a pharmacologic CA inhibitor and used clinically for the treatment of glaucoma and high-altitude pulmonary edema.¹³ In the lungs, acetazolamide lowers normoxic pulmonary artery pressure¹⁴ and inhibits hypoxic pulmonary vasoconstriction in animals¹⁵ and humans.^{16,17} When studied in isolated pulmonary artery smooth

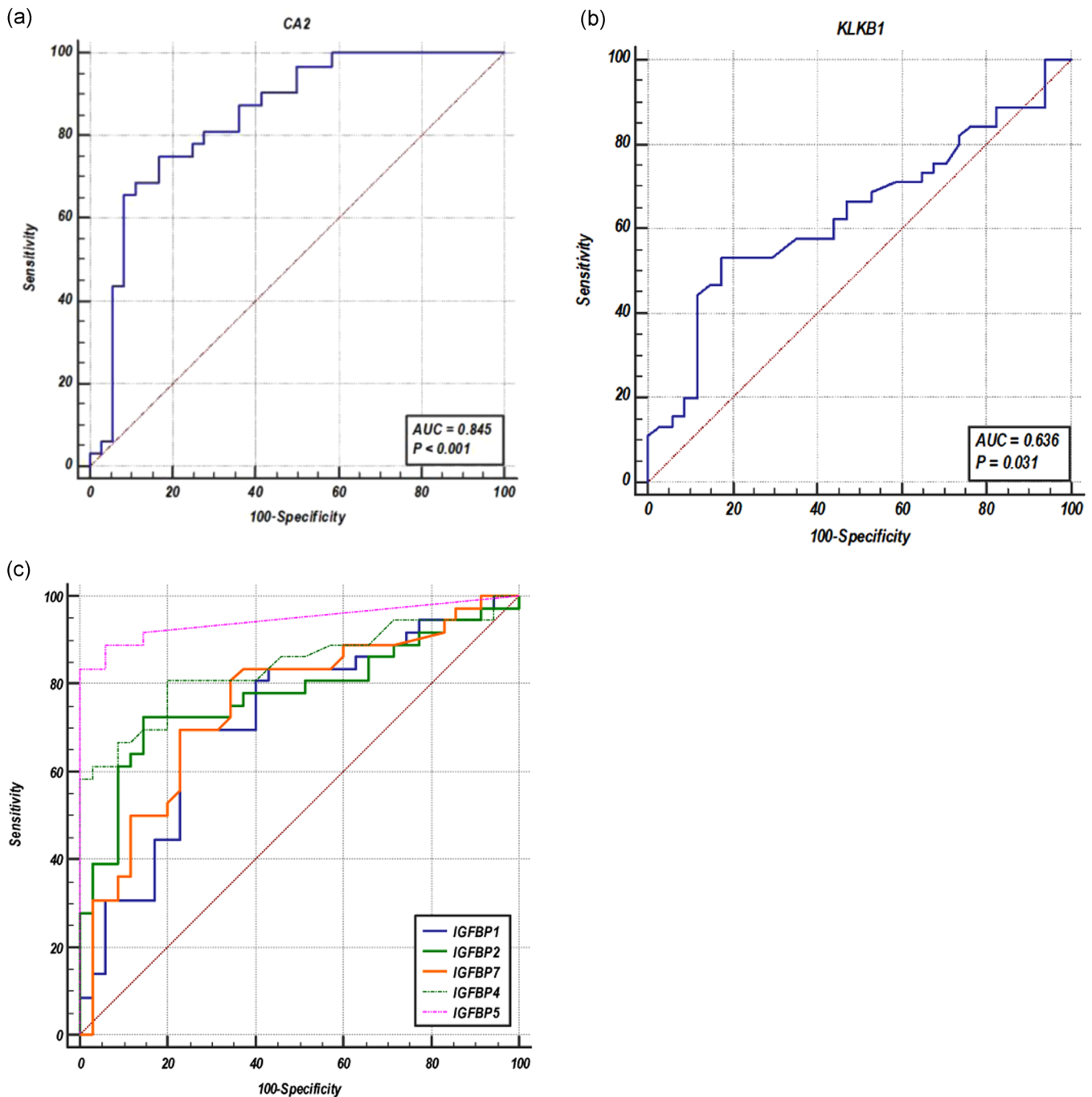


FIGURE 4 Receiver operating characteristic (ROC) curves for CA2 (a), KLKB1 (b), comparison of ROC (c) for IGFBP1, IGFBP2, IGFBP4, IGFBP5, and IGFBP7 as predictors of pulmonary arterial hypertension for 36 subjects from the PAH cohort and 35 healthy control subjects. AUC, area under the curve

muscle cells, acetazolamide (10–100 μ M) blocks the rise in intracellular cytosolic calcium that occurs with hypoxia, that would initiate membrane depolarization and smooth muscle contraction, thus inhibiting pulmonary artery smooth muscle contraction.¹⁷ In an animal model of lung ischemia/reperfusion injury, acetazolamide treatment attenuated ischemia/reperfusion induced lung injury, decreasing pulmonary vascular hyper-

permeability, pulmonary edema, pulmonary hypertension and neutrophilic sequestration.¹⁸

In this study, CA2 was one of the best performing biomarkers alone to discriminate PAH from healthy controls, although circulating CA2 levels were not associated with any clinical or hemodynamic measures. CA2 may have diagnostic potential as a pulmonary vascular marker but may not be involved in the downstream

TABLE 7 The association of biomarkers with functional and hemodynamics measures

	6MWD (m, N = 32)		CO (L/min, N = 36)		PVR (WU, N = 36)		mPAP (mmHg, N = 36)		mRAP (mmHg, N = 36)		PCWP (mmHg, N = 36)	
	R	p-value	R	p-value	R	p-value	R	p-value	R	p-value	R	p-value
CA2	-0.04	0.824	0.301	0.079	-0.126	0.472	-0.183	0.271	-0.011	0.948	-0.097	0.562
KLKB1	-0.037	0.828	-0.1	0.513	0.289	0.054	0.217	0.151	0.143	0.35	-0.154	0.314
IGFBP1	-0.332	0.0633	-0.384	0.0209	0.345	0.0395	0.161	0.348	0.169	0.325	0.2	0.2414
IGFBP2	-0.42	0.0182	-0.326	0.0522	0.36	0.031	0.289	0.087	0.143	0.406	0.23	0.1776
IGFBP3	0.088	0.6313	-0.163	0.3425	0.067	0.6959	-0.037	0.832	-0.141	0.411	-0.266	0.1175
IGFBP4	-0.63	0.0001	-0.127	0.4613	0.159	0.3544	0.237	0.164	0.288	0.089	0.376	0.0238
IGFBP5	0.069	0.7076	-0.101	0.556	0.159	0.3543	0.202	0.237	0.239	0.16	0.286	0.0907
IGFBP6	-0.171	0.3504	0.134	0.4367	-0.108	0.5313	0.079	0.648	-0.019	0.914	0.211	0.2156
IGFBP7	-0.45	0.0093	-0.185	0.2807	0.254	0.1344	0.272	0.108	0.241	0.157	0.134	0.4349

Note: Bold numbers indicate statistically significant differences.

Abbreviations: CO, cardiac output; mPAP, mean pulmonary artery pressure; mRAP, mean right atrial pressure; PVR, pulmonary vascular resistance; 6MWD, 6-min walk distance.

cardiopulmonary interactions that drive outcomes in this disease.

Similar to CA2, we confirmed that plasma kallikrein (KLKB1) was significantly increased in PAH compared to controls. KLKB1 is a member of the plasma kallikrein kinin system (plasma KKS),^{19–21} which has powerful vascular inflammation and permeability effects by activating bradykinin and des-Arg-bradykinin.²² Pharmacologic studies have shown that KLKB1 induces retinal vascular hyperpermeability in diabetes, whereas a small-molecular inhibitor of KLKB1 reversed the effect.²³ Given the known microvascular damage in PAH, the role of KLKB1 and other members of plasma kallikrein kinin system may be significant cardio-pulmonary regulators in PAH and require further investigation.

The insulin-like growth factor family is a group of essential proteins including IGF1, IGF2, the respective receptors, as well as IGF binding proteins. IGF1 promotes angiogenesis by stimulating endothelial cell proliferation and survival.²⁴ IGF1 also has a protective role in cardiovascular disease with maintaining cardiomyocyte and vascular smooth muscle cells function and survival;²⁵ indeed, knockout of IGF1 in mice proved lethal.²⁶ Typically, free IGF1 is thought of as a primary hormone in cell signaling, with the IGF binding proteins acting to bind and sequester IGF1, thus preventing its degradation.²⁷ We previously showed that IGFBP2 was a predictor of PAH disease severity and survival,⁴ but did not find a similar association for IGF1 or IGF2, which are typically thought to be the important effect molecule in vascular signaling. In this study, IGFBP2 was once again found to be significantly elevated in PAH plasma, however we also found that the other IGF binding proteins were significantly dysregulated in PAH with associations with severity and survival.

The IGF binding protein (IGFBP) family originally consisted of six binding proteins, all of which bind IGF1 and IGF2 with high affinity.^{26,27} The newest member, IGFBP7, also known as IGFBPRP1, or prostacyclin-stimulating factor, is similar but has a lower binding affinity with IGFs compared to other IGFBPs.^{28,29} The primary function of IGFBPs is thought to be a reservoir for IGF1 and IGF2, forming a protein complex with the IGFBPs. Thus extending their half-life in the circulation, as well as assisting in delivery of IGF proteins to cell surface receptors.^{24–27} Comparative genomic studies found that all IGFBPs evolved from a single ancestor through gene duplication; the function of the original IGFBPs in amphioxus chordate was independent of IGF; binding and sequestering IGF appears to be a trait that was acquired later during gene evolution.³⁰ Thus, the current family of human IGFBPs in appears to function as a reservoir for IGF1, but also may have IGF

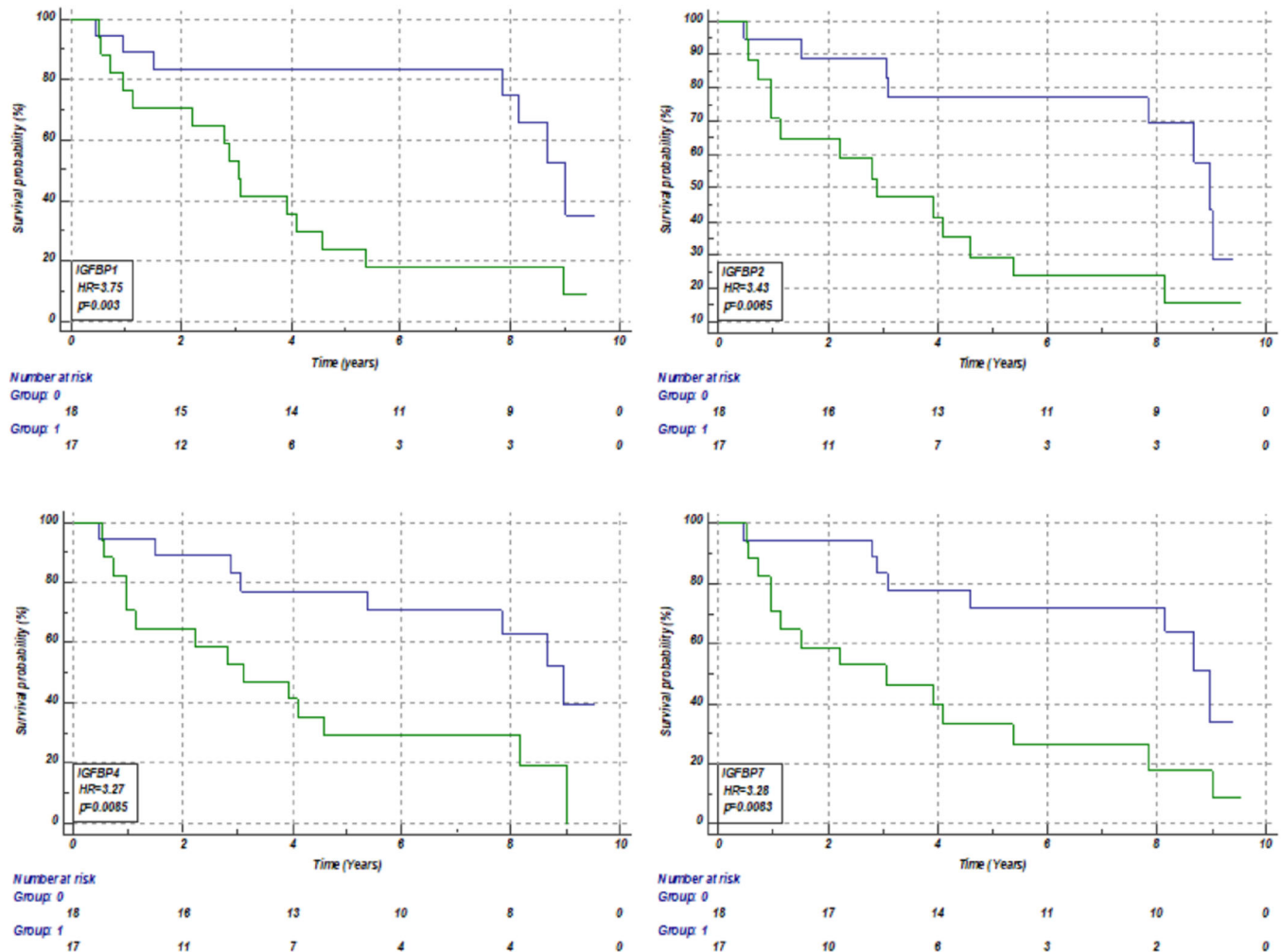


FIGURE 5 Kaplan–Meier survival curve for IGFBP1, IGFBP2, IGFBP4, and IGFBP7 in the pulmonary arterial hypertension (PAH) cohort. The curves represent survival analysis of PAH cohort dichotomized by median serum protein level of each individual IGFBP separately. HR, hazard ratio

TABLE 8 Multivariable Cox proportional hazard models for IGF binding proteins adjusted for age

	HR	95% CI	p-value
IGFBP1 (ng/ml)	2.31	1.38, 3.87	0.002*
IGFBP2 (ng/ml)	2.12	1.16, 3.88	0.015*
IGFBP3 (ng/ml)	0.79	0.19, 3.33	0.8
IGFBP4 (ng/ml)	4.9	1.87, 12.9	0.001*
IGFBP5 (ng/ml)	1.04	0.85, 1.28	0.7
IGFBP6 (ng/ml)	0.71	0.30, 1.66	0.4
IGFBP7 (ng/ml)	2.22	1.01, 4.84	0.046*

Abbreviations: CI, confidence interval; HR, hazard ratio.

* $p < 0.05$.

independent functions.^{31,32} In this study, multiple IGFbps had altered levels in PAH and were associated with survival as a measure of PAH severity. Using a complementary targeted aptamer array based approach,

Rhodes et al.³³ identified that IGFBP1 increased in PAH and associated with survival. Presently, it is unknown whether these IGF binding proteins are regulating IGF1, with ultimate effect on the vascular smooth muscle and cardiomyocytes, or whether they have an independent role in PAH pathobiology. Given the complexity of the IGF axis, an integrated approach to analyze all IGF binding proteins together in PAH will be needed.

The study is limited in that it is a cross-sectional discovery study, using samples from patients with end-stage disease before transplant and obtained at a single time point. Thus, the markers identified may be markers of more severe disease rather than early changes. Further, to eliminate the analysis bias, the verification cohort consisted of patients with less severe disease than the discovery cohort, namely longer 6MWD, better hemodynamics measures and milder NYHA classification groups. Since this is a cross-sectional study, we are unable to show how these markers changed over time, which may reveal a proteomic marker or pattern

of changes signaling disease progression. The verification cohort in this study was also relatively small, although significant clinical associations were still found. The control cohort did not perfectly match the disease cohort for age, however, age was adjusted in the followed analysis. Future studies should focus on validation of these markers in larger cohorts, including generating and testing cutoff values to distinguish PAH from controls, and a systematic approach to understanding the function of these protein networks together in PAH and other PH types.

ACKNOWLEDGMENTS

We thank contributors, including the Pulmonary Hypertension Centers who collected samples used in this study, as well as patients and their families, whose help and participation made this study possible. This study was supported by the National Institutes of Health/National Heart, Lung, and Blood Institute award 1R03HL110830 (Allen D. Everett, Melanie K. Nies, and Jun Yang), R01HL135114, and R01HL150070 (Allen D. Everett, Jun Yang, Rachel Damico, Dhananjay Vaydia, Dunbar D. Ivy, and Eric D. Austin). The Johns Hopkins Innovation Proteomics Center in Heart Failure (NHLBI-HV-10-05 (2) contract, Jennifer E. Van Eyk) supported the development of the proteomics discovery platform. Melanie K. Nies was also supported by the Matthew and Michael Wojciechowski Pulmonary Hypertension Pediatric Proof-of-Concept Grant (Dr. Robyn J. Barst Pediatric PH Research and Mentoring Fund Grant). Megan Griffiths was supported by the Pediatric Scientist Development Program. The Pediatric Scientist Development Program is supported by Award Number K12-HD000850 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development. Serum/tissue samples were provided by PHBI under the PHBI. Funding for the PHBI is provided under an NHLBI R24 grant, R24HL123767, and by the Cardiovascular Medical Research and Education Fund (CMREF).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

ETHICS STATEMENT

All study cohorts were approved by Johns Hopkins University Institutional Review Board, and informed consent was obtained from all subjects by the PHBI (<https://www.ipahresearch.org>).

AUTHOR CONTRIBUTIONS

Jun Yang, Melanie K. Nies, Jennifer E. Van Eyk, and Allen D. Everett planned the project, analyzed the data, and wrote the manuscript; Jun Yang, Melanie K.

Nies, Jie Zhu, Zongming Fu, Stephanie Brandal, and Megan Griffiths performed the experiments and interpreted the results; Dhananjay Vaydia, Rachel Damico, Jun Yang, and Megan Griffiths performed statistical analysis; Melanie K. Nies, Rachel Damico, Dunbar D. Ivy, Eric D. Austin, Paul M. Hassoun, and Allen D. Everett recruited subjects and performed research; all authors reviewed, revised, and approved the manuscript for submission.

ORCID

Jun Yang  <http://orcid.org/0000-0001-9273-4578>

REFERENCES

1. Simonneau G, Montani D, Celermajer DS, Denton CP, Gatzoulis MA, Krowka M, Williams PG, Souza R. Haemodynamic definitions and updated clinical classification of pulmonary hypertension. *Eur Respir J*. 2019;53(1):1801913.
2. Takase H, Dohi Y. Kidney function crucially affects B-type natriuretic peptide (BNP), N-terminal proBNP and their relationship. *Eur J Clin Invest*. 2014;44(3):303–8.
3. Lammers AE, Hislop AA, Haworth SG. Prognostic value of B-type natriuretic peptide in children with pulmonary hypertension. *Int J Cardiol*. 2009;135(1):21–6.
4. Yang J, Griffiths M, Nies MK, Brandal S, Damico R, Vaidya D, Tao X, Simpson CE, Kolb TM, Mathai SC, Pauciulo MW, Nichols WC, Ivy DD, Austin ED, Hassoun PM, Everett AD. Insulin-like growth factor binding protein-2: a new circulating indicator of pulmonary arterial hypertension severity and survival. *BMC Med*. 2020;18(1):268.
5. Simpson RJ, Greening DW, Sheng S, Skalninkova H, Meng A, Tra J, Fu Q, Everett A, Van Eyk J. Intact protein separation by one- and two-dimensional liquid chromatography for the comparative proteomic separation of partitioned serum or plasma. *Methods Mol Biol*. 2011;728:29–46.
6. Sheng S, Chen D, Van Eyk JE. Multidimensional liquid chromatography separation of intact proteins by chromatographic focusing and reversed phase of the human serum proteome: optimization and protein database. *Mol Cell Proteomics*. 2006;5(1):26–34.
7. Chong J, Soufan O, Li C, Caraus I, Li S, Bourque G, Wishart DS, Xia J. MetaboAnalyst 4.0: towards more transparent and integrative metabolomics analysis. *Nucleic Acids Res*. 2018;46(W1):W486–94.
8. Chong J, Yamamoto M, Xia J. MetaboAnalystR 2.0: from raw spectra to biological insights. *Metabolites*. 2019;9:3.
9. Chong J, Xia J. MetaboAnalystR: an R package for flexible and reproducible analysis of metabolomics data. *Bioinformatics*. 2018;34(24):4313–14.
10. Xia J, Wishart DS. Using MetaboAnalyst 3.0 for comprehensive metabolomics data analysis. *Curr Protoc Bioinformatics*. 2016;55:14.10.1–14.10.91.
11. Chen J, Lecuona E, Briva A, Welch LC, Sznajder JJ. Carbonic anhydrase II and alveolar fluid reabsorption during hypercapnia. *Am J Respir Cell Mol Biol*. 2008;38(1):32–7.
12. Esbaugh AJ, Tufts BL. The structure and function of carbonic anhydrase isozymes in the respiratory system of vertebrates. *Respir Physiol Neurobiol*. 2006;154(1–2):185–98.

13. Luks AM, Swenson ER, Bärtsch P. Acute high-altitude sickness. *Eur Respir Rev.* 2017;26(143):160096.
14. Teppema LJ, Balanos GM, Steinback CD, Brown AD, Foster GE, Duff HJ, Leigh R, Poulin MJ. Effects of acetazolamide on ventilatory, cerebrovascular, and pulmonary vascular responses to hypoxia. *Am J Respir Crit Care Med.* 2007;175(3):277–81.
15. Höhne C, Krebs MO, Seiferheld M, Boemke W, Kaczmarczyk G, Swenson ER. Acetazolamide prevents hypoxic pulmonary vasoconstriction in conscious dogs. *J Appl Physiol* (1985). 2004;97(2):515–21.
16. Ke T, Wang J, Swenson ER, Zhang X, Hu Y, Chen Y, Liu M, Zhang W, Zhao F, Shen X, Yang Q, Chen J, Luo W. Effect of acetazolamide and ginkgo biloba on the human pulmonary vascular response to an acute altitude ascent. *High Alt Med Biol.* 2013;14(2):162–7.
17. Shimoda LA, Luke T, Sylvester JT, Shih HW, Jain A, Swenson ER. Inhibition of hypoxia-induced calcium responses in pulmonary arterial smooth muscle by acetazolamide is independent of carbonic anhydrase inhibition. *Am J Physiol Lung Cell Mol Physiol.* 2007;292(4):L1002–L12.
18. Lan CC, Peng CK, Tang SE, Huang KL, Wu CP, Palaniyar N. Carbonic anhydrase inhibitor attenuates ischemia-reperfusion induced acute lung injury. *PLoS One.* 2017;12(6):e0179822.
19. Bryant JW, Shariat-Madar Z. Human plasma kallikrein-kinin system: physiological and biochemical parameters. *Cardiovasc Hematol Agents Med Chem.* 2009;7(3):234–50.
20. Rhaleb NE, Yang XP, Carretero OA. The kallikrein-kinin system as a regulator of cardiovascular and renal function. *Compr Physiol.* 2011;1(2):971–93.
21. Feener EP, Zhou Q, Fickweiler W. Role of plasma kallikrein in diabetes and metabolism. *Thromb Haemost.* 2013;110(3):434–41.
22. Marceau F, Regoli D. Bradykinin receptor ligands: therapeutic perspectives. *Nat Rev Drug Discov.* 2004;3:845–52.
23. Clermont A, Murugesan N, Zhou Q, Kita T, Robson PA, Rushbrooke LJ, Evans DM, Aiello LP, Feener EP. Plasma kallikrein mediates vascular endothelial growth factor-induced retinal dysfunction and thickening. *Invest Ophthalmol Vis Sci.* 2016;57(6):2390–99.
24. Bach LA. Endothelial cells and the IGF system. *J Mol Endocrinol.* 2015;54(1):R1–R13.
25. Troncoso R, Ibarra C, Vicencio JM, Jaimovich E, Lavandero S. New insights into IGF-1 signaling in the heart. *Trends Endocrinol Metab.* 2014;25(3):128–37.
26. Jones JI, Clemmons DR. Insulin-like growth factors and their binding proteins: biological actions. *Endocr Rev.* 1995;16(1):3–34.
27. Collett-Solberg PF, Cohen P. Genetics, chemistry, and function of the IGF/IGFBP system. *Endocrine.* 2000;12(2):121–36.
28. Jin L, Shen F, Weinfeld M, Sergi C. Insulin growth factor binding protein 7 (IGFBP7)-related cancer and IGFBP3 and IGFBP7 crosstalk. *Front Oncol.* 2020;10:727.
29. Zhou J, Xiang J, Zhang S, Duan C. Structural and functional analysis of the amphioxus IGFBP gene uncovers ancient origin of IGF-independent functions. *Endocrinology.* 2013;154(10):3753–63.
30. Burger AM, Leyland-Jones B, Banerjee K, Spyropoulos DD, Seth AK. Essential roles of IGFBP-3 and IGFBP-rP1 in breast cancer. *Eur J Cancer.* 2005;41(11):1515–27.
31. LeRoith D, Holly JMP, Forbes BE. Insulin-like growth factors: ligands, binding proteins, and receptors. *Mol Metab.* 2021;4:101245.
32. Allard JB, Duan C. IGF-binding proteins: why do they exist and why are there so many? *Front Endocrinol (Lausanne).* 2018;9:117.
33. Rhodes CJ, Wharton J, Ghataorhe P, Watson G, Girerd B, Howard LS, Gibbs JSR, Condliffe R, Elliot CA, Kiely DG, Simonneau G, Montani D, Sitbon O, Gall H, Schermuly RT, Ghofrani HA, Lawrie A, Humbert M, Wilkins MR. Plasma proteome analysis in patients with pulmonary arterial hypertension: an observational cohort study. *Lancet Respir Med.* 2017;5:717–26.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Nies MK, Yang J, Griffiths M, Damico R, Zhu J, Vaydia D, Fu Z, Brandal S, Austin ED, Ivy DD, Hassoun PM, Van Eyk JE, Everett AD. Proteomics discovery of pulmonary hypertension biomarkers: insulin-like growth factor binding proteins are associated with disease severity. *Pulm Circ.* 2022;12:e12039. <https://doi.org/10.1002/pul2.12039>