Mouse Genome-Wide Association Study of Preclinical Group II Pulmonary Hypertension Identifies Epidermal Growth Factor Receptor

Neil J. Kelly^{1*}, Josiah E. Radder^{1*}, Jeffrey J. Baust², Christine L. Burton¹, Yen-Chun Lai^{1,2}, Karin C. Potoka^{1,3}, Brittani A. Agostini¹, John P. Wood¹, Timothy N. Bachman², Rebecca R. Vanderpool², Nadine Dandachi¹, Adriana S. Leme¹, Alyssa D. Gregory¹, Alison Morris¹, Ana L. Mora^{1,2}, Mark T. Gladwin^{1,2}, and Steven D. Shapiro¹

¹Department of Medicine, ²Vascular Medicine Institute, and ³Department of Pediatrics, University of Pittsburgh and University of Pittsburgh Medical Center, Pittsburgh, Pennsylvania

Abstract

Pulmonary hypertension (PH) is associated with features of obesity and metabolic syndrome that translate to the induction of PH by chronic high-fat diet (HFD) in some inbred mouse strains. We conducted a genome-wide association study (GWAS) to identify candidate genes associated with susceptibility to HFD-induced PH. Mice from 36 inbred and wild-derived strains were fed with regular diet or HFD for 20 weeks beginning at 6–12 weeks of age, after which right ventricular (RV) and left ventricular (LV) end-systolic pressure (ESP) and maximum pressure (MaxP) were measured by cardiac catheterization. We tested for association of RV MaxP and RV ESP and identified genomic regions enriched with nominal associations to both of these phenotypes. We excluded genomic regions if they were also associated with LV MaxP, LV ESP, or body weight. Genes within significant regions were scored based on the shortest-path betweenness centrality, a measure of network connectivity, of their human orthologs in a gene interaction network of human PH-related genes. WSB/Eij, NON/ShiLtJ, and AKR/J mice had the largest increases in RV MaxP after high-fat feeding. Network-based scoring of GWAS candidates identified epidermal growth factor receptor

(*Egfr*) as having the highest shortest-path betweenness centrality of GWAS candidates. Expression studies of lung homogenate showed that EGFR expression is increased in the AKR/J strain, which developed a significant increase in RV MaxP after high-fat feeding as compared with C57BL/6J, which did not. Our combined GWAS and network-based approach adds evidence for a role for *Egfr* in murine PH.

Keywords: pulmonary hypertension; metabolic syndrome; genome-wide association study

Clinical Relevance

The results of our study will be broadly useful to investigators in the field, both in facilitating greater understanding of the pathophysiology of pulmonary hypertension (PH) and in identifying novel targets for therapeutic intervention. In addition, the results of our strain study will be widely applicable to investigators in PH, as the choice of background strain is crucial to developing a robust disease model.

Knowledge of the molecular events contributing to the pathogenesis of pulmonary hypertension (PH) has increased substantially in recent decades; however, our understanding remains incomplete, as

evidenced by the scarcity of medical interventions to prolong the survival of patients suffering from this deadly disease (1). PH occurs in a variety of settings, including left-sided heart disease, chronic lung

disease, intrinsic disease of the pulmonary arteries (pulmonary arterial hypertension [PAH]), and others (2). Despite its distinct triggers, PH subtypes share many features in common (2, 3), including pathogenic

(Received in original form June 7, 2016; accepted in final form December 30, 2016)

*These authors contributed equally to this work.

This work was supported by National Institutes of Health grants 5P01HL103455 (A.M., A.L.M., M.T.G., and S.D.S.), 5T32HL094295 (N.J.K.), and 5T32GM008208 (N.J.K.), and the Flight Attendant Medical Research Institute (A.D.G.).

Author Contributions: Conception and design—N.J.K., J.E.R., A.M., A.L.M., M.T.G., S.D.S.; data collection—N.J.K., J.J.B., C.L.B., B.A.A., J.P.W., A.S.L., A.D.G.; Analysis and interpretation: N.J.K., J.E.R., Y.-C.L., K.C.P., T.N.B., R.R.V., N.D.; drafting the manuscript for important intellectual content—all authors.

Correspondence and requests for reprints should be addressed to Steven D. Shapiro, M.D., 600 Grant Street, Suite 6250, Pittsburgh, PA 15219. E-mail: shapirostd@upmc.edu

This article has an online supplement, which is accessible from this issue's table of contents at www.atsjournals.org

Am J Respir Cell Mol Biol Vol 56, Iss 4, pp 488–496, Apr 2017

Copyright © 2017 by the American Thoracic Society

Originally Published in Press as DOI: 10.1165/rcmb.2016-0176OC on January 13, 2017

Internet address: www.atsjournals.org

vascular remodeling and progressive hypertrophy, dilation, and failure of the right ventricle.

Discoveries from human genetics- and genomics-based approaches form the core of our current understanding of PH as a disease of excessive proliferation and impaired apoptosis in the pulmonary vasculature (4). Early studies investigating the segregation of PAH cases in families (FPAH) identified causal roles for mutations in *BMPR2* (5, 6) and other genes linked to proliferative and antiapoptotic signaling by the transforming growth factor- β superfamily (7), including *ALK1* (8), *SMAD9* (9), and *ENG* (8). More recently, the caveolar protein, *CAV1* (10), pH-sensitive potassium channel, *KCNK3* (11), and eukaryotic translation initiation factor, *EIF2AK4* (12, 13), have been linked to FPAH. Although these genes are not known to be associated with transforming growth factor- β signaling, they are also believed to contribute to excessive proliferation and apoptosis resistance (4). Although currently known genetic abnormalities explain greater than 80% of FPAH cases (7), the Mendelian inheritance patterns observed in FPAH are confounded by low penetrance (14), and account for a small minority of total PH cases (15), underscoring the complex and multifactorial nature of PH pathogenesis (16). Genome-wide association study (GWAS) offers the potential to account for this complexity by recognizing loci associated with PH susceptibility, and the recent success of GWAS in identifying *CBLN2* as a candidate PAH disease gene (17) suggests that it may be useful in identifying PH disease genes more broadly.

In recent years, a new metabolic hypothesis has emerged to unite the multitude of abnormalities involved in the pathogenesis of PH (18). This theory has gained traction from observations in

rodent models that (1) metabolic syndrome induces experimental PH (19, 20), and (2) drugs that treat metabolic abnormalities also reverse experimental PH (21, 22). Indeed, the antidiabetic drug, metformin, is therapeutic in diverse rodent models of PH, including the monocrotaline-induced model of PAH, the chronic hypoxia-induced model of PH secondary to hypoxic lung disease, and experimental PH secondary to diastolic left heart failure (21, 23).

In this study, we sought to leverage the power of GWAS in mice to identify novel candidate genes in PH pathogenesis as it relates to high-fat feeding. Murine GWAS have been successfully employed to identify candidate genes associated with other pulmonary disorders, including asthma (24) and lung cancer (25). To induce disease, we used a high-fat diet (HFD)-induced mouse model, which has been shown to produce a robust PH phenotype in mice (20). We hypothesized that testing for association with this phenotype would identify genes that are involved in PH pathogenesis. Here, we report the results of a genome-wide scan testing for association with HFD-induced changes in right ventricular (RV) maximum pressure (MaxP) in 36 inbred and wild-derived mouse strains. Some of the results of these studies have been previously reported in the form of abstracts (26–28).

Materials and Methods

Animals

Male mice from 36 inbred and wild-derived strains (Table 1) were purchased from Jackson Laboratories (Bar Harbor, ME). All animal experiments were performed in accordance with the Institutional Animal Care and Use Committee of the University of Pittsburgh School of Medicine

(Pittsburgh, PA; protocol no. 13,112,793). Animals were housed within a pathogen-free barrier facility that maintained a 12-hour light/dark cycle in Plexiglas cages (one to four mice per cage) with free access to autoclaved water and irradiated pellet food. Animal health, weight, and overall behavior were monitored throughout the experiments.

High-Fat Feeding

Male mice were fed an open regular diet (RD; 15% lipids/kcals) or HFD (60% lipids/kcals; Research Diets, New Brunswick, NJ) for 20 weeks beginning at 6–12 weeks of age.

Hemodynamics

In vivo pressure–volume (PV) loop measurements of RV function were performed by a PV catheter in anesthetized animals, as described previously (20). Briefly, animals were weighed and then anesthetized with isoflurane (5% for induction, 2% during surgery, and 1% while performing PV loop measurements). A four-electrode PV catheter (Scisense, Inc., London, ON, Canada) attached to the data acquisition system (EMKA Instruments, Falls Church, VA) was inserted into the apex of the RV. After the acquisition of RV PV loops, the procedure was repeated on the left ventricle (LV). Data were acquired using the PowerLab data acquisition system and LabChart Pro software (AD Instruments, Colorado Springs, CO). PV loop quality control was performed by a blinded observer.

Genetic Association Testing

Genome-wide scans were performed using the genome-wide efficient mixed model algorithm (29) using 4 million single-nucleotide polymorphisms (SNPs) made publicly available by the National Institute of Environmental Health Sciences and imputed from 138,980 SNPs published in the Mouse HapMap project (available at <http://mouse.cs.ucla.edu/mousehapmap/full.html>).

Region Analysis

Regions were constructed by beginning with a nominally significant SNP ($P < 10^{-3}$) and iteratively adding flanking nominally significant SNPs within ± 1 megabase (Mb) until there were no nominally significant SNPs in the 1 Mb upstream or downstream of the region. Region *P* values were calculated by the hypergeometric survival

Table 1. List of Mouse Strains Used in the Strain Study

129S1/SvImJ	C57BL/10J	DBA/2J	NZB/BinJ
A/J	C57BL/6J	FVB/NJ	NZO/HiLtJ
AKR/J	C57BLKS/J	KK/HIJ	NZW/LacJ
BALB/cByJ	C57BR/cdJ	LG/J	PWK/PhJ
BALB/cJ	C57L/J	I/LnJ	PL/J
BPN/3J	CAST/EiJ	LP/J	RIIS/J
BTBRT+tf/J	CBA/J	MRL/MpJ	SJL/J
BUB/BnJ	CE/J	NOD/ShiLtJ	SWR/J
C3H/HeJ	DBA/1J	NON/ShiLtJ	WSB/EiJ

function based on the total number of SNPs, the number of SNPs in the region, the total number of significant SNPs ($P < 10^{-3}$), and the number of significant SNPs in the region. Regions with hypergeometric P less than 2.5×10^{-3} were considered significant. Mouse genes with exons contained wholly or partially within a significant region were considered candidate genes.

Statistical Analysis

P values for fold change between dietary groups were calculated by Student's two-sided t test. P values for fold change in gene expression between dietary groups and strains were calculated by two-way ANOVA. Spearman correlation coefficients and two-tailed P values were calculated in PRISM v6.04 (GraphPad Software, Inc., La Jolla, CA). P values less than 0.05 were considered statistically significant.

Overlapping Publication

The fold change in RV MaxP with high-fat feeding by background strain (Figure 1A) is also reported in an article by Meng and colleagues (30) to identify an ideal strain for the development of a murine model of heart failure with preserved ejection fraction.

Additional details on these methods are provided in the online supplement.

Results

Interstrain Differences in HFD-Induced PH Susceptibility

To quantify the effects of chronic high-fat feeding on pulmonary hemodynamics, we fed mice from 36 inbred or wild-derived strains (Table 1) with a RD (15% lipids/kcal) or HFD (60% lipids/kcal) beginning at 6–12 weeks of age. To eliminate the confounding effects of sex (31), only male mice were used in this study. After 20 weeks on their respective diets, RV and left-ventricular (LV) PV loops were obtained by terminal cardiac catheterization, allowing determination of the fold change in RV MaxP, RV end-systolic pressure (ESP), RV end-diastolic pressure, LV MaxP, LV ESP, and LV end-diastolic pressure after high-fat feeding (Figures 1A–1D, and see Figure E1 in the online supplement). Fold change in body weight was also recorded (Figure 1E), and RV hypertrophy was assessed by

measuring the fold change in its free weight (Figure E2A) or that of its weight expressed as a fraction of either left tibial length (Figure E2B) or the mass of the LV plus the interventricular septum (Fulton index, Figure E2C). Similarly, LV hypertrophy was assessed by measuring the fold change in its free weight, including the interventricular septum (Figure E3A), or their weights expressed as a fraction of left tibial length (Figure E3B). Raw values for each mouse in the study are available in the online supplement (Table E1).

In this HFD-induced model of PH, we assessed the degree of PH as the fold change in RV MaxP in HFD versus RD. This metric ranged from a 30.59% decline in the NOD/ShiLtJ strain ($P = 0.3049$) to increases of 30.40 and 35.74% in the common laboratory strains, AKR/J ($P = 0.0040$) and NON/ShiLtJ ($P < 0.0001$), respectively. The wild-derived WSB/Eij strain (32) showed the greatest increase in RV MaxP (69.67%), although the change did not meet criteria for significance between dietary groups ($P = 0.0529$). To assess the relationship between right- and left-sided pressures and weight gain, we measured the correlation between fold change in RV MaxP and fold changes in RV ESP, LV MaxP, LV ESP, and body weight. As expected, fold change in RV MaxP was significantly positively correlated with that of RV ESP (Figure E4A, Spearman's $r = 0.9423$, $P < 0.0001$). In addition, the fold change in both LV MaxP (Figure E4B, $r = 0.4208$, $P = 0.0106$) and LV ESP (Figure E4C, $r = 0.3490$, $P = 0.0369$) showed positive and significant association with the fold change in RV MaxP, suggesting that the two phenotypes may be related as seen in PH secondary to left-sided heart disease; this finding is explored in detail in an accompanying article by Meng and colleagues (30). Interestingly, although our model of PH relied on intervention through HFD, fold change in body weight was not significantly correlated with fold change in RV MaxP (Figure E4D, $r = 0.1320$, $P = 0.4427$), suggesting that factors besides susceptibility to HFD-induced weight gain determine the extent of PH.

Genome-Wide Association Identifies Regions Associated with Right-Ventricular Hemodynamics

After data acquisition, we identified SNPs associated with each of five phenotypes—encompassing RV MaxP, RV ESP, LV

MaxP, LV ESP, and body weight (29, 33)—and all SNPs with P less than 10^{-3} were considered potentially significant (Tables E2–E6). To account for the influence of RD phenotypes on HFD phenotypes, we expressed each quantitative phenotype with HFD as a fold change from the mean RD value of its strain. Due to the significant linkage disequilibrium that exists between inbred strains, we constructed genomic regions of interest for each phenotype by combining 1-Mb flanking segments of significant SNPs ($P < 10^{-3}$). Significant regions were defined as regions that were enriched for significant SNPs by hypergeometric enrichment with P less than 2.5×10^{-3} (see MATERIALS AND METHODS). All genes within these regions were chosen as potential candidate genes.

To filter our initial list of 1,511 potential candidate genes located in regions associated with RV MaxP, we excluded genes that were also located within genomic regions associated with either body weight, LV MaxP, or LV ESP, reasoning that these genes could function upstream of the molecular events of most interest to our study. We also excluded genes that were not present in genomic regions associated with RV ESP, as genes associated with both phenotypes were less likely to be false positives. Following this protocol, we obtained filtered genomic regions (Figure 2) consisting of 880 candidate genes.

Network Analysis Identifies Candidate Genes Associated with Human PH

We next sought to identify genes that are closely related to what is currently known about PH. Protein interaction networks have been successfully used to gain novel insights into PH pathogenesis (22, 34). Thus, we examined the PH relatedness of the human orthologs (35) of our candidate genes based on their connectedness to a subnetwork of known PH genes, which we refer to as the PH interactome (34), within a curated global network of protein–protein interactions (36) (Figure 3A). Of our remaining genes, 31 were network-connected genes, which we defined as genes that were either members or first-degree neighbors of the PH interactome. We hypothesized that, if our network-based filtering strategy is valid, then the SNPs with $-\log_{10}(P \text{ value})$ values for association with RV MaxP above a

Figure 1. Fold change right ventricular (RV) and left ventricular (LV) systolic pressures and body weight. A total of 36 strains (*horizontal axis*) of inbred or wild-derived mice were fed with a regular diet (RD) or high-fat diet (HFD) for 20 weeks beginning at 6–12 weeks of age ($n = 3–8$ mice/diet per strain). RV and LV hemodynamics were measured by terminal catheterization. Fold change in (A) RV maximum pressure (MaxP), (B) RV end-systolic pressure (ESP), (C) LV MaxP, (D) LV ESP, and (E) body weight were calculated as the HFD value divided by the average RD value. Data are mean ($\pm$ SEM).

Figure 2. Region-based filtering of genome-wide association study (GWAS) results. Genome-wide single-nucleotide polymorphism (SNP) association P values were calculated by genome-wide efficient mixed model for fold change (HFD versus RD) in RV MaxP, RV ESP, LV MaxP, LV ESP, and body weight (Weight). Black and gray dots represent SNP P values (left axis) for the respective phenotypes (right axis) at their chromosomal position (horizontal axis). Colored strips represent chromosomes, and black bars show regions enriched (hypergeometric $P < 2.5 \times 10^{-3}$) for significant SNPs ($P < 10^{-3}$). Regions were filtered (Filt.) as chromosomal regions that were significant in RV MaxP and RV ESP phenotypes, but not Weight, LV MaxP, nor LV ESP phenotypes.

sliding threshold should be progressively enriched for SNPs in network-connected genes versus SNPs in other unfiltered genes as the threshold increases from zero to the maximum value in the GWAS (Figure 3B). We saw that the hypergeometric P value of SNPs in network-connected genes became increasingly significant, to P values of less than 10^{-10} , as the threshold approached the maximum (Figure 3C). These data suggest that, by the above standard, the network-based filter is indeed valid.

We next scored the PH relatedness of our network-connected candidate genes based on their shortest path betweenness centrality (37) in a subnetwork consisting of the candidate and the PH interactome (22). Shortest path betweenness centrality of a node, v , measures the fraction of shortest

paths between two other network nodes, s and t , that pass through v ; in essence, shortest path betweenness centrality measures the amount of network connectivity that would be lost if v were to be removed from the network. We normalized each value to its theoretical maximum for an undirected graph (Table E7). We found that epidermal growth factor receptor (*Egfr*), a gene with a hypothesized association to PH pathogenesis (38), was the candidate gene with the highest connectivity to the PH network as assessed by betweenness centrality.

Throughout the filtering process, beginning with the initial list of 1,511 candidate genes in genomic regions developed from the GWAS results of fold change in RV MaxP, the filtering of

candidates was independent of the strengths of their associations to the original phenotype; this strategy ultimately led to the identification of *Egfr* as the highest-ranked gene (Figure 4A). Interestingly, 13 of the 19 most significant SNPs in our original association study with the phenotype of fold change in RV MaxP (see Table E2) were located within the genomic coordinates of the mouse *Egfr* gene (Figure 4B). To further test the possibility that genetic differences in *Egfr* contribute to PH susceptibility, we examined *Egfr* expression in the lungs of two common mouse strains with different hemodynamic responses to HFD: C57BL/6J, which developed a statistically insignificant 9.37% increase in RV MaxP with HFD compared with RD ($n = 6-7$ mice/diet, $P = 0.4299$), and AKR/J, which developed a 30.40%

Figure 3. Network-based filtering of candidate genes. (A) Graph showing network-connected candidate genes (*gray*) and their interactions with each other and with other known pulmonary hypertension (PH) genes (*white*). (B) Manhattan plot of SNPs in candidate genes filtered by phenotype screening and mapped to human orthologs. *P* values correspond to each SNP's strength of association with the phenotype in RV MaxP. (C) Plot of $-\log_{10}(P \text{ value})$ for the hypergeometric enrichment of SNPs in network-connected genes versus the threshold $-\log_{10}(\text{SNP } P \text{ value})$ above which their enrichment was calculated. *Dashed gray line with arrows* (B and C) shows the visualization of the threshold for $-\log_{10}(\text{SNP } P \text{ value})$.

increase between the two groups ($n = 7$ mice/diet, $P = 0.0040$). The expression of *Egfr* mRNA (Figure 4C) and protein (Figure 4D) was significantly increased in the lungs of AKR/J mice as compared with C57BL/6J mice, suggesting that increased *Egfr* expression may contribute to PH susceptibility in our model. Taken together, these results offer new genomic, computational, and biologic evidence of a role for *Egfr* in the pathogenesis of PH.

Discussion

In this report, we present the results of a murine GWAS of RV pressure phenotypes after chronic high-fat feeding. Through region-, phenotype-, and network-based filtering steps with biologic evidence

from gene expression studies, we identified *Egfr* as a candidate gene that may be related to murine PH.

In this study, we used network-based analyses to identify *Egfr* as the candidate gene most likely to be associated with changes in RV MaxP after high-fat feeding. Network-based analyses have recently been employed to gain biologic insights into the pathogenesis of PH (22, 34), and the emerging application of network methods to GWAS interrogation has shown significant promise (39–41). In a key study, Barrenäs and colleagues (42) showed that highly interconnected genes in multiple disease networks were enriched for disease-associated SNPs. Comparably, in our study, *Egfr* was found to be the most integral GWAS candidate gene in the PH network by shortest path betweenness centrality, while also harboring 13 of the 19 SNPs

most associated with the phenotype of fold change in RV MaxP. Interestingly, although Barrenäs and colleagues combined human networks with human genomics, our article is the first, to our knowledge, that pairs a human network with mouse genomics; nonetheless, we still reached compatible conclusions. Hence, the marriage of human network-based analytics with murine GWAS may offer the opportunity to elicit novel insights from both existing and future murine phenotyping surveys.

EGFR has long been known to regulate the growth and proliferation of vascular smooth muscle cells, making it an attractive drug target in PH (43, 44). However, pharmacologic inhibitors of EGFR have had mixed results in animal models of PH, and the potential for their use remains controversial (38). Although EGFR

Figure 4. GWAS filtering strategy identifies epidermal growth factor receptor (*Egfr*) as highest-ranked candidate gene. (A) Workflow summary showing filtering steps leading to identification of *Egfr* (BW, body weight; 1°, first degree). (B) SNPs located within the genomic coordinates of the mouse *Egfr* gene and their strengths of association with the phenotype of fold change in RV MaxP. C57BL/6J and AKR/J mice were fed RD or HFD diet for 20 weeks beginning at 6–12 weeks of age. (C) Expression of *Egfr* mRNA relative to that of a *Gapdh* endogenous control was measured by quantitative real-time PCR and expressed proportionally to its level in C57BL/6J mice fed an RD ($n = 4$ mice/diet per strain). P values were calculated by two-way ANOVA. (D) Levels of EGFR and GAPDH proteins were assessed by Western blot. Western blot is representative of two independent experiments ($n = 4$ mice/diet per strain).

inhibitors had therapeutic efficacy in monocrotaline-induced PH in rats (44), they had no effect on the severity of PH in a mouse model of chronic hypoxia-induced PH (38). Interestingly, those mouse studies were conducted in C57BL/6J mice, and studies in AKR/J mice, which express higher levels of EGFR, may yield different results. In addition, EGFR-based interventions may depend on the specific type of PH—in contrast to chronic hypoxia-induced PH, our HFD-induced model appears to mimic PH secondary to left-sided heart disease (30). It is worth noting that, although the expression of *Egfr* was higher in the PH-susceptible AKR/J strain than in C57BL/6J, its expression in lung tissue was unchanged between RD and HFD groups. However, a recent study has shown that post-translation redox

mechanisms can account for increased EGFR signaling in the setting of PH, despite constant levels of EGFR protein, including in lung tissue from human subjects with PAH (45). Hence, higher baseline levels of *Egfr* expression may be sufficient to increase disease susceptibility. Although limited to mice and intended for hypothesis generation, the results of this study add new genomic evidence for a role of *Egfr* in experimental PH. Additional work in humans will be necessary to confirm the role of *Egfr* in disease pathogenesis.

Although the focus of this study was on the pulmonary circulation and RV hemodynamics, our results should be widely applicable to a variety of disease states. For example, it is well known that high-fat feeding also contributes to diseases

of the systemic circulation (46, 47). Hence, the results of this comprehensive study may be of use to investigators interested in left-sided cardiovascular disease. In addition, the results of our strain survey should be useful to all investigators studying PH, as this strain survey should provide useful insight into the ideal background strains for future genetic studies of PH susceptibility.

In summary, this report describes the results of a strain study that identified *Egfr* as a candidate gene in HFD-induced PH susceptibility. Our findings and data will be useful to investigators in multiple fields examining the relationship between diet and right- and left-sided cardiovascular hemodynamics. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

References

1. Michelakis ED. Pulmonary arterial hypertension: yesterday, today, tomorrow. *Circ Res* 2014;115:109–114.
2. Simonneau G, Gatzoulis MA, Adatia I, Celermajer D, Denton C, Ghofrani A, Gomez Sanchez MA, Krishna Kumar R, Landzberg M, Machado RF, et al. Updated clinical classification of pulmonary hypertension. *J Am Coll Cardiol* 2013;62(25 suppl):D34–D41.
3. Hoeper MM, Bogaard HJ, Condliffe R, Frantz R, Khanna D, Kurzyna M, Langleben D, Manes A, Satoh T, Torres F, et al. Definitions and diagnosis of pulmonary hypertension. *J Am Coll Cardiol* 2013;62(25 suppl):D42–D50.
4. Austin ED, Loyd JE. The genetics of pulmonary arterial hypertension. *Circ Res* 2014;115:189–202.
5. Lane KB, Machado RD, Pauculo MW, Thomson JR, Phillips JA III, Loyd JE, Nichols WC, Trembath RC; International PPH Consortium. Heterozygous germline mutations in BMPR2, encoding a TGF- β receptor, cause familial primary pulmonary hypertension. *Nat Genet* 2000;26:81–84.
6. Deng Z, Morse JH, Slager SL, Cuervo N, Moore KJ, Venetos G, Kalachikov S, Cayanis E, Fischer SG, Barst RJ, et al. Familial primary pulmonary hypertension (gene PPH1) is caused by mutations in the bone morphogenetic protein receptor-II gene. *Am J Hum Genet* 2000;67:737–744.
7. Soubrier F, Chung WK, Machado R, Grünig E, Aldred M, Geraci M, Loyd JE, Elliott CG, Trembath RC, Newman JH, et al. Genetics and genomics of pulmonary arterial hypertension. *J Am Coll Cardiol* 2013;62(25 suppl):D13–D21.
8. Harrison RE, Flanagan JA, Sankelo M, Abdalla SA, Rowell J, Machado RD, Elliott CG, Robbins IM, Olschewski H, McLaughlin V, et al. Molecular and functional analysis identifies ALK-1 as the predominant cause of pulmonary hypertension related to hereditary haemorrhagic telangiectasia. *J Med Genet* 2003;40:865–871.
9. Shintani M, Yagi H, Nakayama T, Saji T, Matsuoka R. A new nonsense mutation of SMAD8 associated with pulmonary arterial hypertension. *J Med Genet* 2009;46:331–337.
10. Austin ED, Ma L, LeDuc C, Berman Rosenzweig E, Borczuk A, Phillips JA III, Palomero T, Sumazin P, Kim HR, Talati MH, et al. Whole exome sequencing to identify a novel gene (caveolin-1) associated with human pulmonary arterial hypertension. *Circ Cardiovasc Genet* 2012;5:336–343.
11. Ma L, Roman-Campos D, Austin ED, Eyries M, Sampson KS, Soubrier F, Germain M, Tréguët DA, Borczuk A, Rosenzweig EB, et al. A novel channelopathy in pulmonary arterial hypertension. *N Engl J Med* 2013;369:351–361.
12. Eyries M, Montani D, Girerd B, Perret C, Leroy A, Lonjou C, Chelghoum N, Coulet F, Bonnet D, Dorfmueller P, et al. EIF2AK4 mutations cause pulmonary veno-occlusive disease, a recessive form of pulmonary hypertension. *Nat Genet* 2014;46:65–69.
13. Best DH, Sumner KL, Austin ED, Chung WK, Brown LM, Borczuk AC, Rosenzweig EB, Bayrak-Toydemir P, Mao R, Cahill BC, et al. EIF2AK4 mutations in pulmonary capillary hemangiomatosis. *Chest* 2014;145:231–236.
14. Larkin EK, Newman JH, Austin ED, Hemnes AR, Wheeler L, Robbins IM, West JD, Phillips JA III, Hamid R, Loyd JE. Longitudinal analysis casts doubt on the presence of genetic anticipation in heritable pulmonary arterial hypertension. *Am J Respir Crit Care Med* 2012;186:892–896.
15. Lai YC, Potoka KC, Champion HC, Mora AL, Gladwin MT. Pulmonary arterial hypertension: the clinical syndrome. *Circ Res* 2014;115:115–130.
16. Chan SY, Loscalzo J. Pathogenic mechanisms of pulmonary arterial hypertension. *J Mol Cell Cardiol* 2008;44:14–30.
17. Germain M, Eyries M, Montani D, Poirier O, Girerd B, Dorfmueller P, Coulet F, Nadaud S, Maugenre S, Guignabert C, et al. Genome-wide association analysis identifies a susceptibility locus for pulmonary arterial hypertension. *Nat Genet* 2013;45:518–521.
18. Paulin R, Michelakis ED. The metabolic theory of pulmonary arterial hypertension. *Circ Res* 2014;115:148–164.
19. Lawrie A, Hameed AG, Chamberlain J, Arnold N, Kennerley A, Hopkinson K, Pickworth J, Kiely DG, Crossman DC, Francis SE. Paigen diet-fed apolipoprotein E knockout mice develop severe pulmonary hypertension in an interleukin-1-dependent manner. *Am J Pathol* 2011;179:1693–1705.
20. Kelley EE, Baust J, Bonacci G, Golin-Bisello F, Devlin JE, St Croix CM, Watkins SC, Gor S, Cantu-Medellin N, Weidert ER, et al. Fatty acid nitroalkenes ameliorate glucose intolerance and pulmonary hypertension in high-fat diet-induced obesity. *Cardiovasc Res* 2014;101:352–363.
21. Agard C, Rolli-Derkinderen M, Dumas-de-La-Roque E, Rio M, Sagan C, Savineau JP, Loirand G, Pacaud P. Protective role of the antidiabetic drug metformin against chronic experimental pulmonary hypertension. *Br J Pharmacol* 2009;158:1285–1294.
22. Bertero T, Lu Y, Annis S, Hale A, Bhat B, Saggar R, Wallace WD, Ross DJ, Vargas SO, et al. Systems-level regulation of microRNA networks by miR-130/301 promotes pulmonary hypertension. *J Clin Invest* 2014;124:3514–3528.
23. Lai YC, Tabima DM, Dube JJ, Hughan KS, Vanderpool RR, Goncharov DA, St Croix CM, Garcia-Ocaña A, Goncharova EA, Tofovic SP, et al. Sirt3-amp-activated protein kinase activation by nitrite and metformin improves hyperglycemia and normalizes pulmonary hypertension associated with heart failure with preserved ejection fraction. *Circulation* 2016;133:717–731.
24. XHimes BE, Sheppard K, Berndt A, Leme AS, Myers RA, Gignoux CR, Levin AM, Gauderman WJ, Yang JJ, Mathias RA, et al. Integration of mouse and human genome-wide association data identifies KCNP4 as an asthma gene. *PLoS One* 2013;8:e56179.
25. Berndt A, Cario CL, Silva KA, Kennedy VE, Harrison DE, Paigen B, Sundberg JP. Identification of fat4 and tsc22d1 as novel candidate genes for spontaneous pulmonary adenomas. *Cancer Res* 2011;71:5779–5791.
26. Kelly N, Fitch A, Michael H, Ghedin E, Tipton L, Lucht L, Gladwin MT, Shapiro SD, Morris AM. Susceptibility to high-fat diet-induced pulmonary hypertension is associated with reduced GI α -diversity and altered species composition in mice [abstract]. *Am J Respir Crit Care Med* 2015;191:A3961.
27. Kelly N, Radder J, Baust J, Burton C, Agostini B, Wood J, Gregory A, Leme A, Shapiro SD. Genome-wide association study with interactome-based scoring identifies Dok4 and Ppp2ca as novel candidate genes in high-fat diet-induced pulmonary hypertension [abstract]. *Am J Respir Crit Care Med* 2016;193:A2219.
28. Meng Q, Bueno M, Kelly N, Lai Y, Mora AL, Gladwin MT. Evaluation of the therapeutic effects of nitrite and metformin in a mouse model of pulmonary hypertension associated with heart failure and preserved ejection fraction [abstract]. *Am J Respir Crit Care Med* 2016;193:A3888.
29. Zhou X, Stephens M. Genome-wide efficient mixed-model analysis for association studies. *Nat Genet* 2012;44:821–824.
30. Meng Q, Lai Y-C, Kelly NJ, Bueno M, Baust J, Bachman T, Vanderpool RR, Radder JE, Hu J, Morris A, et al. Development of a mouse model of metabolic syndrome, pulmonary hypertension, and heart failure with preserved ejection fraction (PH-HFpEF). *Am J Respir Cell Mol Biol* [online ahead of print] 24 Jan 2017; DOI: 10.1165/rcmb.2016-0177OC.
31. Umar S, Rabinovitch M, Eghbali M. Estrogen paradox in pulmonary hypertension: current controversies and future perspectives. *Am J Respir Crit Care Med* 2012;186:125–131.
32. Petkov PM, Cassell MA, Sargent EE, Donnelly CJ, Robinson P, Crew V, Asquith S, Haar RV, Wiles MV. Development of a SNP genotyping panel for genetic monitoring of the laboratory mouse. *Genomics* 2004;83:902–911.
33. Zhou X, Stephens M. Efficient multivariate linear mixed model algorithms for genome-wide association studies. *Nat Methods* 2014;11:407–409.
34. Parikh VN, Jin RC, Rabello S, Gulbahce N, White K, Hale A, Cottrill KA, Shaik RS, Waxman AB, Zhang YY, et al. MicroRNA-21 integrates pathogenic signaling to control pulmonary hypertension: results of a network bioinformatics approach. *Circulation* 2012;125:1520–1532.

35. Eppig JT, Blake JA, Bult CJ, Kadin JA, Richardson JE; Mouse Genome Database Group. The Mouse Genome Database (MGD): facilitating mouse as a model for human biology and disease. *Nucleic Acids Res* 2015;43:D726–D736.
36. Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, Barabási AL. Disease networks: uncovering disease–disease relationships through the incomplete interactome. *Science* 2015; 347:1257601.
37. Brandes U. On variants of shortest-path betweenness centrality and their generic computation. *Soc Networks* 2008;30:136–145.
38. Dahal BK, Cornitescu T, Tretyn A, Pullamsetti SS, Kosanovic D, Dumitrascu R, Ghofrani HA, Weissmann N, Voswinckel R, Banat GA, *et al.* Role of epidermal growth factor inhibition in experimental pulmonary hypertension. *Am J Respir Crit Care Med* 2010;181: 158–167.
39. Silverman EK, Loscalzo J. Network medicine approaches to the genetics of complex diseases. *Discov Med* 2012;14:143–152.
40. Jia P, Zhao Z. Network-assisted analysis to prioritize GWAS results: principles, methods and perspectives. *Hum Genet* 2014;133: 125–138.
41. Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, Wu W, Uitdehaag BM, Kappos L, Polman CH, *et al.*; GeneMSA Consortium. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009;18:2078–2090.
42. Barrenäs F, Chavali S, Alves AC, Coin L, Jarvelin MR, Jörnsten R, Langston MA, Ramasamy A, Rogers G, Wang H, *et al.* Highly interconnected genes in disease-specific networks are enriched for disease-associated polymorphisms. *Genome Biol* 2012;13:R46.
43. Jones PL, Crack J, Rabinovitch M. Regulation of tenascin-C, a vascular smooth muscle cell survival factor that interacts with the $\alpha v \beta 3$ integrin to promote epidermal growth factor receptor phosphorylation and growth. *J Cell Biol* 1997;139:279–293.
44. Merklinger SL, Jones PL, Martinez EC, Rabinovitch M. Epidermal growth factor receptor blockade mediates smooth muscle cell apoptosis and improves survival in rats with pulmonary hypertension. *Circulation* 2005;112:423–431.
45. Rafikova O, Rafikov R, Kangath A, Qu N, Aggarwal S, Sharma S, Desai J, Fields T, Ludewig B, Yuan JX, *et al.* Redox regulation of epidermal growth factor receptor signaling during the development of pulmonary hypertension. *Free Radic Biol Med* 2016;95:96–111.
46. Nakashima Y, Plump AS, Raines EW, Breslow JL, Ross R. ApoE-deficient mice develop lesions of all phases of atherosclerosis throughout the arterial tree. *Arterioscler Thromb* 1994;14:133–140.
47. Van Gaal LF, Mertens IL, De Block CE. Mechanisms linking obesity with cardiovascular disease. *Nature* 2006;444:875–880.