Journal of Human Genetics (2010) 55, 13–17

www.nature.com/jhg

ORIGINAL ARTICLE

Common genetic variants in pre-microRNAs are associated with risk of coal workers' pneumoconiosis

Meilin Wang[1,3], Yang Ye[1,3], Haiyang Qian[1], Zhifang Song[2], Xiaomin Jia[2], Zhengdong Zhang[1], Jianwei Zhou[1] and Chunhui Ni[1]

microRNAs (miRNAs) are an abundant class of small noncoding RNA molecules thought to be involved in biological functions, including embryonic development, chromosome architecture, cell proliferation and apoptosis. We hypothesized that common variants in the miRNAs are associated with risk of coal workers' pneumoconiosis (CWP). In a case–control study of 496 CWP patients and 513 control subjects frequency matched by exposure years and work types, we genotyped four single-nucleotide polymorphisms (SNPs) (rs2910164, rs2292832, rs11614913 and rs3746444) in pre-miRNAs (miR-146a, miR-149, miR-196a2 and miR-499) and assessed the associations with risk of CWP. A significantly increased risk of CWP was found for the miR-149 rs2292832 TT genotype (odds ratio (OR), 1.31; 95% confidence interval (CI), 1.01–1.69), compared with the CT/CC genotypes, and this increased risk was evident among subgroups of those aged ⩾68 years (OR=1.52, 95% CI=1.03–2.25), dust exposure ⩾26 years (OR=1.42, 95% CI=1.04–1.93) and ever smokers (OR=1.48, 95% CI=1.00–2.20). Furthermore, a significant association was observed between the genotypes and patients with stages II and III (OR=1.50, 95% CI=1.05–2.14 for stage II, and OR=3.33, 95% CI=1.67–6.65 for stage III). These results suggest that miR-149 rs2292832 polymorphism is involved in susceptibility to developing CWP.
Journal of Human Genetics (2010) **55,** 13–17; doi:10.1038/jhg.2009.112; published online 30 October 2009

Keywords: coal workers' pneumoconiosis; genetic susceptibility; molecular epidemiology; pre-microRNAs

INTRODUCTION

Coal workers' pneumoconiosis (CWP) is a lung disease produced by the inhalation and deposition of occupational coal dust. There are more than 100 000 coal miners in the United States of America. In China, there are more than 5 000 000 coal miners and over 600 000 cases with silicosis, which are half of the cases in the world.[1] CWP is characterized by the accumulation of inflammatory cells in the lung, thickening of the alveolar walls and formation of fibrotic nodules.[2] However, only a part of the individuals exposed to coal dust or silica develop CWP in their life, suggesting that genetic susceptibility factors also have a role in the development of CWP.

microRNAs (miRNAs) are a family of endogenous, small and noncoding RNAs, 17–27nt in length, that negatively regulate gene expression by suppressing translation or degrading mRNAs.[3] Strongly conserved among distantly related organisms, miRNAs are involved in a variety of biologic processes, including cell cycle regulation, differentiation, development, metabolism, neuronal patterning and aging.[4] It has been estimated that the human genome contains about 1000 miRNAs[5] and that they could regulate nearly 30% of human genes.[6] Genetic variants including single-nucleotide polymorphisms (SNPs) or mutations in miRNAs may alter miRNA expression and/or maturation. For example, an SNP located at the eighth nucleotide of mature miR-125a in normal subjects blocks pri-miR-125a processing to pre-miRNA and thus reduces the miRNA-mediated translational suppression of its target.[7] Recently, Hu *et al.*[8] conducted a screening for common SNPs in miRNA sequences and identified four SNPs (rs2910164, rs2292832, rs11614913 and rs3746444) located at the pre-miRNA regions of miR-146a, miR-149, miR-196a2 and miR-499, respectively (Figure 1). Furthermore, they also evaluated associations between these SNPs in the pre-miRNAs and lung cancer susceptibility.[9] As CWP is a dust-associated lung disease, it could have similar mechanisms underlying the development of lung cancer.

In this study, we hypothesized that the four SNPs in the pre-miRNAs are associated with risk of CWP. To test this hypothesis, we genotyped the polymorphisms and assessed the associations with risk of CWP in our ongoing case–control study in the Chinese population.

MATERIALS AND METHODS

Study subjects

Four hundred ninety-six CWP patients and 513 controls were recruited from the coal mines of Xuzhou Mining Business Group, China, between January 2006 and December 2008. High kilovolt chest X-ray and physical examinations

[1]Department of Occupational Medicine and Environmental Health, Nanjing Medical University, Nanjing, China and [2]General Hospital of Xuzhou Mining Business Group Co Ltd, Xuzhou, China
Correspondence: Dr C Ni, Department of Occupational Medicine and Environmental Health, School of Public Health, Nanjing Medical University, 140 Hanzhong Road, Nanjing 210029, China.
E-mail: chni@njmu.edu.cn
[3]These authors contributed equally to this work.
Received 20 July 2009; revised 21 September 2009; accepted 7 October 2009; published online 30 October 2009

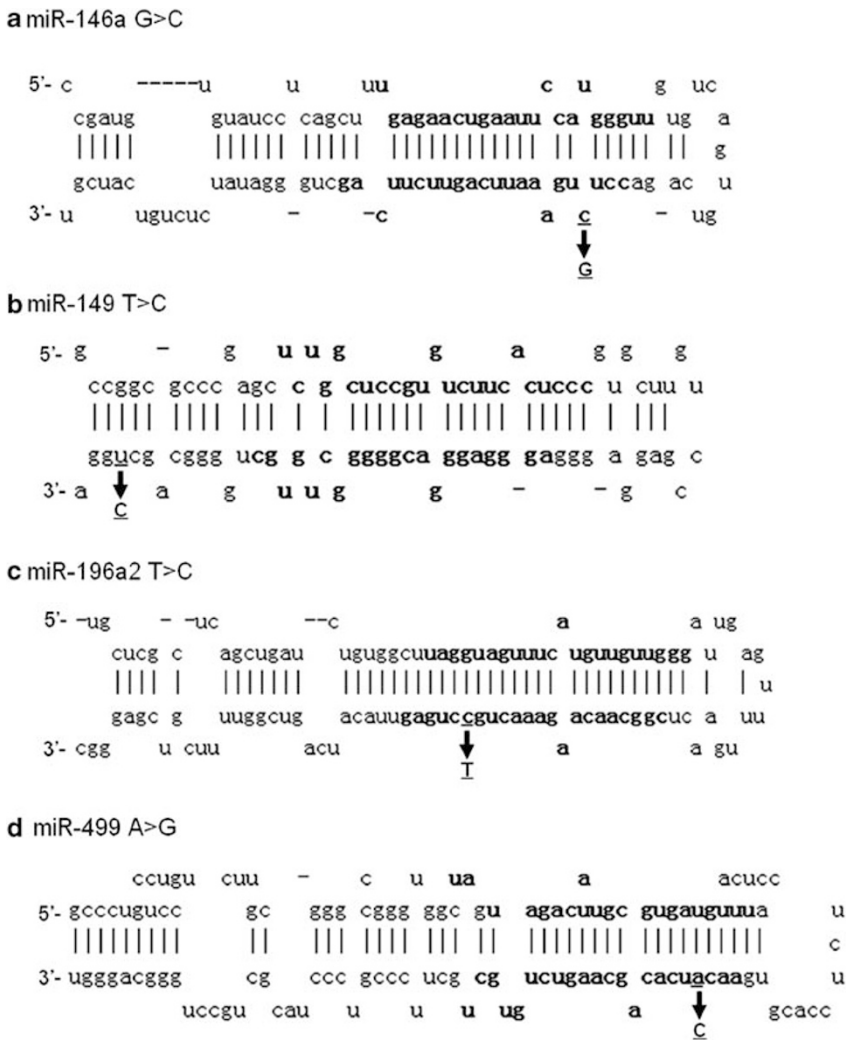

Figure 1 Sequences and mature forms of microRNAs (miRNAs) (**a**–**d**). Sequence for mature miRNA is bolded, and the position of single-nucleotide polymorphism in the miRNA is underlined.

were performed for all, reconfirming the diagnoses based on the 2002 Chinese Roentgenodiagnostic Criteria of Pneumoconiosis, which is the same as the 1980 International Labour Office Classification of Pneumoconioses in the judgment of opacity profusion. The pneumoconiosis patients were classified into stage I, stage II and stage III according to the size, profusion and distribution range of opacities on chest X-ray. The participation rate of the cases was approximately 96%. The control subjects were workers matched with the CWP cases for dust exposure years and work types from the same company. The response rate of those controls contacted for participation in the study was approximately 88%. Those who refused to participate in the study did not want to donate their blood samples for the genetic testing needed in this study. Information about individual demographics as well as data on smoking status was obtained through face-to-face interviews. The research protocol was approved by the institutional review board of Nanjing Medical University.

SNP selection and genotyping

The study for SNP selection has been described in detail elsewhere.[8] Briefly, there were a total of five common SNPs associated with 400 known human pre-miRNAs (miR-146a rs2910164, mirR-149 rs2292832, miR-196a2 rs11614913, miR-499 rs3746444 and miR-423 rs6505162) for the Chinese listed in the public database miRBase. As miR-423 rs6505162 SNP did not affect the hydrogen bands or predict secondary structure free energy, only four SNPs were included in the study. The polymorphism was determined using the PCR-restriction fragment length polymorphism (RFLP) method. The primers, lengths and restriction enzymes have been previously described.[8] The gel electrophoresis pictures of the four SNPs are shown in Figure 2. The polymorphism analysis was performed by two persons independently in a blind manner. More than 10% of the samples were randomly selected for confirmation, and the results were 100% concordant. In addition, the PCR-RFLP results were confirmed by the direct sequencing method to verify the genotypes. As a result, we found that the genotypes using direct DNA sequencing were identical with the PCR-RFLP method.

Statistical analysis

Differences in selected demographic variables, dust exposure years, the known risk factors, such as tobacco smoking, and each allele and genotype of four polymorphisms between the cases and controls were evaluated by using Student's *t*-test (for continuous variables) and χ^2-test (for categorical variables). The Hardy–Weinberg equilibrium of the controls' genotype distributions was tested by a goodness-of-fit χ^2-test. Unconditional univariate and multivariate logistic regression analyses were performed to obtain crude and adjusted odds ratios (ORs) for risk of CWP and their 95% confidence intervals (CIs). All tests were two-sided by using the SAS software (version 9.1; SAS Institute, Cary, NC, USA), unless indicated otherwise.

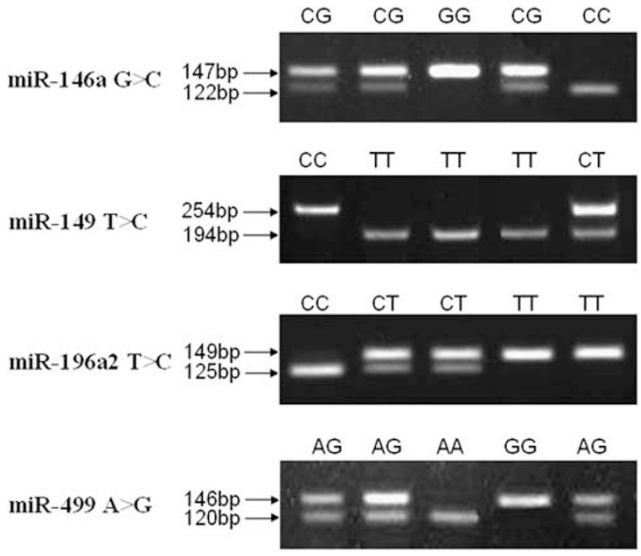

Figure 2 miR-146a rs2910164, miR-149 rs2292832, miR-196a2 rs11614913 and miR-499 rs3746444 genotypes determined by PCR-restriction fragment length polymorphism.

RESULTS

Characteristics of the study subjects

The frequency distributions of selected characteristics of the cases and controls are presented in Table 1. There was no significant difference in the distribution of exposure years and tobacco smoking between the cases and controls (P=0.658 for exposure years, and P=0.110 for tobacco smoking). However, there was a significant difference between the cases and controls in terms of age ($P<0.001$). Specifically, light smokers (0–20 pack-years) had a 2.22-fold (95% CI, 1.61–3.06) increased risk compared with nonsmokers. In addition, of the 496 CWP patients, 228 (46.0%) were stage I, 213 (42.9%) were stage II and the remaining 55 (11.1%) were stage III.

Association between the pre-miRNA polymorphisms and CWP risk

As shown in Table 2, for the miR-149 rs2292832 polymorphism, the difference in the frequencies of genotypes and alleles between cases and controls was statistically significant (P=0.035 for genotypes and P=0.024 for alleles). The genotype distributions of these SNPs in our controls were in agreement with the Hardy–Weinberg equilibrium (P=0.060 for mirR-149 rs2292832, P=0.206 for miR-196a2 rs11614913 and P=0.136 for miR-499 rs3746444), but not for one SNP (P=0.009 for miR-146a rs2910164). In this study, the miR-149 rs2292832 CT genotype was statistically significantly associated with a deceased risk of CWP (adjusted OR=0.75, 95% CI=0.58–0.98), compared with the TT genotype. Furthermore, when we used the CT/CC genotypes as the reference, we found that the TT genotype was associated with a significantly increased risk of CWP (adjusted OR=1.31, 95% CI=1.01–1.69). However, no significant association with CWP risk was identified for the other SNPs in the pre-miRNAs examined in this study.

Stratification analyses between the genotypes of four SNPs in the pre-miRNAs and risk of CWP

In the stratification analyses, we found that individuals with the TT genotype had a significantly increased risk of CWP than those with the CT/CC genotypes, and this increased risk was more pronounced

Table 1 Distribution of selected variables between the coal workers' pneumoconiosis (CWP) cases and control subjects

	CWP (n=496)		*Controls (n=513)*		
Variables	n	%	n	%	P-*value*
Age (mean ± s.d.)	69.5 ± 9.7		65.7 ± 6.8		<0.001
Exposure years (mean ± s.d.)	27.0 ± 8.7		27.2 ± 7.6		0.658
Smoking status					0.110
Never	282	56.9	317	61.8	
Ever	214	43.2	196	38.2	
Former	107	21.6	23	4.5	
Current	107	21.6	173	33.7	
Pack-years smoked					<0.001
0	282	56.9	317	61.8	
0–20	148	29.8	75	14.6	
>20	66	13.3	121	23.6	
Stage					
I	228	46.0			
II	213	42.9			
III	55	11.1			

Table 2 Genotype and allele frequencies of pre-miRNA polymorphisms among the coal workers' pneumoconiosis (CWP) cases and controls and their associations with risk of CWP

	CWP cases (n=496)		*Controls (n=513)*			
Genotypes	n	%	n	%	P-*value*[a]	*OR (95% CI)*[b]
miR-146a						
GG	145	29.2	164	32.0	0.162	1.00
CG	260	52.4	277	54.0		1.06 (0.80–1.42)
CC	91	18.4	72	14.0		1.42 (0.96–2.10)
C allele	0.446		0.410		0.110	
miR-149						
TT	299	60.3	268	52.2	0.035	1.00
CT	173	34.9	217	42.3		0.75 (0.58–0.98)
CC	24	4.8	28	5.5		0.86 (0.48–1.54)
CT/CC	197	39.7	245	47.8		1.00
TT	299	60.3	268	52.2	0.010	1.31 (1.01–1.69)
C allele	0.223		0.266		0.024	
miR-196a2						
TT	122	24.6	143	27.9	0.273	1.00
CT	285	57.5	269	52.4		1.31 (0.97–1.77)
CC	89	17.9	101	19.7		1.07 (0.73–1.57)
C allele	0.467		0.459		0.730	
miR-499						
AA	358	72.2	349	68.0	0.352	1.00
AG	129	26.0	154	30.0		0.84 (0.63–1.11)
GG	9	1.8	10	2.0		0.79 (0.30–2.03)
G allele	0.148		0.170		0.189	

Abbreviations: CI, confidence interval; OR, odds ratio.
[a]Two-sided χ^2-test.
[b]Adjusted for age, exposure years and pack-years of smoking in logistic regression model.

Table 3 Stratification analyses between the genotypes of miR-149 rs2292832 polymorphism and coal workers' pneumoconiosis (CWP) risk

		Genotypes (cases/controls)					
		CT/CC		*TT*			
Variables	*Cases/controls*	n	*%*	n	*%*	P-*value*	*OR (95% CI)*[a]
Total	496/513	197/245	39.7/47.8	299/268	60.3/52.2	0.010	1.31 (1.01–1.69)
Age, years							
<68	182/351	89/174	48.9/49.6	93/177	51.1/50.4	0.840	1.04 (0.72–1.49)
⩾68	314/162	108/71	34.4/43.8	206/91	65.6/56.2	0.035	1.52 (1.03–2.25)
Exposure years							
<26	177/170	68/77	38.4/45.3	109/93	61.6/54.7	0.205	1.32 (0.86–2.03)
⩾26	319/343	129/168	40.4/49.0	190/175	59.6/51.0	0.027	1.42 (1.04–1.93)
Smoking status							
Never	282/317	115/151	40.8/47.6	167/166	59.2/52.4	0.092	1.32 (0.96–1.83)
Ever	214/196	82/94	38.3/48.0	132/102	61.7/52.0	0.049	1.48 (1.00–2.20)
Stage							
I	228/513	105/245	46.0/47.8	123/268	54.0/52.2	0.708	1.06 (0.78–1.45)
II	213/513	79/245	37.1/47.8	134/268	62.9/52.2	0.027	1.50 (1.05–2.14)
III	55/513	13/245	23.6/47.8	42/268	76.4/52.2	0.007	3.33 (1.67–6.65)

Abbreviations: CI, confidence interval; OR, odds ratio.
[a]Adjusted for age, exposure years and pack-years of smoking in logistic regression model.

among subgroups of those aged ⩾68 years (OR=1.52, 95% CI=1.03–2.25), dust exposure ⩾26 years (OR=1.42, 95% CI=1.04–1.93) and ever smokers (OR=1.48, 95% CI=1.00–2.20) (Table 3). In addition, significant associations were observed between the genotypes and patients with stages II and III (OR=1.50, 95% CI=1.05–2.14 for stage II, and OR=3.33, 95% CI=1.67–6.65 for stage III). However, no statistical evidence was found for the gene–environment interaction (data not shown).

DISCUSSION

In this study, we found that the miR-149 rs2292832 TT genotype was associated with a significantly increased risk of CWP compared with the CT/CC genotypes. In addition, the increased CWP risk was more significant in older smokers with long exposure years. Furthermore, statistical evidence was observed for the polymorphism and CWP patients with stages II and III. To the best of our knowledge, this is the first study of the association of common SNPs in pre-miRNAs with CWP risk.

The biogenesis of miRNAs involves a complex protein system, including members of the Argonaute family, Pol II-dependent transcription and the RNAase IIIs Drosha and Dicer.[10] miRNAs are involved in crucial biological processes, including development, differentiation, apoptosis and proliferation.[3,11] From the systematic expression analysis of 217 mammalian miRNAs from 334 samples by Lu *et al.*,[12] miR-149 was widely detected in human lung normal and tumor tissues. Sequence variations within the miRNA genes could potentially influence the processing and/or target selection of miRNAs.[7] A few studies have been performed that examined sequence variations in miRNA regions.[7,8,13,14] For example, Hu *et al.*[8] found that rs11614913 in the pre-miRNA-196a2 was associated with an increase in mature miRNA expression, which contributes to the development and prognosis of lung cancer. In this study, we found a positive association between rs2292832 in miR-149 and CWP risk. However, the potential of mechanism underlying this effect remains unknown. Whether the SNP rs2292832 in miR-149 could result in the expression of mature miRNA or not should be confirmed in further studies. In addition, many human miRNA transcripts have target sequences in the 3′-UTR to which miRNAs bind and exert posttranscriptional repression. As shown in Table 4, some of the predicted probable targets of miR-149 have been experimentally verified. For example, studies showed that N-deacetylase/N-sulfotransferase-1 (NDST1) is an essential NDSR isozyme in mouse embryonic development,[15] and disruption of NDST1, resulting in severe malformations in the lung.[16] Therefore, miR-149 was likely to affect the expression of the miRNA target gene NDST1 and contribute to the susceptibility of CWP. However, functional characterizations of SNP rs2292832 in miR-149 and its target mRNAs in CWP are warranted.

In this study, we also found that the increased risk associated with the rs2292832 TT genotype was more pronounced among older smokers with long exposure years. It is possible that the significant difference in the results may be due to the differences in the age of study subjects. For example, the control population is younger than the study population and may have lesser risk of CWP. Another possible explanation is that individuals in those subgroups may be more likely to have been exposed to some risk factors involved in the etiology of CWP risk, such as tobacco smoking, alcohol use or occupational coal mine dust.[17] However, the lack of significant interactions suggested that the sample size of our study may be insufficient. After stratification for CWP stage, it appeared that the miR-149 rs2292832 TT genotype had an increased risk of CWP with stages II and III, but not with stage I. It has been posed that different stages of CWP have different etiology involving different genetic defects.[18] Therefore, it seems plausible that the miR-149 rs2292832

Table 4 Predicted most probable targets of miR-149

Rank[a]	*Human refseq ID*	*Gene*	*PicTar score*
1	NM_001543	NDST1, N-deacetylase/N-sulfotransferase (heparan glucosaminyl)-1	11.21
2	NM_000599	IGFBP5, insulin-like growth factor binding protein-5	9.95
3	NM_001002762	DNAJB12, DnaJ (Hsp40) homolog, subfamily B, member-12	9.43
4	NM_021116	ADCY1, adenylate cyclase-1	9.32
5	NM_021020	LZTS1, leucine zipper, putative tumor suppressor-1	9.22
6	NM_018425	PI4KII, phosphatidylinositol 4-kinase type II	7.94
7	NM_003085	SNCB, synuclein, β, transcript variant-2	7.77
8	NM_001001502	SNCB, synuclein, β, transcript variant-1	7.77
9	NM_015075	IQSEC2, IQ motif and Sec7 domain-2	7.62
10	NM_001969	EIF5, eukaryotic translation initiation factor-5	7.50

[a]The 10 targeted genes with the highest PicTar scores.

polymorphism is involved in the higher stage of CWP patients. However, this hypothesis needs to be confirmed in further studies.

In conclusion, this study indicates that the SNP rs2292832 in miR-149 modulates the risk of CWP in the Chinese population. These findings should be further validated by larger, preferably prospective studies with more diverse ethnic groups.

ACKNOWLEDGEMENTS

This study was partly supported by the National Natural Science Foundation of China (30872093).

1 Liu, B & Li, Y. Overview and prospect on mechanisms of pneumonoconiosos in China. *Chinese Journal of Industrial Medicine* **20,** 3–5 (2007).
2 Schins, R. P. & Borm, P. J. Mechanisms and mediators in coal dust induced toxicity: a review. *Ann. Occup. Hyg.* **43,** 7–33 (1999).
3 Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116,** 281–297 (2004).
4 Ambros, V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* **113,** 673–676 (2003).
5 Bentwich, I., Avniel, A., Karov, Y., Aharonov, R., Gilad, S., Barad, O. *et al.* Identification of hundreds of conserved and nonconserved human microRNAs. *Nat. Genet.* **37,** 766–770 (2005).
6 Berezikov, E., Guryev, V., van de Belt, J., Wienholds, E., Plasterk, R. H. & Cuppen, E. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120,** 21–24 (2005).
7 Duan, R., Pak, C. & Jin, P. Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. *Hum. Mol. Genet.* **16,** 1124–1131 (2007).
8 Hu, Z., Chen, J., Tian, T., Zhou, X., Gu, H., Xu, L. *et al.* Genetic variants of miRNA sequences and non-small cell lung cancer survival. *J. Clin. Invest.* **118,** 2600–2608 (2008).
9 Tian, T., Shu, Y., Chen, J., Hu, Z., Xu, L., Jin, G. *et al.* A functional genetic variant in microRNA-196a2 is associated with increased susceptibility of lung cancer in Chinese. *Cancer Epidemiol. Biomarkers Prev.* **18,** 1183–1187 (2009).
10 Kim, V. N. & Nam, J. W. Genomics of microRNA. *Trends Genet.* **22,** 165–173 (2006).
11 Harfe, B. D. MicroRNAs in vertebrate development. *Curr. Opin. Genet. Dev.* **15,** 410–415 (2005).
12 Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D. *et al.* MicroRNA expression profiles classify human cancers. *Nature* **435,** 834–838 (2005).
13 Saunders, M. A., Liang, H. & Li, W. H. Human polymorphism at microRNAs and microRNA target sites. *Proc. Natl Acad. Sci. USA* **104,** 3300–3305 (2007).
14 Iwai, N. & Naraba, H. Polymorphisms in human pre-miRNAs. *Biochem. Biophys. Res. Commun.* **331,** 1439–1444 (2005).
15 Hu, Z., Wang, C., Xiao, Y., Sheng, N., Chen, Y., Xu, Y. *et al.* NDST1-dependent heparan sulfate regulates BMP signaling and internalization in lung development. *J. Cell. Sci.* **122,** 1145–1154 (2009).
16 Ringvall, M., Ledin, J., Holmborn, K., van Kuppevelt, T., Ellin, F., Eriksson, I. *et al.* Defective heparan sulfate biosynthesis and neonatal lethality in mice lacking N-deacetylase/N-sulfotransferase-1. *J. Biol. Chem.* **275,** 25926–25930 (2000).
17 Ng, T. P. & Chan, S. L. Factors associated with massive fibrosis in silicosis. *Thorax.* **46,** 229–232 (1991).
18 Chang, K. C., Leung, C. C. & Tam, C. M. Tuberculosis risk factors in a silicotic cohort in Hong Kong. *Int. J. Tuberc. Lung Dis.* **5,** 177–184 (2001).