Epistatic interactions in idiopathic pulmonary arterial hypertension

Shivani Vadapalli, Satyanarayana M. L., Chaitra K. L., Surekh Rani H., Sastry B.K.S.[1], Pratibha Nallari

Department of Genetics, Osmania University, Hyderabad,
[1]Department of Cardiology, CARE Hospitals, Hyderabad, Andhra Pradesh, India.

BACKGROUND: Idiopathic pulmonary arterial hypertension (IPAH) is a poorly understood complex disorder, which results in progressive remodeling of the pulmonary artery that ultimately leads to right ventricular failure. A two-hit hypothesis has been implicated in pathogenesis of IPAH, according to which the vascular abnormalities characteristic of PAH are triggered by the accumulation of genetic and/or environmental insults in an already existing genetic background. The multifactor dimensionality reduction (MDR) analysis is a statistical method used to identify gene–gene interaction or epistasis and gene–environment interactions that are associated with a particular disease. The MDR method collapses high-dimensional genetic data into a single dimension, thus permitting interactions to be detected in relatively small sample sizes.
AIM: To identify and characterize polymorphisms/genes that increases the susceptibility to IPAH using MDR analysis.
MATERIALS AND METHODS: A total of 77 IPAH patients and 100 controls were genotyped for eight polymorphisms of five genes (*5HTT*, *EDN1*, *NOS3*, *ALK-1*, and *PPAR-γ2*). MDR method was adopted to determine gene–gene interactions that increase the risk of IPAH.
RESULTS: With MDR method, the single-locus model of 5HTT (L/S) polymorphism and the combination of *5HTT*(L/S), *EDN1*(K198N), and *NOS3*(G894T) polymorphisms in the three-locus model were attributed to be the best models for predicting susceptibility to IPAH, with a *P* value of 0.05.
CONCLUSION: MDR method can be useful in understanding the role of epistatic and gene–environmental interactions in pathogenesis of IPAH.

Introduction

Idiopathic pulmonary arterial hypertension (IPAH) is a debilitating disorder characterized by progressive narrowing of the pulmonary arterioles that leads to increased resistance to the flow of blood from right ventricle to the lungs and eventually results in death due to right ventricular failure. The disease is clinically defined as sustained elevation of the pulmonary artery pressures of >25 mmHg at rest or 30 mmHg during exercise, in absence of any underlying cause. IPAH is a rare and fatal disease with extremely poor prognosis that is more commonly seen in women than men.[1]

The pathology of pulmonary hypertension includes endothelial, smooth-muscle, and/or adventitial abnormalities, which result in obliterative remodeling of the pulmonary circulation. It is characterized by vasoconstriction, occlusion of the lumen in medium-sized and small pulmonary arteries due to excessive cellular proliferation in the vascular wall, and *in situ* thrombosis, with loss of microvessels and capillaries.[2]

Mutations in two receptors in the TGF-β superfamily: *BMPR2* and *ALK-1* have been recognized as the cause in most familial PAH cases and 10%–40% of sporadic cases. In addition, other signaling systems have also been found to participate in PAH, including K channels,

Address for correspondence: Prof. Pratibha Nallari, Department of Genetics, Osmania University, Hyderabad – 500 004, Andhra Pradesh, India. E-mail: prathinallari@yahoo.com

endothelial mediators, serotonin, angiopoietin, and cyclooxygenases.[3] Current theories on pathogenesis focus on abnormalities in interaction between endothelial and smooth-muscle cells coupled with imbalanced activation of other TGF- receptors due to reduced activity of mutated *BMPR2/ALK1*. Endothelial-cell injury may result in an imbalance in endothelium-derived mediators, favoring vasoconstriction, due to elevated levels of endothelin-1, thromboxane, and serotonin accompanied with reduced levels of nitric oxide and prostacyclin. Defects in ion-channel activity in smooth-muscle cells in the pulmonary artery may contribute to vasoconstriction and vascular proliferation.

However, it has been postulated that presence of a mutation in *BMPR2/ALK1* is not solely responsible in the causation of IPAH, rather, 'multiple hits' are required for the onset of diseased phenotype. This difference of susceptibility may result from mutations in other genes, inherited polymorphisms in various genes, environmental factors, or acquired factors. Thus, identification of various other genes and their possible gene interactions is warranted to understand the complexity of this disease.[4]

Multifactor dimensionality reduction (MDR) is a novel and powerful statistical tool for detecting and modeling epistasis. MDR is a data reduction method for detecting multilocus genotype combinations that predict disease risk for common, complex disease.[5,6] MDR pools genotypes into "high-risk" and "low-risk" groups to reduce multidimensional data into one dimension. Based on MDR analysis, many studies have observed that complex interactions among multiple genes may contribute genetically to complex disorders.[7,8]

The pathogenesis of IPAH suggests the complexity of the disease, wherein multiple genetic and environmental factors are more likely to be involved in its etiology. Based on this hypothesis, MDR analysis was carried out to detect multiple loci interaction that increase the risk IPAH and this may include epistatic interactions, with the combined effect being greater (or less) than that expected by multiplying their individual main effects.

A total of eight SNPs for five genes involved in endothelial dysfunction and smooth muscle cell proliferation were selected for their possible role in IPAH. These genes/pathways were selected for analysis based on their given role in pathogenesis of IPAH. This includes the promoter polymorphism in the serotonin transporter (*5HTT*-L/S), two endothelin-1 polymorphisms (*EDN1*- +138 Ins A, K198N), three polymorphism of nitric oxide synthase [*NOS3* – T-786C, Intron 4(a/b), and G894T], one polymorphism in the intron 3 of *ALK-1* (c.313+11 C>T), and P12A polymorphism of peroxisome proliferators activated receptor – gamma 2 (*PPARγ2*-P12A).

Materials and Methods

The study was approved by the ethics committee of Care Hospitals, Hyderabad. All the patients included in the study were confirmed IPAH cases, referred by the cardiologist. The study included 77 PAH patients (73 IPAH and 5 FPAH). Randomly selected 100 healthy subjects without history of cardiac and systemic disorders were included as controls.

Genotyping

The eight polymorphisms included in the study were examined by polymerase chain reaction (PCR), restriction fragment length polymorphism (RFLP) analysis, single strand conformation polymorphism (SSCP) analysis, and agarose and polyacrylamide gel electrophoresis. The type of method adopted for each SNP is given in Table 1, along with the nucleotide position.

Table 1: Polymorphisms included in multifactor dimensionality reduction analyses of idiopathic pulmonary arterial hypertension in a case-control study

Gene	Polymorphism typing method	Nucleotide position	Variation/ codon	Minor allele	Minor allele frequency	
					Cases	Controls
NOS3	RFLP	-786	T/C	C	0.27	0.25
	Genotyping	Intron 4	4a/4b	4a	0.175	0.195
	RFLP	894	G/T (298)	T	0.27	0.28
EDN1	SSCP	+138	3A/4A	4A		
	SSCP	5665	G/T (198)	T	0.357	0.20
ALK-1	SSCP	Intron 3 c.313+11	C/T	T	0.28	0.24
5HTT	Genotyping	5'UTR	L/S	L	0.538	0.335
PPARγ2	RFLP	Exon B	C/G (12)	G	0.071	0.175

Statistical analysis

Potential locus–locus interactions were performed by using the nonparametric MDR software (*version 2.0*) with eight SNPs. The MDR analysis redefines a variable of high and low risk by combining information on several loci that may interact in the disease etiology. The fitness of an MDR model was assessed by estimating the testing accuracy. Models that were true positive would have estimated testing accuracy of $P \leq 0.0.5$. The cross-validation consistency was a measure of the number of times of 10 divisions of the data that MDR found in the same best model.

Interaction dendrogram was also employed. The attributes (i.e., SNPs) that are strongly interacting appear close together at the leaves of the tree, whereas those that do not interact appeared far from one another. The colors used in the dendrogram comprise a spectrum of colors representing a continuum from synergy to redundancy. The colors range from red representing a high degree of synergy, orange a lesser degree, and gold the midway point between synergy and redundancy. On the redundancy end of the spectrum, the highest degree was represented by the blue color with a lesser degree represented by green.

Results

The observed allelic frequency of the eight SNPs in both the cases and controls in the present study is given in Table 1. The odds test of association for each SNP with IPAH is given in Table 2. We have previously reported the role of *EDN1* +138 Ins A, K198N, and *NOS3* polymorphisms (T-786C and G894T) in IPAH.[9] When analyzed independently, the odds ratio for 4 SNPs - *5HTT* (LL/LS vs. SS), *EDN1* (3A/4A+4A/4A vs. 3A/3A: N198N/K198N vs. K198K), and *PPAR*γ2 (CC/CG vs. GG) were observed to be significantly associated with IPAH [Table 2].

On MDR analysis, in the one factor model, *5HTT* (SNP7) was the best attribute in the prediction of risk to IPAH having a testing accuracy of 73.16% and cross-validation consistency (CVC) of 10/10 [Table 3]. Overall among the eight predicted models, the best interaction model was the three factor model, i.e., the combination of *NOS3*–G894T polymorphism, *EDN1* K198N, and *5HTT* polymorphism (SNP3, SNP5, and SNP7), with an improved testing accuracy of 73.36% and a CVC of 9/10 that was found to be statistically significant ($p \leq 0.05$), as determined empirically by the permutation testing [Table 3]. All models including four or more factors had a decrease in testing accuracy and CVC. While the combination of all eight SNP's showed a testing accuracy of 62.21% and a perfect CVC of 10/10, the

Table 2: Association between genotypes of polymorphisms included in the present study and the risk of idiopathic pulmonary arterial hypertension

Gene/ polymorphism	Genotypes	Odds ratio (95% CI)	P
5HTT (L/S)	LL vs. SS	3.125 * (1.39-6.33)	<0.0001
	LS vs. SS	4.80* (2.24-10.29)	
EDN1			
3A/4A	4A/4A vs. 3A/3A	2.95 (0.26-33.26)	
	3A/4A vs. 3A/3A	3.44* (1.26-9.44)	0.031*
	NN vs. KK	3.38* (1.22-9.36)	
K198N	KN vs. KK	2.35* (1.21-4.57)	0.0076
NOS3			
T-786C	CC vs. TT	0.71 (0.12-4.08)	
	CT vs. TT	1.25 (0.68-2.30)	0.67
Intron 4 (4a/4b)	4a/4a vs. 4b/4b	1.25 (0.07-19.92)	
	4a/4b vs. 4b/4b	0.80 (0.43-1.5)	0.43
G894T	TT vs. GG	0.66 (0.19-2.30)	
	GT vs. GG	1.53 (0.82-2.87)	0.25
ALK-1			
c.313+11 C>T	TT vs. CC	1.18 (0.385-3.59)	0.9
	CT vs. CC	1.423 (0.70-2.89)	0.3
PPAR 2			
P12A	A12A + P12A vs. P12P	0.37 (0.17-0.79)	0.0072*
	P12P vs. A12A + P12A	2.35* (1.08-5.11)	0.043*

*$P \leq 0.05$ is considered significant

Table 3: Gene-to-gene interaction in determining the risk to idiopathic pulmonary arterial hypertension

Genes included in best combination in each model	CVC	Training accuracy	Testing accuracy	P
SNP7	10/10	0.7316	0.7316	0.05*
SNP5, SNP7	6/10	0.7486	0.6636	0.2
SNP3, SNP5, SNP7	9/10	0.7893	0.7336	0.05*
SNP1, SNP3, SNP5,SNP7	8/10	0.8297	0.6886	0.11
SNP1, SNP3, SNP5, SNP7, SNP8	7/10	0.8697	0.6321	0.26
SNP1, SNP3, SNP4, SNP5, SNP7, SNP8	4/10	0.8938	0.6316	0.3
SNP1, SNP2, SNP3, SNP4, SNP5, SNP7, SNP8	7/10	0.9135	0.6276	0.2
SNP1, SNP2, SNP3, SNP4, SNP5, SNP6, SNP7, SNP8	10/10	0.9245	0.6221	0.2

SNP1 – T-786C, SNP2- 4a/4b, SNP3- G894T, SNP4- +138A, SNP5- K198N, SNP6- P12A , SNP7- 5HTT, SNP8- ALK-1 (c.313+11 C>T); CVC -Cross-validation consistency. *$P \leq 0.05$ is considered significant.

interaction was not found to be statistically significant (*p* - 0.2) [Table 3].

A graphical depictions of the combined effect of the three-locus model of *NOS3* G894T, *EDN1* K198N, and *5HTT* as high- and low-risk groups and statistical interactions determined by MDR are shown [Figure 1a–c]. The darker-shaded cells show higher-risk combinations, whereas lighter-shaded cells show combinations not associated with elevated risk to IPAH. No shading or white cells represent genotype combinations for which no data were observed. The case–control ratio known as the threshold ratio in the present study was set at 0.77 (77/100). Bars represent hypothetical distributions of cases (left) and controls (right) with each multifactor combination. Cells were labeled as high risk if the ratio of the percentage of cases to controls met or exceeded the threshold of 0.77 and cells were labeled as low risk if the threshold has not exceeded 0.77.

As per the three-locus model, the individuals homozygous for the wild-type allele in all the three polymorphisms, i.e., combination of *5HTT* SS, *EDN1* K198K, and *NOS3* G894G, were placed in the low-risk group, while the combinations of *5HTT* LL genotype with either heterozygotes for K198N or G894T polymorphism were placed in high-risk group. Individuals heterozygous for all the three polymorphisms (*5HTT* LS, *EDN1* K198N, and *NOS3* G894T) were placed in the highest risk group. Individuals heterozygous for at least two of the three polymorphisms belonged to the high-risk group [Figure 1a–c].

The dendrograms provided by MDR were examined to assist in the visualization and interpretation of potential interactions.[7] As observed in the dendrogram, the SNP5, SNP7, and SNP8 belonged to one cluster, while SNP4, SNP6, SNP3, and SNP1 belonged to another cluster. The two clusters showed synergistic interaction, while SNP2 showed a synergistic association with both the clusters, in predicting susceptibility to IPAH. However, every other combination of the interaction provided redundant information [Figure 2]. Redundancy refers to the situation in which the entropy-based interaction between two SNPs provides less information than the entropy-based correlation between the pair.

Discussion

Based on the MDR method, the *5HTT* marker was observed to be the best single-locus model, with a "p" value of 0.05, suggesting that the power of the marker in predicting susceptibility to IPAH is very strong. When tested independently also, the *5HTT* gene was found to be associated with IPAH. The role of *5HTT* promoter polymorphism has been implicated in pathogenesis of IPAH by a number of studies.[10,11]

Among the eight models with different loci tested (eight SNPs of five candidate genes), the best gene–gene interaction model identified was a three-locus model including the *5HTT* L/S promoter polymorphism, the *EDN-1* K198N polymorphism and *NOS3* G894T polymorphism. Individuals heterozygous in at least two out of three loci
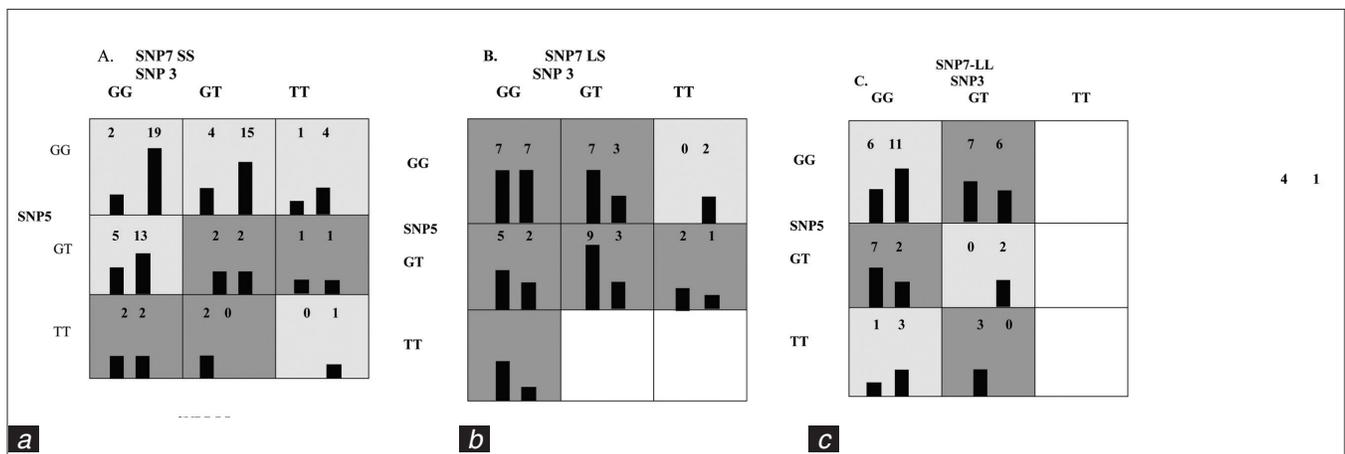
Figure 1: (a–c) The best three-locus SNP model selected by MDR. interaction of NOS3 G894T, 5HTT, and EDN1 K198N polymorphisms
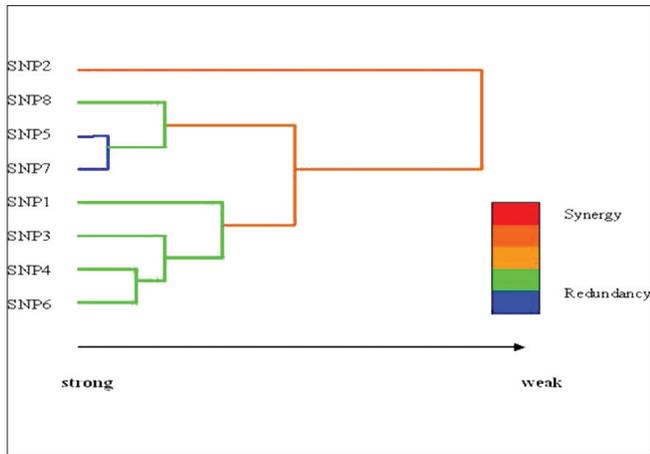
Figure 2: The dendrogram demonstrates the nature of the interactions between SNPs. The colors used in the dendrogram comprise of a spectrum of colors representing a continuum from synergy to redundancy. Red denotes a high degree of interaction, orange denotes a lesser degree of interaction, and green denotes a weak interaction. Blue lines indicate no interaction or redundancy

were clustered in the high-risk group. In this model, the combination of the *5HTT* SS homozygotes, *EDN-1* K/K homozygotes, and *NOS3* exon7 GG homozygotes were associated with a reduced risk to IPAH. The G894T *NOS3* polymorphism is associated with reduced basal NO production.[12] As reported previously, the G894T polymorphism when studied independently did not show any association with IPAH.[9] In MDR, by contrast, the *NOS3* G894T polymorphism, in combined effect with other two variants, was associated with increased risk for IPAH, which implies a genetic interaction between these genes in the pathogenesis of the disease. Further, dendrogram also revealed a synergistic interaction of the *NOS3* 27-bp VNTR Intron 4 polymorphism with all the SNPs studied, indicating the important epistatic interactions of *NOS3* in pathogenesis of IPAH.

The three genes, *5HTT, EDN1*, and *NOS3*, play a crucial role in pathogenesis of IPAH. Serotonin [5-hydroxytryptamine (5-HT)] alters vascular tone and promotes vascular smooth muscle cell growth. The serotonin transporter (5HTT) acts by uptake of free 5-HT. Serotonin activates the Ras or Rac or both pathways that activate NADP(H) oxidase producing a reactive oxygen species (ROS), which activates ERK and/or MAP kinase pathways, inducing hyperplasia. The L variant of the *5HTT* gene polymorphism is associated with 5HTT over

expression and was found to be more common in patients with IPAH.[10] The pulmonary artery smooth muscle cells (PA-SMCs) with LL genotype of *5HTT* from IPAH patients have been shown to proliferate faster than PA-SMCs derived from with the LS or SS genotypes.

Endothelin-1 (EDN1) and nitric oxide are important endothelial mediators having opposing action. Imbalances in the expression of endothelium-derived vasoactive substances are thought to contribute to the pathogenesis of IPAH. Plasma levels of EDN1 are increased in IPAH, which is known to cause profound and persistent pulmonary vasoconstriction.[13,14] Elevated levels of vasoconstrictors are accompanied with reduced levels of two important vasodilators NO and PGI2 in IPAH. The exact function of the K198N polymorphism in expression of *EDN1* is not known, though it has been associated with many diseases. The K198N polymorphism was also found to be associated with increased risk to IPAH, in our previous study.[9]

One possible pathway linking these three genes is the ROS pathway. Enhanced production of ROS, especially superoxide ($\bullet O_2$), is also known to decrease NO bioavailability.[15] The uptake of 5HT by the 5HTT transporter and its subsequent oxidative deamination by monoamine oxidase A (MAO-A) and breakdown by NADPH oxidase in PASMC also leads to generation of ROS. Hypoxia is recognized as a trigger in onset and progression of IPAH. Both hypoxia and ROS induce the expression of hypoxia inducible factor (HIF1α), which in turn can induce *EDN1* expression leading to further elevated EDN1 levels and pulmonary vascular resistance. Changes in ROS and nitric oxide levels upon hypoxia can result in oxidative damage and subsequent apoptosis of endothelium and vascular smooth muscle cells.[16]

It is also observed that all the SNP's included in the study have weak synergistic effect with each other. This further corroborates that the pathogenesis of IPAH involves an interplay of a variety of susceptibility alleles and external stimuli, which, in turn, may influence disease severity and natural history. Thus, more genes need to be analyzed/identified to clarify the nature of interactions among the SNPs/genes studied and in better understanding of the pathway implicated. Individuals may

also differ in their susceptibility to environmental risk factors and hence role of gene–environment interaction in IPAH also needs to be identified. However, as different populations have distinct genetic backgrounds, it is necessary to validate or replicate such associations with independent samples collected especially from other ethnic groups or populations.

Conclusion

Thus, using MDR analysis, the present study gives an insight into the possible epistatic interactions between the *5HTT*(L/S), *EDN1*(K198N), and *NOS3*(G894T) polymorphisms/genes in susceptibility to IPAH is established. Studies have suggested that gene interaction is not only possible but is probably ubiquitous in determining susceptibility to complex human diseases.[8] Further studies on the epistatic interactions are therefore warranted to elucidate their possible underlying role in the pathogenesis and delineation of pathways leading to IPAH.

Acknowledgement

References

1. Rich S, Dantzker DR, Ayres SM, Bergofsky EH, Brundage BH, Detre KM, *et al*. Primary pulmonary hypertension. A national prospective study. Ann Intern Med 1987;107: 216-23.
2. Rich, S (ed). Executive summary from the World Symposium on Primary Pulmonary Hypertension, Evian, France, September 6-10, 1998, co-sponsored by The World Health Organization.
3. Newman JH, Phillips JA 3rd, Loyd JE. Narrative review: The enigma of pulmonary arterial hypertension: New insights from genetic studies. Ann Intern Med 2008;148:278-83
4. Machado RD, James V, Southwood M, Harrison RE, Atkinson C, Stewart S, *et al*. Investigation of second genetic hits at the *BMPR2* locus as a modulator of disease progression in familial pulmonary arterial hypertension. Circulation 2005;111:607-13.
5. Ritchie MD, Hahn LW, Roodi N, Bailey LR, Dupont WD, Parl FF, *et al*. Multifactor-dimensionality reduction reveals high-order interactions among estrogen-metabolism genes in sporadic breast cancer. Am J Hum Genet 2001;69:138-47.
6. Hahn LW, Ritchie MD, Moore JH. Multifactor dimensionality reduction software for detecting gene-gene and gene-environment interactions. Bioinformatics 2003;19:376-82.
7. Moore JH. A global view of epistasis. Nat Genet 2005;37:13-4.
8. Thornton-Wells TA, Moore JH, Haines JL. Genetics, statistics and human disease: Analytical retooling for complexity. Trends Genet 2004;20:640-7.
9. Shivani Vadapalli, H.Surekha Rani, B.K.S. Sastry, Pratibha Nallari. Endothelin-1 and endothelial nitric oxide polymorphisms in idiopathic pulmonary arterial hypertension. Int J Mol Epidemiol Genet 2010;1:208-13.
10. Eddahibi S, Humbert M, Fadel E, Raffestin B, Darmon M, Capron F, *et al*. Serotonin transporter overexpression is responsible for pulmonary artery smooth muscle hyperplasia in primary pulmonary hypertension. J Clin Invest 2001;108:1141-50.
11. Marcos E, Fadel E, Sanchez O, Humbert M, Dartevelle P, Simonneau G, *et al*. Serotonin-induced smooth muscle hyperplasia in various forms of human pulmonary hypertension. Circ Res 2004;94:1263-70.
12. Veldman BA, Spiering W, Doevendans PA, Vervoort G, Kroon AA, de Leeuw PW, *et al*. The Glu298Asp polymorphism of the *NOS3* gene as a determinant of the baseline production of nitric oxide. J Hypertens 2002;20:2023-7.
13. Stewart DJ, Levy RD, Cernacek P, Langleben D. Increased plasma endothelin-1 in pulmonary hypertension: Marker or mediator of disease? Ann Intern Med 1991;114:464-9.
14. Sujana K, Sastry BK, Nallari P. Endothelin 1: A vasoconstrictor in primary pulmonary hypertension. J Cell Tissue Res 2007;7:867-70.
15. Harrison DG. Cellular and molecular mechanisms of endothelial cell dysfunction. J Clin Invest 1997;100: 2153-7.
16. Esteve JM, Launay JM, Kellermann O, Maroteaux L. Functions of serotonin in hypoxic pulmonary vascular remodeling. Cell Biochem Biophys 2007;47:33-43.