# Check for updates

# CLINICAL RESEARCH ARTICLE OPEN Interactions between PDA-associated polymorphisms and genetic ancestry alter ductus arteriosus gene expression

Ronald I. Clyman<sup>1</sup>, Nancy K. Hills<sup>2</sup>, John M. Dagle<sup>3</sup>, Jeffrey C. Murray<sup>3</sup> and Keegan Kelsey<sup>3</sup>

**BACKGROUND:** DNA polymorphisms in *PTGIS* and *TFAP2B* have been identified as risk factors for patent ductus arteriosus (PDA) in a population composed of preterm infants with European genetic ancestry but not in more genetically diverse populations. **GOAL:** To determine if the effects of TFAP2B and *PTGIS* polymorphisms on ductus arteriosus (DA) gene expression differ based on genetic ancestry.

**METHODS:** DA from 273 human second trimester fetuses were genotyped for *TFAP2B* and *PTGIS* polymorphisms and for polymorphisms distributing along genetic ancestry lines. RT-PCR was used to measure the RNA expression of 49 candidate genes involved with DA closure.

**RESULTS:** Seventeen percent of the DA analyzed were of European ancestry. In multivariable regression analyses we found consistent associations between four PDA-related *TFAP2B* polymorphisms (rs2817399(A), rs987237(G), rs760900(C), and rs2817416 (C)) and expression of the following genes: *EPAS1*, *CACNB2*, *ECE1*, *KCNA2*, *ATP2A3*, *EDNRA*, *EDNRB*, *BMP9*, and *BMP10*, and between the PTGIS haplotype rs493694(G)/rs693649(A) and *PTGIS* and *NOS3*. These changes only occurred in DA with European ancestry. No consistent positive or negative associations were found among DA samples unless an interaction between the polymorphisms and genetic ancestry was taken into account.

**CONCLUSION:** *PTGIS* and *TFAP2B* polymorphisms were associated with consistent changes in DA gene expression when present in fetuses with European ancestry.

Pediatric Research (2022) 91:903-911; https://doi.org/10.1038/s41390-021-01506-6

## **IMPACT:**

- DNA polymorphisms in PTGIS and TFAP2B have been identified as risk factors for patent ductus arteriosus (PDA) in a population composed primarily of preterm infants with European genetic ancestry but not in more genetically diverse populations.
- The same PTGIS and TFAP2B polymorphisms are associated with changes in ductus gene expression when present in ductus from fetuses with European genetic ancestry.
- No consistent associations with gene expression can be found unless an interaction between the polymorphisms and genetic ancestry is taken into account.

## INTRODUCTION

In contrast with full-term infants, those born before 28 weeks' gestation frequently fail to close their ductus arteriosus (DA) after birth. Persistent DA patency alters cerebral, mesenteric, and renal blood flow, impairs pulmonary mechanics, increases the risk of pulmonary hemorrhage, and prolongs the need for mechanical ventilation. Prior studies have shown that immature gestation, absence of antenatal glucocorticoid exposure, and mother's self-identified race are the most consistent independent risk factors for identifying preterm newborn infants who fail to close their patent ductus arteriosus (PDA) either spontaneously or with inhibitors of prostaglandin production like indomethacin and ibuprofen.<sup>1–5</sup> Both immature gestation and absence of antenatal betamethasone decrease the expression of a wide range of DA genes involved in oxygen-induced constriction (e.g., calcium channels, potassium channels, and endothelin signaling), contractile protein

maturation, prostaglandin- and nitric oxide-mediated relaxation, and tissue inflammation and remodeling.  $^{\rm 5-7}$ 

There is growing evidence from monozygotic twin studies that genetic risk factors may act in concert with gestational age to alter the ability of the DA to close in preterm infants.<sup>8,9</sup> We previously identified several single-nucleotide polymorphisms (SNPs) in genes encoding transcription factor AP-2 beta (*TFAP2B*, the gene mutated in Char syndrome) and prostacyclin synthase (*PTGIS*), which are associated with isolated (non-syndromic) PDAs in preterm infants.<sup>10</sup> *PTGIS* and its vasodilatory lipid product, prostacyclin (PGI2), play an important role in maintaining preterm DA patency.<sup>11</sup> Similarly, TFAP2B, a transcription factor that regulates endothelin, hypoxia inducible factor 2-alpha (HIF2 alpha), and calponin, plays an important role in DA smooth muscle development.<sup>10,12,13</sup> We previously examined one of the *TFAP2B* polymorphisms (SNP *rs2817399(A*)) that has been

Received: 16 February 2021 Revised: 12 March 2021 Accepted: 16 March 2021 Published online: 9 April 2021

<sup>&</sup>lt;sup>1</sup>Department of Pediatrics and Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA, USA; <sup>2</sup>Departments of Epidemiology and Biostatistics, and Neurology, University of California San Francisco, San Francisco, CA, USA and <sup>3</sup>Department of Pediatrics, University of Iowa, Iowa City, IA, USA Correspondence: Ronald I. Clyman (clymanr@peds.ucsf.edu)

Interactions between PDA-associated polymorphisms and genetic ancestry... RI Clyman et al.

904

associated with persistent DA patency, for its effects on human fetal DA gene expression and found that it decreased several of the same calcium- and potassium-channel genes previously shown to be involved with oxygen-induced constriction of the DA.<sup>6</sup>

In contrast with our findings, two subsequent epidemiologic studies<sup>14,15</sup> failed to find an association between the same SNPs we identified in *TFAP2B* and *PTGIS* and alterations in DA closure. Although differences in both the definition of PDA and the strategies used to treat the PDA might account for the discordant results between studies, another explanation might be the significant differences in genetic ancestry among the study populations. Ninety percent of mothers in our original lowabased, single center study self-identified as White/European ancestry.<sup>10</sup> In contrast, 50 and 0% of the populations in the subsequent two studies self-identified as European ancestry.<sup>14,15</sup> In addition, the lowa study utilized a family-based approach, which is less susceptible to the effects of population stratification compared to the case–control design used in the latter studies.

We designed the following study to determine whether the PDA-associated SNPs in *TFAP2B* and *PTGIS* that we previously identified are indeed associated with unique alterations in gene expression. Our goal was to test the reproducibility of our prior findings in fetal DA obtained from a population with diverse genetic ancestry and to expand the list of genes that might be affected by the *TFAP2B* and *PTGIS* polymorphisms. We hypothesized that an interaction exists between the fetus's genetic ancestry and the SNPs in *TFAP2B* and *PTGIS* such that the effects of the SNPs on gene expression only occur in DA with European genetic ancestry.

## METHODS

We used de-identified DNA and RNA samples, collected as part of a prior study,<sup>7</sup> to determine whether common genetic variants in *TFAP2B* and *PTGIS*, which have been associated with a PDA in preterm newborns, are associated with unique patterns of gene expression in the human fetal DA. The study was reviewed by the Institutional Review Board of the University of California San Francisco and given an exempt status.

## Tissue

Human tissue was obtained under the oversight of the Institutional Review Board at University of California San Francisco. Midgestation (13<sup>0/7</sup>–23<sup>6/7</sup> weeks) human fetal DA and ascending aorta were collected from elective pregnancy terminations in healthy women with no known fetal abnormalities. Consent for the use of fetal tissue for research purposes was obtained by the clinic staff, who had been trained in human subjects' protections. The consent for the use of fetal tissue for research purposes is separate from the consent for the clinical procedure. Researchers have no patient contact and only receive de-identified tissues.

Prostaglandins were not used during the terminations. Cervical ripening was performed with laminaria (compressed seaweed). Fetal tissue was immediately submerged in calcium- and magnesium-free phosphate-buffered saline at 4 °C following delivery. The DA and aorta were dissected in the chilled buffer solution and the isolated DA and aorta were snap frozen in liquid nitrogen (between 1.5 and 2 h after delivery). Gestational age was determined by fetal foot length.<sup>16</sup> De-identified tissues were individually labeled and stored for later analysis. Individual samples were analyzed in "batches" of 90 samples. There was no "pooling" or combining of tissues during the analyses.

During the period of the study, women who donated tissue selfidentified their racial origins to the clinic staff as White/European ancestry = 21%, Non-White/Non-European ancestry = 76%, and unknown = 3%. The data on self-reported racial origins were available solely as a population-level statistic. Individual descriptors were not linked to de-identified tissues samples. No clinical information was available for analysis.

Preparation of total RNA, reverse transcription, and quantitative PCR

We examined the RNA expression of 49 "DA closure genes" in each of the 273 human DA samples (Table 1). The "DA closure genes" were chosen because: (1) their expression in the DA has previously been shown to differ from their expression in the aorta, and (2) their mutations or polymorphisms (or their pharmacologic inhibition) has been shown to affect DA closure (see refs. <sup>7,6</sup> for references for "DA closure genes").

Total RNA was isolated from each individual DA and cDNA was generated as described elsewhere.<sup>6,17</sup> We used the TaqMan Universal PCR master mix of PE Applied Biosystems (Foster City, CA) to quantify gene expression in a 96-well format. TagMan probes were designed using the Primer Express program and labeled with fluorophores FAM (6-caboxy-fluorescein) and TAMRA (6 carboxy-tetramethyl-rhodamine) as reporter and guencher dyes, respectively. An ABI PRISM 7500 Sequence detection system was used to determine the cycle threshold (CT). Reactions were carried out in triplicate. Data were analyzed using the Sequence Detector version 1.6.3 program. The degree of expression of the gene of interest was determined using the relative gene expression method. Malate dehydrogenase (MDH) was used as an internal control to normalize the data.<sup>6,18</sup>  $\Delta$ CT represents the difference in cycle threshold (CT) between the expression of the housekeeping gene (MDH) and the gene of interest. Each unit of  $\Delta$ CT represents a twofold change in mRNA levels. The more negative the  $\Delta$ CT, the fewer the number of starting copies of a gene's mRNA.

DNA genotyping of fetal ductus arteriosus to determine the presence or absence of several TFAP2B and PTGIS SNPs as well as to infer genetic ancestry

DNA was extracted from the ascending aorta of each of the fetal samples using the QIAamp DNA mini kit (Qiagen Inc., Valencia, CA). DNA was quantified spectrophotometrically. Allelic variation was determined by using the TaqMan genotyping system (Applied Biosystems, Foster City, CA), as previously described.<sup>19</sup> Allele scoring was performed using the Sequence Detection Systems 2.2 software (Applied Biosystems).

We examined the DNA for the presence of several SNPs in the *TFAP2B* and *PTGIS* genes that have been associated with altered DA closure in preterm infants (Fig. 1)<sup>10</sup> (Dagle et al., unpublished results). Specifically, we examined the DNA for four SNPs in *TFAP2B* (rs2817399: (A allele); rs987237: (G allele); rs760900: (C allele); rs2817416 (C allele)) that have been associated with delayed DA closure (even in the presence of indomethacin). We also determined the presence of two *TFAP2B* SNPs that are unrelated to the timing of DA closure (rs2817419: G allele; and rs2635727: T allele). In addition, we examined one haplotype combination of two neighboring SNPs in the gene *PTGIS* (rs493694 (G allele) and *PTGIS* rs693649 (A allele)) that is negatively associated with PDA, suggesting a preventative effect of the allele combination against PDA.<sup>10</sup>

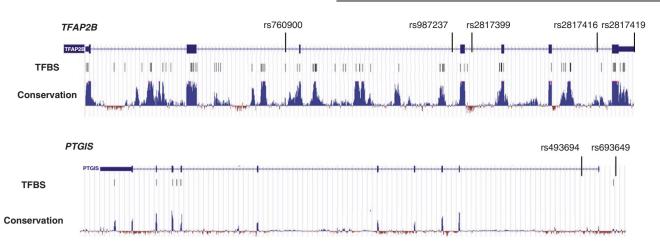
To identify the genetic ancestry of the fetal tissues, we examined four genes whose sequence polymorphisms are distributed along genetic lines between European ancestry and Non-European ancestry (African, Chinese, and Japanese) populations.<sup>20–22</sup> These include *SLC24A5* SNP rs1426654 (100% European = **A** allele; 98% Non-European = **G** allele;<sup>20,21</sup>; *SLC45A2* SNP rs16891982 (100% European = **G** allele; 100% Non-European = **C** allele;<sup>20,21</sup>; *DARC* SNP rs2814778 (100% European = **A** allele; 85–100% Non-European = **G** allele;<sup>22</sup>; and *HERC2* SNP rs12913832 (79% European = **G** allele; 100% Non-European = **A** allele).<sup>21</sup> Each fetal sample possessed between 0 and 8 of the Non-European ancestry alleles. Although we recognized that the clinic's population data for race

Genes/aliases	Regression	coefficient	Regression coefficients for Gestation <sup>a</sup>	'nª				Regressio	n coefficien	Regression coefficients for non-European ancestry <sup>b</sup>	uropean an	cestry <sup>b</sup>		
	TFAP2B rs760900	TFAP2B rs987237	TFAP2B rs2817399	TFAP2B rs2817416	TFAP2B rs2817419	TFAP2B rs2635727	<i>PTGIS</i> Haplotype	TFAP2B 760900	TFAP2B 987237	TFAP2B 2817399	TFAP2B 2817416	TFAP2B 2817419	TFAP2B 2635727	<i>PTGIS</i> Haplotype
	U	IJ	A	U	ט	г		U	ט	٩	U	ט	μ	
Ca <sup>2+</sup> signaling														
ATP2A3/SERCA	-0.032**	-0.031**	$-0.029^{*}$	-0.033**	-0.031**	-0.029*	-0.032**							
CACNA1C/CaV1.2	-0.056**	-0.058**	-0.052**	-0.059**	-0.056**	-0.053**	-0.055**							
CACNA1G/Cav3.1	-0.136**	-0.139**	-0.138**	$-0.141^{**}$	-0.137**	-0.137**	-0.139**	0.164						
CACNB2/Cavbeta2	-0.061**	-0.062**	-0.061**	-0.064**	-0.062**	-0.059**	-0.061**							
RHOB	0.039**	0.042**	0.042**	0.041**	0.041**	0.044**	0.041**							
ROCK1	0.048*	0.042	0.031	0.036	0.042	0.044*	0.043							
${f K}^+$ channels														
KCNA2/Kv1.2	-0.208**	-0.203**	-0.204**	-0.207**	-0.207**	-0.195**	-0.201**							
KCNA5/Kv1.5														
KCNB1/Kv2.1	-0.076**	-0.081**	-0.075**	-0.086**	$-0.081^{**}$	-0.075**	-0.08**							
KCNS3/Kv9.3	-0.235**	-0.236**	-0.246**	$-0.241^{**}$	-0.233**	-0.222**	-0.239**							
KCNJ8/Kir6.1	0.034*	0.032*		0.028	0.032*	0.04**	0.036*							
ABCC9/SUR2B	0.056**	0.05**	0.043**	0.045**	0.051**	0.055**	0.051**							
<b>Contractile proteins</b>														
ACTA2	0.068**	0.069**	0.07**	0.066**	0.068**	0.071**	0.069**							
CNN1/Calponin	0.125**	0.121**	0.119**	0.121**	0.122**	0.124**	0.122**							
MYH11/SM1	0.101**	0.095**	0.092**	0.095**	0.098**	0.099**	0.095**							
MYH11/SM2	0.148**	0.146**	0.137**	0.144**	0.146**	0.148**	0.144**							
MYLK	0.063**	0.063**	0.062**	0.057**	0.063**	0.066**	0.063**							
MYOCD/Myocardin	-0.044**	-0.045**	-0.038**	-0.042**	-0.045**	-0.041**	-0.044**							
TPM1/Tropomyosin	0.044**	0.041**	0.045**	0.038**	0.042**	0.045**	0.042**							
Endothelin signaling														
ECE1	0.028**	0.025**	0.021**	0.022**	0.026**	0.027**	0.024**							
EDNRA/EtA-receptor	-0.066**	-0.068**	$-0.068^{**}$	-0.073**	-0.067**	-0.067**	-0.068**							
EDNRB/EtB-receptor	-0.068**	-0.066**	-0.073**	-0.074**	-0.066**	-0.064**	-0.068**							
Prostaglandin signaling														
PTGS1/COX1	0.09**	0.084**	0.091**	0.084**	0.085**	0.087**	0.086**							
PTGS2/COX2	0.05**	0.048**	0.047**	0.041**	0.049**	0.053**	0.047**	0.287**	0.238*	0.24*	0.243*	0.278**	0.241*	0.25*
CYP8A1/PTGIS	0.045**	0.041**	0.04**	0.04**	0.041**	0.044**	0.042**							
PTGER4/EP4	0.074**	0.07**	0.067**	0.069**	0.069**	0.072**	0.07**							
PDE1B	-0.051**	-0.053**	-0.052**	-0.055**	-0.052**	-0.048**	-0.053**							
PDE3B	0.038**	0.039**	0.039**	0.039**	0.038**	0.043**	0.039**							
PDE4D	0.044**	0.037**	0.031*	0.031*	0.037**	0.039**	0.036**							

**SPRINGER NATURE** 

	-	gression	coefficients	Regression coefficients for Gestation <sup>a</sup>	n <sup>a</sup>				Regressio	n coefficien	Regression coefficients for non-European ancestry <sup>b</sup>	uropean ar	icestry <sup>b</sup>		
NHTC oxide signaling         Notic oxide signaling           No53(eVCS         0.066**         0.064**         0.065**         0.065**         0.065**           PDESA         PDESA         PDESA         0.065**         0.065**         0.065**         0.065**           PDESA         PDESA         PDESA         PDESA         PDESA         PDESA           PDESA         PDESA         PDESA         PDESA         PDESA         PDESA           AGTRI         -013**         -013**         -0132**         -0132**         -0136**           AGTRI         -0104**         -0.047**         -0.047**         -0.038**         -0.038**           BMPA         -0.047**         -0.047**         -0.038**         -0.038**         -0.038**         -0.038**           BMPA         -0.047**         -0.037**         -0.038**         -0.038**         -0.038**         -0.038**           BMPA         -0.037**         -0.058**         -0.058**         -0.038**         -0.038**         -0.038**           BMPA         -0.037**         -0.068**         -0.028**         -0.038**         -0.038**           BMPA         -0.037**         -0.068**         -0.047**         -0.047**           BMPA			TFAP2B rs987237 G	TFAP2B rs2817399 A	<i>TFAP2B</i> rs2817416 C	<i>TFAP2B</i> rs2817419 G	TFAP2B rs2635727 T	<i>PTGIS</i> Haplotype	<i>TFAP2B</i> 760900 C	TFAP2B 987237 G	<i>TFAP2B</i> 2817399 A	<i>TFAP2B</i> 2817416 C	<i>TFAP2B</i> 2817419 G	<i>TFAP2B</i> 2635727 Т	<i>PTGIS</i> Haplotype
NOS3/eNOS         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.064**         0.065**         0.065**         0.065**         0.065**         0.065**         0.065**         0.064**         0.013**         0.023**															
PESA         Inflammation and tendeling         AGTR       -0.134**       -0.141**       -0.142**       -0.132**       -0.136**         AGTR       -0.134**       -0.134**       -0.141**       -0.142**       -0.136**         AGTR       -0.104**       -0.141**       -0.142**       -0.136**       -0.014**         AGTR       -0.044**       -0.044**       -0.044**       -0.044**       -0.034**         AMPI       -0.044**       -0.304**       -0.034**       -0.034**       -0.034**         BMP       -0.312**       -0.034**       -0.034**       -0.024**       -0.232**         BMP       -0.034**       -0.034**       -0.034**       -0.044**       -0.044**         BMP       -0.244**       -0.034**       -0.044**       -0.044**       -0.044**         AGT       -0.244**       -0.044**       -0.044**       -0.044**       -0.044***         AGT		.068**	0.064**	0.06**	0.059**	0.065**	0.067**	0.065**							
Inflammation and termination and activity of the second and activity of the second and activity of the second activity activity of the second activity of the se	PDE5A														
AGTR1         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.134**         -0.014**         0.009**         0.009**         0.009**         0.004**         0.004**         0.004**         0.004**         0.014**         0.014**         0.014**         0.014**         0.014**         0.005**         0.035**         0.035**         0.036**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.035**         0.025***         0.025***         0.0	Inflammation and remodeling														
AGTR2         0.107**         0.101**         0.010**         0.00**         0.009**         0.101**         0.101**           BMMP         -0.047**         -0.049**         -0.043**         -0.048**         -0.048**           BMMP1         -0.047**         -0.049**         -0.032**         -0.038**         -0.327**           BMP1         -0.345**         -0.339**         -0.333**         -0.337**         -0.327**           BMP10         -0.345**         -0.339**         -0.328**         -0.232**         -0.232***           BMP10         -0.312**         -0.339**         -0.232**         -0.232***         -0.232***           BMP10         -0.312**         -0.304**         -0.232**         -0.232***         -0.232***           BMP10         -0.312**         -0.204**         -0.232**         -0.232***         -0.232***           FN1         0.025**         0.001**         0.025**         0.002**         0.035***           JG1         -0.245**         -0.204**         -0.204**         -0.204***         -0.204***           JG1         -0.045**         -0.048**         0.007***         0.007***         0.019***********************************	-	.134**	-0.139**	-0.141**	-0.142**	-0.139**	-0.132**	-0.136**							
BMP40.447* - 0.049* - 0.042**0.053**0.048**0.048**0.048**0.048**0.048**0.349**0.331**0.332**0.331**0.332**0.331**0.331**0.331**0.332**0.328**0.237**0.207**0.217**0.217**0.217**0.217**0.217**0.217**0.217**0.207**0.217**0.207**0.217**0.207**0.217**0.207**0.217**0.207**0.217**0.207**0.217**0.207**0.207**0.217**0.207**0.217**0.207**0.217**0.207**0.207**0.207**0.217**0.207**0.217**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.207**0.045**0.04		.107**	0.101**	0.09**	0.09**	0.099**	0.102**	0.101**							
BMP9 - 0.345 <sup>++</sup> - 0.339 <sup>++</sup> - 0.338 <sup>++</sup> - 0.333 <sup>++</sup> - 0.322 <sup>++</sup> - 0.322 <sup>++</sup> - 0.321 <sup>++</sup> - 0.296 <sup>++</sup> - 0.286 <sup>++</sup> - 0.296 <sup>++</sup> - 0.216 <sup>++</sup> - 0.021 <sup>++</sup>	Ι	.047**	-0.049**	-0.042**	-0.053**	-0.048**	-0.044**	-0.048**							
BMP10 -0.312** -0.304** -0.302** -0.296** -0.296** -0.286** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.282** -0.284** -0.284** -0.284** -0.284** -0.284** -0.21** -0.24*** -0.24*** -0.24*** -0.24*** -0.24*** -0		.345**	-0.349**	-0.331**	-0.338**	-0.333**	-0.322**	-0.321**							
EPAS1/HIF2 alpha         0.059**         0.056**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.059**         0.002**         0.021**         0.031**         0.019*         0.010**         0.019*         0.019*         0.019*         0.019*         0.019*         0.010**         0.010**         0.010**         0.010**         0.010**         0.010**		.312**	-0.304**	-0.302**	-0.296**	-0.295**	-0.286**	-0.282**							
FNI         0.022**         0.021**         0.021**         0.022**         0.021**         0.022**         0.021**         0.022**         0.021**         0.022**         0.021**         0.019         0.016*         0.013**         0.019         0.016*         0.016*         0.021**         0.023**         0.019         0.019         0.016*         0.021**         0.023**         0.019         0.019         0.019         0.019         0.011*         0.021**         0.033**         0.033**         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0.019         0		.059**	0.059**	0.056**	0.056**	0.059**	0.059**	0.059**							
IGF10.214** -0.218** -0.209** -0.216** -0.215** -0.207** -0.21** ILK 0.075** 0.068** 0.062** 0.059** 0.068** 0.072** 0.048** -0.047** JAG10.045** -0.047** -0.044** -0.049** -0.048** -0.047** MAPKI/ERK2 0.023* 0.019 0.016* 0.016* 0.02 0.023* 0.019 PDGFB/PDGF-B chain -0.042** -0.042** -0.042** -0.039** -0.047** PTPN11 - 0.042** -0.042** -0.042** -0.031** -0.039** -0.047** PTPN11 - 0.042** -0.042** -0.042** -0.042** -0.039** -0.047** TFPN11 - 0.042** -0.042** -0.042** -0.042** -0.039** -0.039** -0.045** TFPN11 - 0.042** -0.042** -0.042** -0.042** -0.039** -0.039** -0.045** TFPN11 - 0.034** -0.031** 0.031** -0.031** -0.031** -0.031** -0.034** TFPN25 beta 0.049** 0.046** 0.031** 0.048** 0.048** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.046** 0.031** 0.048** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.048** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.048** 0.048** 0.046** TFAP2B/TFAP2 beta 0.051** 0.046** 0.031** 0.045** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.046** 0.051** 0.045** 0.048** 0.046** TFAP2B/TFAP2 beta 0.046** 0.051** 0.045** 0.045** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.046** 0.051** 0.048** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.045** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.046** 0.046** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.046** 0.046** 0.046** TFAP2B/TFAP2 beta 0.049** 0.045** 0.045** 0.046** 0.046** 0.046** TFAP2B/TFAP2 beta 0.046** 0.051** 0.045** 0.046** 0.046** 0.046** 0.046** TFAP2B/TFAP2 beta 0.046** 0.051** 0.045** 0.046** 0.046** 0.046** 0.046** 0.046** 0.046** 0.046** 0.046** 0.046** TFAP2B/TFAP2 beta 0.049** 0.046** 0.051** 0.046** 0.04		.022**	0.021**	0.021**	0.02*	0.021**	0.022**	0.02**							
ILK $0.075^{**}_{5*}$ $0.062^{**}_{7*}$ $0.062^{**}_{7*}$ $0.066^{**}_{7*}$ $0.066^{**}_{7*}$ $0.066^{**}_{7*}$ $0.066^{**}_{7*}$ JAG1 $-0.045^{**}_{7*}$ $-0.047^{**}_{7*}$ $-0.034^{**}_{7*}$	I	.214**	-0.218**	-0.209**	-0.216**	-0.215**	-0.207**	-0.21**							
JG10.045** -0.047** -0.047** -0.047** -0.042** -0.042** -0.047** -0.047** -0.047** -0.047** -0.047** -0.045** -0.045** -0.047** -0.045** -0.047** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045*** -0.045**** -0.045****0.045*** -0.045*** -0.045*** -0.045**		.075**	0.068**	0.062**	0.059**	0.068**	0.072**	0.066**							
$ \begin{array}{llllllllllllllllllllllllllllllllllll$	Ι	.045**	-0.047**	-0.044**	-0.049**	-0.045**	-0.048**	-0.047**							
PDGFB/PDGF-B chain         -0.042**         -0.043**         -0.043**         -0.045**           PTPN11         -0.042**         -0.043**         -0.043**         -0.045**           PTPN11         -0.045**         -0.043**         -0.045**         -0.045**           SMARCA4/BRG1         -0.348**         -0.034**         -0.036**         -0.031**         -0.034**           TGFB1/TGF beta1         0.051**         0.039**         0.048**         0.047**         -0.034**           TFAP2B/TFAP2 beta         0.049**         0.048**         0.047**         0.046**         0.047**           TRAF1         0.049**         0.045**         0.046**         0.046**         0.046**           Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods" Each model included the SNP of interest, gestation:           accestry.         ****         0.046**         0.046**         0.046**		.023*	0.019	0.016*	0.016*	0.02	0.023*	0.019							
PTN11       -0.015       -0.018*       -0.014         SMARCA4/BRG1       -0.348**       -0.034**       -0.035**       -0.034**       -0.034**         SMARCA4/BRG1       -0.348**       -0.034**       -0.036**       -0.036**       -0.034**         TGFB1/TGF beta1       0.051**       0.046**       0.033**       -0.034**       0.043**         TFAP2B/TFAP2 beta       0.049**       0.031**       0.048**       0.046**       0.031**         TFAP2B/TFAP2 beta       0.049**       0.046**       0.048**       0.046**       0.046**         TFAP2B/TFAP2 beta       0.049**       0.048**       0.046**       0.046**       0.045**         TFAP2B/TFAP2 beta       0.049**       0.048**       0.046**       0.045**       0.045**         TFAP2       Exercise       0.049**       0.048**       0.046**       0.046**         TRAF1       0.049**       0.046**       0.046**       0.046**       0.046**         Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods" Each model included the SNP of interest, gestation:         accestry.       ************************************	I	.042**	-0.045**	-0.042**	-0.051**	-0.043**	-0.039**	-0.045**							
SMARCA4/BRG1 -0.348** -0.034** -0.032** -0.035** -0.033** -0.034** -0.034** TGFB1/TGF beta1 0.051** 0.046** 0.051** 0.039** 0.048** 0.049** 0.047** TFAP2B/TFAP2 beta 0.049** 0.046** 0.031** 0.045** 0.048** 0.048** 0.046** TRAF1 Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods". Each model included the SNP of interest, gestation: ancestry. Positive regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gen	PTPN11		-0.015		-0.018*	-0.014									
TGFB1/TGF beta1 0.051** 0.046** 0.051** 0.039** 0.048** 0.048** 0.047** TFAP2B/TFAP2 beta 0.046** 0.046** 0.031** 0.045** 0.048** 0.046** TRAF1 Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods". Each model included the SNP of interest, gestati ancestry. Positive regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represe	I	.348**	-0.034**	-0.032**	-0.036**	-0.033**	-0.031**	-0.034**							
TFAP2B/TFAP2 beta 0.049** 0.046** 0.031** 0.045** 0.045** 0.048** 0.046** TRAF1 Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods". Each model included the SNP of interest, gestativ ancestry. Positive regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every destation.		.051**	0.046**	0.051**	0.039**	0.048**	0.049**	0.047**							
TRAF1 Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods". Each model included the SNP of interest, gestati ancestry. "Positive regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for ev gestation.		.049**	0.046**	0.031**	0.045**	0.045**	0.048**	0.046**							
Regression coefficients for gestation and non-European genetic ancestry were derived in the multivariable models described in "Methods". Each model included the SNP of interest, gestation and succestry. ancestry. Positive regression coefficients represent the increase in a gene's ACT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ACT for every estation.	TRAF1														
ancesury. Positive regression coefficients represent the increase in a gene's ΔCT for every increased week of gestation; negative regression coefficients represent the decrease in a gene's ΔCT for every gestation.	Regression coefficients for gestati	ion and n	ion-Europeai	n genetic anc	estry were dei	rived in the m	ultivariable m	odels describe	d in "Methc	ods". Each mo	odel included	l the SNP of	interest, gest	ational age,	and genetic
gestation. br	ancesuy. <sup>a</sup> Positive regression coefficients re	epresent	the increase	in a gene's ∆	CT for every ir	ncreased week	of gestation;	negative regr	ession coeffi	icients repre	sent the deci	rease in a ge	ine's <b>ΔCT</b> for	every increa	sed week o
"Regression coefficients represent the increase in a genes ΔLI when ductus from fetuses with non-European ancestry were compared to those from fetuses with European ancestry.	gestation. <sup>b</sup> Regression coefficients represen	nt the inc	rease in a ge	ene's ΔCT wh	en ductus fro	m fetuses with	non-Europe	an ancestry w	ere compar	ed to those	from fetuses	s with Europ	ean ancestry		

Interactions between PDA-associated polymorphisms and genetic ancestry... RI Clyman et al.



**Fig. 1 Gene structure of human TFAP2B and PTGIS and location of polymorphisms.** The TFAP2B polymorphism rs2635727 is located further to the right of the gene and is not shown on the figure. In contrast to PTGIS, the intronic regions of TFAP2B show many areas of conservation, some with consensus transcription factor-binding sites. Figures were obtained from the UCSC Genome Browser. Note that the two genes are represented in opposite orientations and not to the same scale. The exons of each gene are shown as the thick lines at the top of the figure. TFBS = conserved transcription factor-binding sites; conservation estimated sequence = conservation of genetic loci among 100 vertebrate species (PhyloP).<sup>24</sup>

was an imperfect proxy for genetic ancestry, we used the clinic's statistics to help create a definition for European and Non-European ancestry. Since 76% of the women who donated tissue self-identified as Non-White/Non-European ancestry and 83% of the samples had two or more Non-European ancestry alleles (see "Results"), we defined tissues with two or more Non-European ancestry alleles as "Non-European origin" and those with zero or one Non-European ancestry allele as "European origin" (The bold lettering is to highlight the only change in the SNPs).

#### Statistical analyses

Stata software (Release 16.1; StataCorp LP, College Station, TX) was used for all statistical analyses. We used multivariable linear regression to build statistical models that adjusted for possible confounding effects of gestational age and genetic ancestry on the relationship between a SNP, or 2-SNP haplotype, and the change in RNA expression of each of the 49 "DA closure genes" (represented by their  $\Delta$ CT). The multivariable models were analyzed using generalized estimating equations techniques to account for clustering within each of the 90 sample "batch" assays. Coefficients derived from these models were interpreted as the difference (positive or negative) between the RNA expression in the presence of the SNP in the study population and that in the absence of the SNP while holding gestational age and genetic ancestry constant. Models were run individually for each of the 49 "DA closure genes".

To determine if an SNP's effect on RNA expression differed depending on whether it occurred in a European ancestry or non-European ancestry DA, we added an interaction term to the model (between the SNP in question and genetic ancestry) and reran the regression.

Our study was an exploratory study designed to identify "DA closure genes" that might be altered by the presence of common genetic variants. Because of its exploratory nature, we considered any association between an SNP and a change in gene expression as possible evidence of association if the regression coefficient for RNA expression had a p value < 0.1. Our purpose was to decrease the likelihood of missing true positive signals, knowing that false-positive signals will inevitably be present.

#### RESULTS

We analyzed 273 fetal DA samples in the current study (gestational age =  $19.8 \pm 2.9$  weeks (m ± s.d.)). Seventeen percent

of the samples had zero or one non-European ancestry allele and were assigned as European ancestry. The allele frequencies of the *TFAP2B* polymorphisms associated with an increased incidence of PDA were as follows: rs2817399 (A allele) = 81%; rs987237 (G allele) = 37%; rs760900 (C allele) = 89%; and rs2817416 (C allele) = 25%). The frequencies of the *TFAP2B* polymorphisms that are unrelated to the timing of DA closure were: (rs2817419 (G allele) = 42% and rs2635727 (T allele) = 38%). The frequency of the *PTGIS* haplotype of two neighboring SNPs that is negatively associated with PDA was: (rs493694 (G allele)/rs693649 (A allele) = 21%).

We created multivariable linear regression models to determine the independent effects of gestational age, genetic ancestry, and the SNP alleles on RNA expression of the 49 "DA closure genes". As previously reported, advancing gestational age was independently associated with changes in RNA expression for the majority (92%) of the "DA closure genes" (Table 1). In contrast, genetic ancestry was only consistently and independently associated with RNA expression in 2 genes: *PTGS2/COX2* (cyclooxygenase 2) and *SLOCA2A1* (the prostaglandin transporter which regulates prostaglandin reuptake) (Table 1).

Our main objective was to identify "DA closure genes" that are modified by the *TFAP2B* and *PTGIS* SNPs that have previously been shown to alter DA behavior: rs2817399 (A allele), rs987237 (G allele), rs760900 (C allele), and rs2817416 (C allele). In our initial examination of the general population of 273 samples, we found no consistent independent association between the *TFAP2B* SNPs associated with delayed DA closure and alterations in RNA expression for any of the "DA closure genes" (Table 2—General population).

However, when we tested whether an interaction occurred between the fetus's genetic ancestry and the same PDAassociated TFAP2B SNPs, we found that several of the "DA closure genes" had consistent, independent changes in gene expression when the SNPs occurred in samples with European ancestry. At least three of the four TFAP2B SNPs were associated with changes in expression in each of the following genes: EPAS1 (HIF2 alpha), CACNB2 (Cavbeta2 calcium channel subunit), ECE1 (endothelin converting enzyme), KCNA2 (potassium channel Kv1.2), ATP2A3 (SFRCA sarcoplasmic reticulum Calcium-ATPase), EDNRA (endothelin A-receptor), EDNRB (endothelin B-receptor), BMP9 (bone morphogenetic protein-9), and BMP10 (bone morphogenetic protein-10) (Table 2-European ancestry). None of these changes were seen when the same SNPs were examined in the

	Regressio	n coefficien	ts for <i>TFAP2</i> 1	Regression coefficients for <i>TFAP2B</i> (PDA-associated polymorphisms)	iated polyn	norphisms)							Regression coefficients for <i>TFAP2B</i> (non-PDA-associated polymorphisms)	for <i>TFAP2B</i> ssociated sms)
	General p	General population <sup>a</sup>			European ancestry <sup>b</sup>	ancestry <sup>b</sup>			Non-Eurol	Non-European ancestry <sup>c</sup>	try <sup>c</sup>		European ancestry <sup>b</sup>	ncestry <sup>b</sup>
	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs2817419 G	rs2635727 T
					-0.444*	-0.511**	-0.411*			0.126				
CACNB2/Cavbeta2	-0.215*				-0.357	-0.493**								
					1.361*	-1.301**	-1.084							
													1.361	1.342
				-0.194* -0.194*				-0.267				-0.385** -0.212*		
Contractile proteins														
	0.235*				0.364*	0.25								
					0.832*		0.773**						0.649*	0.712*
					0.509*		0.381**							
Endothelin signaling														
					-0.281**	-0.174*	-0.243*	-0.272**					0.209*	0.191
EDNRA/EtA-receptor	-0.258**			-0.109	-0.329	-0.228	-0.293	-0.27	-0.231*					
EDNRB/EtB-receptor	-0.237			-0.206*	-0.394	-0.515*		-0.58*					0.339	
Prostaglandin Signaling														
					0.207		0.404							
						-0.422**							0.348	0.341
						-0.328*								
						-0.35								
SLCO2A1/PG transporter	-0.259*				-0.361				-0.221					
Nitric oxide signaling														
						-0.297								
Inflammation and remodeling														
						-0.765*				0.37*	0.379*			
						-0.361*								
	-1.13*				-2.079**	-1.841*	-1.909*							
	-0.891*				-1.531**	-1.598**	-1.438*							
EPAS1/HIF2 alpha					-0.238		-0.235	-0.341**						
										0 328				
						-0.93/*	-0.92			0000				

# Interactions between PDA-associated polymorphisms and genetic ancestry... RI Clyman et al.

**SPRINGER NATURE** 

Table 2. continued														
Genes/Aliases	Regressio	n coefficien	Regression coefficients for <i>TFAP2B</i> (		PDA-associated polymorphisms)	orphisms)							Regression coefficients for <i>TFAP2B</i> (non-PDA-associated polymorphisms)	TFAP2B ciated
	General p	General population <sup>a</sup>			European ancestry <sup>b</sup>	ancestry <sup>b</sup>			Non-Europ	Non-European ancestry <sup>c</sup>	y℃		European ancestry <sup>b</sup>	stry <sup>b</sup>
	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs760900 C	rs987237 G	rs2817399 A	rs2817416 C	rs760900       rs987237       rs2817416       rs2817399       rs2817416       rs2815727         C       G       A       C       C       G       A       C       G       A       C       G       T	635727
PTPN11 SMARCA4/BRG1 TRAF1						0.208 0.173 0.346		-0.198		0.111			0.172	
The bold values indicate the genes where at least three of the four TFAP2B SNPs were associated with changes in expression. Regression coefficient represents the increase in a gene's $\Delta CT$ when the TFAP2B allele was present (compared with when it was absent). Regression coefficients are only listed in the table if the association with the "ductus closure gene" has a <i>p</i> value < 0.10. Negative regression coefficients are only listed in the table if the association with the "ductus closure gene" has a <i>p</i> value < 0.10. Negative regression coefficients are only listed in the table if the association with the "ductus closure gene" has a <i>p</i> value < 0.10. Negative regression coefficients are only listed for gestational age and genetic ancestry (European or Non-European) without an interaction term between the SNP allele and genetic ancestry: <sup>a</sup> General population: multivariate analyses were adjusted for gestational age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry: regression coefficients were obtained for the referent value: European genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry: regression coefficients were obtained for the referent value: non-European ancestry: multivariate analyses were adjusted for gestational age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry: regression coefficients were obtained for the referent value: non-European ancestry: multivariate analyses were adjusted for gestational age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry: regression coefficients were obtained for the referent value: non-European ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry: regression coefficients were obtained for the referent value: non-European genetic ancestry (European or Non-Europe	the where at the increase on coefficien: analyses wei e referent vi riate analyse	: least three the age three the age only lis the adjusted re adjusted f alue: Europe: es were adju treferent val	of the four TI ACT when the sted in the tal for gestationa an genetic ar isted for gest ue: non-Eurol	FAP2B SNPs v TFAP2B allell se TFAP2B allell all age and gen l age and ger cestry. cational age a	vere associal evere associal ciation with enetic ances netic ancestr ind genetic ancestry.	ted with chi ti (compared the "ductus stry (European y (European ancestry (Eu	anges in expr a with when i closure gene an or Non-Eu or Non-Eurol iropean or N	ession. t was absent) " has a <i>p</i> valu rropean) with pean) plus an on-European	. Regression e < 0.10. Nei out an inter i interaction ) plus an in	coefficients gative regres action term term betwe teraction ter	were calcula ision coeffici between th en the SNP a m between	ted for each of ted for each of ents are in itale of SNP allele allele and gen dilele and gen the SNP allel	22B SNPs were associated with changes in expression. AP2B allele was present (compared with when it was absent). Regression coefficients were calculated for each of the 49 "ductus closure if the association with the "ductus closure gene" has a $p$ value < 0.10. Negative regression coefficients are in italics. * $p < 0.05$ ; ** $p < 0.01$ . age and genetic ancestry (European or Non-European) without an interaction term between the SNP allele and genetic ancestry. ge and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; regression estry (european or Non-European) plus an interaction term between the SNP allele and genetic ancestry; not ge and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; stry. onal age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and age and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and genetic ancestry; and genetic ancestry (European or Non-European) plus an interaction term between the SNP allele and gene	closure > < 0.01. stry. jression ncestry;

**Table 3.** Multivariable regression models examining the independent effects of the *PTGIS* SNP haplotype rs493694(G)/rs693649(A) (associated with early ductus closure) on the RNA expression of "ductus closure genes" (n = 273).

Genes/aliases	Regression coefficien	ts for PTGIS
	Haplotype: rs493694(	(G)/rs693649(A)
	General population <sup>a</sup>	European ancestry <sup>b</sup>
Ca <sup>2+</sup> signaling		
CACNB2/Cavbeta2		-0.469*
K <sup>+</sup> channels		
KCNS3/Kv9.3		-2.454*
Contractile proteins		
MYH11/SM1		-0.636
Endothelin signaling		
Prostaglandin signaling		
CYP8A1/PTGIS		-0.461**
Nitric oxide signaling		
NOS3/eNOS		-0.618*
Inflammation and remodeling		
BMP9		-2.22*
BMP10	-0.513	-1.36*
Regression coefficient represe PTG/S SNP haplotype was pro- Regression coefficients were of genes" listed in Table 1. Regre the association with the "ductu	esent (compared with v calculated for each of th ssion coefficients are onl	when it was absent). ne 49 "ductus closure ly listed in the table if

genes' listed in Table 1. Regression coefficients are only listed in the table if the association with the "ductus closure gene" has a *p* value < 0.10. Negative regression coefficients are in italics. \**p* < 0.05; \*\**p* < 0.01. <sup>a</sup>General population: multivariate analyses were adjusted for gestational age and genetic ancestry (European or Non-European) without an interaction term between the haplotype and genetic ancestry. <sup>b</sup>European ancestry: multivariate analyses were adjusted for gestational age

and genetic ancestry (European or Non-European) plus an interaction term between the haplotype and genetic ancestry; regression coefficients were obtained for the referent value: European genetic ancestry.

samples with Non-European ancestry (Table 2—Non-European ancestry). Nor were the same changes observed when the two *TFAP2B* polymorphisms that are unrelated to the timing of DA closure (rs2817419 (G allele) and rs2635727 (T allele)) were examined in samples with European ancestry (Table 2—European ancestry/*TFAP2B* (Non-PDA-associated polymorphisms)).

A similar phenomenon occurred when we tested whether an interaction occurred between the fetus's genetic ancestry and the 2-SNP haplotype of *PTGIS* that is negatively associated with PDA (rs493694 (G allele)/rs693649 (A allele)). When the *PTGIS* haplotype was present in samples with European ancestry, the haplotype was associated with changes in RNA expression in several "DA closure genes" (the most significant change occurring in *PTGIS* itself) (Table 3).

# DISCUSSION

Premature infants born to mothers who self-identify as White/ European ancestry are less likely to close their PDA following prostaglandin inhibition than infants born to mothers who selfidentify as Non-White/Non-European ancestry.<sup>1-4</sup> This difference does not appear to be due to different rates of indomethacin/ ibuprofen metabolism or different serum prostaglandin E2 concentrations.<sup>1-4</sup> Our current study demonstrates that genetic ancestry is associated with changes in the expression of several

"DA closure genes". This occurs through a direct association between genetic ancestry and a limited number of "DA closure genes" (*SLCO2A1* (the prostaglandin transporter) and *PTGS2* (cyclooxygenase 2)) (Table 1), as well as through a broader, indirect, interactive effect, where genetic ancestry modifies the associations between common genetic polymorphisms and DA gene expression.

We previously identified several polymorphisms in the genes *PTGIS* and *TFAP2B* that were associated with different rates of PDA closure in a population composed primarily of preterm infants with European genetic ancestry.<sup>10</sup> These associations were not replicated by other investigators using populations with different or more diverse genetic origins.<sup>14,15</sup> In line with these discordant observations, our current study found consistent associations between *PTGIS* and *TFAP2B* polymorphisms and the expression of "DA closure genes" in DA with European genetic ancestry. On the other hand, no consistent positive or negative associations could be found in our genetically diverse DA population unless an interaction between the polymorphisms and genetic ancestry was taken into account (Tables 2 and 3).

In DA with European genetic ancestry, the *PTGIS* haplotype (rs493694 (G allele)/rs693649 (A allele)), which is associated with early DA closure, was associated with decreased expression of *PTGIS* itself as well as *NOS3* (endothelial nitric oxide synthase, which regulates nitric oxide production) and several other calcium and potassium regulatory genes (Table 3).

Consistent alterations in gene expression were also found when each of the four *TFAP2B* SNPs (that are associated with persistent PDA) were present in DA with European genetic ancestry. These changes include decreased expression of calcium and potassium signaling genes, as well as decreased expression of genes regulating endothelin and HIF2 alpha (Table 2). It is interesting to note that similar changes in endothelin and HIF2 alpha were previously found in newborn mice with targeted deletions of Tfap2b (the mouse equivalent of *TFAP2B*).<sup>12</sup>

To determine whether the changes in DA gene expression were specific for the *TFAP2B* SNPs associated with persistent PDA, we examined two other *TFAP2B* polymorphisms, rs2817419(G) and rs2635727(T), which are unrelated to the incidence of preterm PDA (Table 2). Neither polymorphism was associated with the changes in gene expression described above (Table 2).

Our study has several limitations. The tissues were from pregnancy terminations, which may have altered the gene expression in the DA before tissue processing. We explored a limited number of candidate genes and may have missed others that might have been detected by genome-wide association studies or pathway-based analyses. There was also a relatively small number of tissue samples and a low proportion of European genetic ancestry in our study population which may have limited our ability to identify smaller effects in the "DA closure genes" we studied.

Since our investigation was an exploratory study, we chose to consider results with a p value < 0.1 as possible evidence of association. Although applying a more stringent p value would have reduced the chance of finding false-positive signals, it might have eliminated our ability to detect true positive signals, especially when the genetic effects are small. Our finding that at least three of the four *TFAP2B* SNPs, that were associated with persistent PDA, also were associated with the same changes in expression of several of the "DA closure genes" (*EPAS1, CACNB2, ECE1, KCNA2, ATP2A3, EDNRA, EDNRB, BMP9*, and *BMP10*) increases the confidence that these may actually represent true positive results. None of these changes were seen when the two *TFAP2B* polymorphisms that were unrelated to the timing of DA closure were examined in samples with European genetic ancestry (Table 2).

As an observational study, we cannot distinguish between causation and association. Nor do we know if the changes in gene expression have a direct effect on DA closure, or if they are merely

an indirect effect of other events that are responsible for its closure. However, our findings do provide biologic plausibility to the concept that the PTGIS and TFAP2B SNPs are either functional polymorphisms or in tight association with functional polymorphisms that play an active role in regulating DA closure. Since the SNPs we studied are present in haplotype blocks, the actual genetic variations responsible for the associated changes in gene expression could lie anywhere within that block. We speculate that the increased rate of DA closure associated with the PTGIS 2-SNP haplotype rs493694(G)/rs693649(A) may be due to the associated decrease in prostaglandin I2 synthase expression (and a subsequent decrease in the potent vasodilator, PGI2). On the other hand, we have no similar explanation for the changes associated with the TFAP2B SNPs since none of the SNPs appear to alter TFAP2B mRNA levels (Table 2). It is worth noting that the TFAP2B SNPs we examined are situated in unique, highly conserved regions, that are located between exons, and in proximity to a number of putative transcription factor-binding sites (Fig. 1). SNPs in or near a gene can affect both the amount and function of the mRNA or protein produced. We speculate that alterations in these unique, highly conserved, noncoding regions might alter TFAP2B splicing such that transcript levels are normal but the transcripts themselves are abnormal; or, they may have distant effects (possibly through altered transcription factor binding or microRNA production) on gene expression even beyond the TFAP2B gene in which they are located. These findings are consistent with our current understanding that many disease-associated common variants are noncoding and are enriched in DNA regulatory elements.<sup>23</sup> Future studies will be needed to determine how these polymorphisms affect the expression of downstream genes.

In conclusion, we found no consistent associations between the presence of polymorphisms in *PTGIS* and *TFAP2B* and the expression of "DA closure genes" unless an interaction between the polymorphisms and genetic ancestry was taken into account. When an interaction between the polymorphisms and ancestry was accounted for, the *PTGIS* and *TFAP2B* polymorphisms were associated with consistent changes in DA gene expression in DA from fetuses with European genetic ancestry.

### DATA AVAILABILITY

The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

#### ACKNOWLEDGEMENTS

This study is dedicated to the memory of our coinvestigator, Dr. Nahid Waleh, who helped conceptualize and design the original study and who meticulously performed all of the RNA analyses. We were not able to list her as an author because her untimely death, prior to the preparation of the manuscript, violated the journal's policy for authorship. We are very grateful to Dr. Eleanor Drey, Janette Alvarez, and all the nursing and counseling personnel at the Women's Option Center at San Francisco General Hospital for their help in enabling our tissue collection. Similarly, we thank Anne Marie Barrette, Hart Horneman, and Christine Roman for their skillful assistance with the sample collection. We also thank Drs. Bruce Gelb and Deepak Srivastava for their helpful insights and suggestions regarding our findings. This work was supported by the National Heart, Lung, and Blood Institute (HL109199) and a gift from the Jamie and Bobby Gates Foundation.

#### AUTHOR CONTRIBUTIONS

The following authors have (1) made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; (2) drafted the article or revised it critically for important intellectual content; and (3) have given final approval of the version to be published: R.I.C., N.K.H., J.M.D., J.C.M., and K.K.

# ADDITIONAL INFORMATION

Competing interests: The authors declare no competing interests.

Patient consent: Patient consent was not required because this study used deidentified data.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## REFERENCES

- Chorne, N., Jegatheesan, P., Lin, E., Shi, R. & Clyman, R. I. Risk factors for persistent ductus arteriosus patency during indomethacin treatment. J. Pediatr. 151, 629–634 (2007).
- Durrmeyer, X. et al. Are cytochrome P450 CYP2C8 and CYP2C9 polymorphisms associated with ibuprofen response in very preterm infants? *PLoS ONE* 5, e12329 (2010).
- Furzan, J. A., Reisch, J., Tyson, J. E., Laird, P. & Rosenfeld, C. R. Incidence and risk factors for symptomatic patent ductus arteriosus among inborn very-low-birthweight infants. *Early Hum. Dev.* **12**, 39–48 (1985).
- Cotton, R. B., Haywood, J. L. & FitzGerald, G. A. Symptomatic patent ductus arteriosus following prophylactic indomethacin. A clinical and biochemical appraisal. *Biol. Neonate* **60**, 273–282 (1991).
- Shelton, E. L. et al. Effects of antenatal betamethasone on preterm human and mouse ductus arteriosus: comparison with baboon data. *Pediatr. Res.* 84, 458–465 (2018).
- Waleh, N. et al. Patterns of gene expression in the ductus arteriosus are related to environmental and genetic risk factors for persistent ductus patency. *Pediatr. Res* 68, 292–297 (2010).
- 7. Waleh, N. et al. Effects of advancing gestation and non-caucasian race on ductus arteriosus gene expression. J. Pediatr. 167, 1033–1041 (2015).
- Bhandari, V. et al. Genetic contribution to patent ductus arteriosus in the premature newborn. *Pediatrics* 123, 669–673 (2009).
- Lavoie, P. M., Pham, C. & Jang, K. L. Heritability of bronchopulmonary dysplasia, defined according to the consensus statement of the National Institutes of Health. *Pediatrics* 122, 479–485 (2008).
- 10. Dagle, J. M. et al. Determination of genetic predisposition to patent ductus arteriosus in preterm infants. *Pediatrics* **123**, 1116–1123 (2009).
- Majed, B. H. & Khalil, R. A. Molecular mechanisms regulating the vascular prostacyclin pathways and their adaptation during pregnancy and in the newborn. *Pharmacol. Rev.* 64, 540–582 (2012).
- 12. Ivey, K. N. et al. Transcriptional regulation during development of the ductus arteriosus. *Circ. Res.* **103**, 388–395 (2008).

- Zhao, F. et al. Novel TFAP2B mutations that cause Char syndrome provide a genotype-phenotype correlation. Am. J. Hum. Genet. 69, 695–703 (2001).
- 14. Kawase, K. et al. Single nucleotide polymorphisms in AGTR1, TFAP2B, and TRAF1 are not associated with the incidence of patent ductus arteriosus in Japanese preterm infants. *Pediatr. Int.* **58**, 461–466 (2016).
- Dagle, J. M. et al. Genetic variants associated with patent ductus arteriosus in extremely preterm infants. J. Perinatol. 39, 401–408 (2019).
- Merz, E., Oberstein, A. & Wellek, S. Age-related reference ranges for fetal foot length. Ultraschall Med. 21, 79–85 (2000).
- Bouayad, A. et al. Characterization of PGE2 receptors in fetal and newborn lamb ductus arteriosus. Am. J. Physiol. Heart Circ. Physiol. 280, H2342–H2349 (2001).
- Waleh, N. et al. Prostaglandin E2-mediated relaxation of the ductus arteriosus: effects of gestational age on g protein-coupled receptor expression, signaling, and vasomotor control. *Circulation* **110**, 2326–2332 (2004).
- Ehn, N. L. et al. Evaluation of fetal and maternal genetic variation in the progesterone receptor gene for contributions to preterm birth. *Pediatr. Res.* 62, 630–635 (2007).
- Soejima, M. & Koda, Y. Population differences of two coding SNPs in pigmentationrelated genes SLC24A5 and SLC45A2. Int J. Leg. Med. 121, 36–39 (2007).
- 21. Beleza, S. et al. Genetic architecture of skin and eye color in an African-European admixed population. *PLoS Genet.* **9**, e1003372 (2013).
- Reiner, A. P. et al. Population structure, admixture, and aging-related phenotypes in African American adults: the Cardiovascular Health Study. *Am. J. Hum. Genet.* 76, 463–477 (2005).
- Maurano, M. T. et al. Systematic localization of common disease-associated variation in regulatory DNA. Science 337, 1190–1195 (2012).
- Siepel, A. et al. Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.* 15, 1034–1050 (2005).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons. org/licenses/by/4.0/.

© The Author(s) 2021