Genetic Epidemiology 13:35–47 (1996)

Segregation Analysis of Two Lung Function Indices in a Random Sample of Young Families: The Humboldt Family Study

Yue Chen, Sandra L. Horne, Donna C. Rennie, and James A. Dosman

Centre for Agricultural Medicine (Y.C., D.C.R., J.A.D.) and Department of Biology (S.L.H.), University of Saskatchewan, Saskatoon, Saskatchewan, Canada

The Humboldt Family Study was conducted in the town of Humboldt, Saskatchewan, in 1993. Familial correlations and segregation analyses of lung function were carried out in 799 individuals in 214 nuclear families that included 214 fathers, 214 mothers, and 371 children. Forced expiratory volume in 1 second (FEV_1) and maximal mid-expiratory flow rate (MMFR) were first regressed on age, height, weight, and their quadratic and cubic terms as well as on smoking status in four groups separately (mothers, fathers, daughters, and sons), with terms significant at the 0.10 level being retained. Residual phenotypes were standardized within the four groups. Class D regressive models were used to perform familial correlations and segregation analyses. For both FEV_1 and MMFR, father–mother correlations were not significantly different from zero, and mother–offspring, father–offspring, and sibling–sibling correlations showed no statistically significant difference from each other. Based on the "polygenic" models, the estimated intraclass correlation is 0.132 ($\pm$0.035) for FEV_1 and 0.171 ($\pm$0.039) for MMFR, and the narrow-sense heritability is 0.264 for FEV_1 and 0.342 for MMFR.

Segregation analysis shows that the "mixed" model with both single locus and polygenic components had a better fit for FEV_1 than single-locus or polygenic-only models. However, the model which included a nontransmitted environmental factor [$\tau(AA) = \tau(AB) = \tau(BB) = q_A$] and polygenic loci had a better fit than the Mendelian model [$\tau(AA) = 1$, $\tau(AB) = 1/2$, $\tau(BB) = 0$] [Akaike's information criterion (AIC) = 2219.47 vs. AIC = 2222.14]. For MMFR, the Mendelian "mixed" model gave a nonsignificant improvement in $\log_e$ likelihood compared to the simple polygenic model. Comparison of the single-locus model and Mendelian

Received for publication May 25, 1995; revision accepted September 19, 1995.

Address reprint requests to Dr. Yue Chen, Centre for Agricultural Medicine, Wing 3E, Royal University Hospital, University of Saskatchewan, 103 Hospital Drive, Saskatoon, Saskatchewan, Canada S7N 0W8.

"mixed" model shows no difference in fitting the data. This study suggests that FEV_1 and MMFR are controlled by many loci with no major effects and/or common environmental factors.

Key words: genetics, families, pulmonary function

INTRODUCTION

There is considerable evidence incriminating risk factors that can act singly or in concert resulting in decrease of lung function and the genesis of chronic obstructive pulmonary disease (COPD). Only cigarette smoking and severe α_1-antitrypsin deficiency, however, are considered established causes of clinically significant COPD in the absence of other agents [U.S. Department of Health and Human Services, 1984]. A number of studies have provided evidence of familial aggregation in lung function measurements. Family studies have demonstrated significant parent–offspring and sibling–sibling correlations for lung function test variables [Cotch et al., 1990; Coultas et al., 1991; Devor and Crawford, 1984; Higgins and Keller, 1975; Kauffman et al., 1989; Lewitter et al., 1984; Schilling et al., 1977; Tager et al., 1976], and have indicated that first-degree relatives of COPD patients have increased rates of impaired lung function when compared to nonpulmonary patients or healthy control subjects [Cohen, 1980; Khoury et al., 1985]. Twin studies have indicated that intrapair correlations for lung function measurements are significantly higher in monozygotic twins than in dizygotic twins [Hubert et al., 1982; Kawakami et al., 1985; Man and Zamel, 1976; Redline et al., 1987, 1989]. Because common environmental determinants do not account for all or most of the familial aggregation, genetic control over lung function is a reasonable probability.

Specific genetic factors in the development of COPD have not been clearly identified, except for Protease inhibitor types. α_1-Antitrypsin deficiency, however, is rare in the general population [Horne et al., 1984, 1992] and accounts for a relatively small proportion of the cases of COPD [Cohen, 1980]. Therefore, other potential genetic mechanisms for lung dysfunction require study.

Rybicki et al. [1990] reported major genetic effects on lung function among COPD families by using regressive models, and found that major gene effects could explain all of the familial correlations for forced expiratory volume in 1 sec (FEV_1) in COPD families. In families of nonpulmonary patients, however, there were no familial correlations for FEV_1 and therefore, no evidence of genetic control of lung function [Rybicki et al., 1990]. If there is a single locus which has a major effect on susceptibility to lung dysfunction, it would be of major importance in understanding the etiology and the prevention of COPD.

We conducted a community-based family study of children aged 6 to 17 years and their biological parents. We searched for potential mechanisms of genetic control for lung function.

MATERIALS AND METHODS

Study Subjects

The Humboldt Family Study was conducted in the town of Humboldt, Saskatchewan, in 1993. The town has a stable population, lack of industrial air pollution, and

a history of cooperation in health surveys [Dosman eta l., 1981; Chen et al., 1991a,b]. The target population of young nuclear families was ascertained through parents with children aged 6 to 17 years living in this area. Potential genetic effects are considered to be less confounded by environmental factors on lung function in a relatively young population. For the childhood portion of the investigation, all schools in the town (one high school and three primary schools) provided lists of enrolled students aged 6 to 17 years. Subjects under 18 years of age who were not attending school were identified by means of a total town canvass that was conducted for the adult portion of a cross-sectional study. Almost all town residents were of Caucasian background.

A nuclear family in this study includes both parents and at least one child. Certain exclusions were made for this analysis. Step-offspring or adopted offspring were excluded. A family with only one parent participating in the study was not included. Two hundred fourteen nuclear families were ascertained, including 214 fathers, 214 mothers, and 371 offspring. In this study, there were no conditions on phenotypes in the selection of nuclear families, and random sampling was assumed [SAGE-REGC program, 1994]. The distribution of nuclear family sizes is presented in Table I.

Data Collection and Measures

Three questionnaires were separately designed for 1) children (6–11 years) 2) adolescents (12–17 years), and 3) their parents. Questionnaires for children and adolescents were completed by their parents. Each adolescent also completed a separate questionnaire for questions on life style. The adult questionnaire was self-administered. The questionnaires included information on socio-demographic factors, alcohol consumption, exercise, the home environment, and individual and family history of pulmonary and cardiovascular diseases and diabetes. Information about smoking was collected on all subjects except children less than 12 years of age.

An appointment was made for each participant for a clinic visit where lung function, height, and weight were measured. Two MedGrahics CPF-S Systems (Medical Graphics Corporation, St. Paul, MN) were used for lung function testing. Each subject was tested until three acceptable forced expiratory maneuvers were obtained. The standard for choosing forced vital capacity (FVC), forced expiratory volume in 1 sec (FEV_1), and maximal mid-expiratory flow rate (MMFR) were 1) the best FVC and FEV_1, not necessarily from the same tracing and 2) the MMFR, which came from the tracing with the best sum of FVC and FEV_1 [American Thoracic Society, 1987].

Weight was measured to the nearest 0.5 kg using a calibrated hospital spring scale with subjects dressed in normal indoor clothing. Height in centimetres was

TABLE I. Distribution of Nuclear Family Sizes

Family sizes	Number	%
3	102	47.7
4	72	33.6
5	36	16.8
6	3	1.4
7	1	0.5
Total	214	100.0

measured against a wall, using a wall-mounted tape measure and a fixed square. Subjects did not wear shoes for weight and height measurements.

Statistical Analysis

Data adjustments. Adjustments for the effects of age, height, weight, and smoking habits on lung function variables were performed separately within four groups (mothers, fathers, daughters, and sons). In this analysis, a smoker was defined as a subject who was currently smoking at least one cigarette (or smoking pipes and cigars) a day every day or almost every day. A person who had formerly smoked regularly but had quit smoking for at least 6 months at the time of the study was defined as an ex-smoker. A subject who had never smoked any kind of tobacco regularly or had smoked a total of less than one-half pack-year (i.e., less than one pack daily for 6 months) was defined as a nonsmoker. Information on smoking for those children aged 6 to 11 years was not collected, and they were treated as nonsmokers. The distribution of characteristics in each group is presented in Table II.

FEV_1 and MMFR were regressed on age, height, weight, and their quadratic and cubic terms as well as on smoking status, with the terms significant at the 0.10 level being retained. In children, smoking was not significantly related to either FEV_1 or MMFR and was excluded from the equations. Residual phenotypes were also stan-

TABLE II. Distribution of Characteristics Among Participants*

	Mothers (n = 214)		Fathers (n = 214)		Girls (n = 182)		Boys (n = 189)	
	Mean	(SD)	Mean	(SD)	Mean	(SD)	Mean	(SD)
Age (years)	37.4	(6.2)	39.7	(7.5)	10.7	(3.4)	10.5	(3.1)
Height (cm)	164.0	(6.4)	177.2	(6.2)	146.5	(17.5)	147.8	(19.8)
Weight (kg)	68.5	(12.6)	87.0	(14.3)	41.0	(15.9)	41.8	(16.4)
FEV_1 (L)	3.11	(0.47)	4.19	(0.76)	2.38	(0.84)	2.56	(1.02)
MMFR (L/s)	3.25	(0.84)	4.16	(1.25)	2.88	(1.13)	2.85	(1.17)
	No.	(%)	No.	(%)	No.	(%)	No.	(%)
Smoking status								
Nonsmokers	124	(57.9)	92	(43.0)	174	(95.6)	188	(99.5)
Ex-smokers	57	(26.6)	78	(36.4)	0	(0.0)	0	(0.0)
Current smokers	33	(15.4)	44	(20.6)	8	(4.4)	1	(0.5)
Household ETS exposure								
No	175	(81.8)	181	(84.6)	131	(72.0)	149	(78.8)
Yes	39	(18.2)	33	(15.4)	51	(28.0)	40	(21.2)
Respiratory allergy								
No	147	(68.7)	160	(74.8)	146	(80.2)	134	(70.9)
Yes	67	(31.3)	54	(25.2)	36	(19.8)	55	(29.1)
Grain farming exposure[a]								
No	204	(95.3)	152	(71.0)	167	(91.8)	151	(79.9)
Yes	10	(4.7)	62	(29.0)	15	(8.2)	38	(20.1)

*Definitions of abbreviations: SD, standard deviations; ETS, environmental tobacco smoke; FEV_1, forced expiratory volume in 1 sec; MMFR, maximal mid-expiratory flow rate.
[a]Different questions were used for adults and children.

dardized within the four groups. The normality of the standardized residual phenotypes was tested by using the SPSS Explore procedure [Norusis, 1993].

Familial correlations and segregation analyses. Familial analyses were performed with the REGC program of the Statistical Analysis for Genetic Epidemiology (SAGE) package [1994]. The regressive models introduced by Bonney [Bonney, 1984] were first used to examine the familial patterns of correlations of lung function variables with no major gene. The dependencies in nuclear families were modeled as a Markovian process by conditioning each individual trait on those of family members. Class D models assume that sibling–sibling (brother–brother, sister–sister, and brother–sister) correlations are equal but not necessarily due to common parentage alone [Demenais and Bonney, 1989].

The genetic heritability is a ratio of the genetic variance to the total phenotypic variance in the population. The genetic variance includes the components of additive variance, dominant variance, and epistatic variance. Heritability in the narrow sense is the proportion of phenotypic variation in a population that is due to the additive effects of alleles at one or more loci. In an additive polygenic model, the narrow-sense heritability is twice the correlation between first-degree relatives [Khoury et al., 1993, pp. 271].

Subsequently, segregation analysis was carried out, which involved fitting single-locus, polygenic, and "mixed" ("mixed" here means including both single-locus and polygenic components in regressive models) models of inheritance. The major gene effects are assumed to result from segregation at a single locus having two alleles, A and B. In this analysis, allele A is associated with lower lung function measurements. The parameters in the models include gene frequencies (q_A), transmission probabilities for each genotype of transmitting A [τ(AA), τ(AB), and τ(BB)], population means (μ), genotypic means [μ(AA), μ(AB), and μ(BB)], genotypic variances (σ^2), and familial correlations (ρ = correlations among all first-degree relatives, ρ_{fm} = father–mother correlations; ρ_{mo} = mother–offspring correlations; ρ_{fo} = father–offspring correlations; and ρ_{sib} = sibling–sibling correlations). The single-locus models that we tested included 1) dominant [μ(AA) = μ(AB)]; 2) recessive [μ(AB) = μ(BB)]; 3) additive {μ(AB) = [μ(AA) + μ(BB)]/2}; and 4) codominant (arbitrary). Under Mendelian transmission, τ(AA) = 1, τ(AB) = 1/2, and τ(BB) = 0. The non-transmitted environmental effect was obtained with the three transmission probabilities being equal to q_A [τ(AA) = τ(AB) = τ(BB) = q_A].

A likelihood-ratio test (LRT) was used to select the most parsimonious model, which is minus twice the difference in the $\log_e$ likelihood ($\ln L$) between models before and after reducing parameters. The LRT is distributed asymptotically as a chi-square with degrees of freedom (df) equal to the difference in the number of parameters between two models. However, if the value of a parameter under the null hypothesis is at the boundary of the parameter space, the LRT statistic does not follow a simple chi-square distribution [Khoury et al., 1993, pp. 216]. In addition, the LRT is based on a comparison of strictly hierarchical models. For several alternative non-hierarchical models, the better-fitting model was considered with a lower value of the Akaike's information criterion [AIC = $-2 \ln L + 2$(number of parameters estimated)] [Akaike, 1974].

RESULTS

Familial Correlations

Familial correlations for residual FEV_1 estimated under the class D model are presented in Table III. The data show that the father–mother correlation was trivial (model 1). When the father–mother correlation was set to zero, model 2 shows no statistically significant difference compared with model 1 (LRT = 0.38, df = 1, P = .538), and therefore, the hypothesis of no father–mother correlation for residual FEV_1 cannot be rejected. Comparison of models 3 and 2 shows no significant difference in mother–offspring and father–offspring correlations (LRT = 0.89, df = 1, P = .346). Parent–offspring correlations were statistically significant (model 5 vs. model 2: LRT = 13.76, df = 2, P = .001). In model 4, we fixed the parent–offspring correlation equal to the sibling–sibling correlation. Since the sibling–sibling correlation is modeled as an intraclass correlation and is constrained to be non-negative (bounded), the LRT for the null hypothesis that the correlation equals zero follows a mixture of a chi-square distribution with 1 df and a degenerate chi-square distribution with 0 df when it is fixed equal to an unbounded parameter [Khoury et al., 1993, pp. 268]. However, the LRT shows that the difference between models 4 and 2 is not statistically significant at either 1 or 2 degrees of freedom [LRT = 1.02, 0.313 (df = 1) < P < .601 (df = 2)]; AIC indicates model 4 has a better fit than other models (Table III).

A series of models of the familial correlations for residual MMFR estimated under the class D model are presented in Table IV. The results were similar to those for FEV_1, and are summarized as following: 1) no father–mother correlation could not be rejected; 2) mother–offspring and father–offspring showed no significant dif-

TABLE III. Familial Correlation (Standard Deviation) for Residual FEV_1 Estimated Under Class D Regressive (No Major Gene) Models*

Model	ρ_{fm}	ρ_{mo}	ρ_{fo}	ρ_{sib}	$-2\ln L$ {parameters}	AIC
1. Arbitrary	0.042 (0.068)	0.106 (0.054)	0.176 (0.053)	0.111 (0.071)	2,250.04 {6}	2,262.04
2. No father–mother	[0][a]	0.099 (0.053)	0.172 (0.053)	0.109 (0.071)	2,250.42 {5}	2,260.42
3. Equal parent–offspring	[0][a]	0.135 (0.036)	$=\rho_{mo}$	0.109 (0.072)	2,251.31 {4}	2,259.31
4. Equal parent–offspring and sibling–sibling	[0][a]	0.132 (0.035)	$=\rho_{mo}$	$=\rho_{mo}$	2,251.44 {3}	2,257.44
5. No parent–offspring	[0][a]	[0][a]	[0][a]	0.112 (0.073)	2,264.18 {3}	2,270.18
6. No correlation	[0][a]	[0][a]	[0][a]	[0][a]	2,266.64 {2}	2,270.64

*Definitions of abbreviations: FEV_1, forced expiratory volume in 1 sec; ρ_{fm}, father–mother correlation; ρ_{mo}, mother–offspring correlation; ρ_{fo}, father-offspring correlation; ρ_{sib}, sibling–sibling correlation; $\ln L$, log likelihood; AIC, Akaike's information criterion. The mean and variance of the standarized residual phenotype are omitted.

[a]Correlation is fixed at zero.

TABLE IV. Familial Correlation (Standard Deviation) for Residual MMFR Estimated Under Class D Regressive (No Major Gene) Models*

Model	ρ_{fm}	ρ_{mo}	ρ_{fo}	ρ_{sib}	$-2 \ln L$ {parameters}	AIC
1. Arbitrary	0.107 (0.067)	0.153 (0.055)	0.222 (0.053)	0.194 (0.067)	2,234.98 {6}	2,246.98
2. No father–mother	[0][a]	0.131 (0.053)	0.209 (0.053)	0.190 (0.067)	2,250.42 {5}	2,247.48
3. Equal parent–offspring	[0][a]	0.169 (0.035)	$= \rho_{mo}$	0.184 (0.067)	2,238.44 {4}	2,246.44
4. Equal parent–offspring and sibling–sibling	[0][a]	0.171 (0.039)	$= \rho_{mo}$	$= \rho_{mo}$	2,238.48 {3}	2,244.48
5. No parent–offspring	[0][a]	[0][a]	[0][a]	0.182 (0.069)	2,259.00 {3}	2,255.65
6. No correlation	[0][a]	[0][a]	[0][a]	[0][a]	2,266.30 {2}	2,270.30

*Definition of abbreviation: MMFR, maximal mid-expiratory flow rate. For other definitions, see Table III. The mean and variance of the standardized residual phenotype are omitted.
[a]Correlation is fixed at zero.

ference; 3) parent–offspring correlation was statistically significant; and 4) the model of equal parent–offspring the sibling–sibling correlations had the best fit (model 4).

Based on model 4, with equal parent–offspring and sibling–sibling correlations (or "polygenic" models), the narrow-sense heritability is 0.264 for FEV_1 and 0.342 for MMFR.

Segregation Analysis

Segregation analysis on polygenic, single-locus, and "mixed" models of inheritance was further performed by using the class D regressive model to determine whether they could account for the familial correlations observed in residual FEV_1 and MMFR. Table V presents the maximum-likelihood estimates of the parameters for FEV_1 under Mendelian inheritance. The general model is taken as the standard, in which all parameters are arbitrary. Comparison of the Mendelian "mixed" and general models shows no significant difference [LRT = 1.21, 0.546 (df = 2) < P < .751 (df = 3)]. The single-locus model with arbitrary genotypic mean had a significantly worse fit compared to the general model (LRT = 25.40, df = 4, P < .001). Other single-locus models including dominant, recessive, and additive models, which are not shown, gave no better fits. The polygenic model also shows a significantly worse fit when compared with the general model (LRT = 42.52, df = 5, P < .001). Therefore, all those models except the Mendelian "mixed" model, were rejected when compared with the general model. AIC indicates that the Mendelian "mixed" model with both single-locus and polygenic components had the best fit (even better than the Mendelian general model).

Transmission parameters were further estimated for FEV_1 under class D regressive models. Table VI shows that both mendelian [τ(AA) = 1, τ(AB) = 1/2, τ(BB) = 0] and environmental [τ(AA) = τ(AB) = τ(BB) = q_A] models cannot be rejected when compared to the general model (arbitrary transmission probabilities). The model which included a nontransmitted environmental factor and polygenic loci

TABLE V. Parameter Estimates From Segregation Analysis of FEV_1 Under Mendelian Transmission–Class D Regressive Models*

	Sporadic	Polygenic	Single locus	"Mixed"	General
q_A	[1.0][a]	[1.0][a]	0.201	0.182	0.182
μ(AA)	0.00	0.00	−2.34	−2.55	−2.54
μ(AB)			−0.22	0.32	0.33
μ(BB)			0.24	0.04	−0.04
σ^2	1.00	1.00	0.75	0.77	0.77
ρ_{fm}	[0][a]	[0][a]	[0][a]	[0][a]	0.071
ρ_{mo}	[0][a]	0.132	[0][a]	0.183	0.184
ρ_{fo}	[0][a]	= ρ_{mo}	[0][a]	= ρ_{mo}	0.210
ρ_{sib}	[0][a]	= ρ_{mo}	[0][a]	= ρ_{mo}	0.155
Parameters	2	3	5	6	9
$-2 \ln L$	2,266.64	2,251.44	2,234.33	2,210.14	2,208.93
AIC	2,270.64	2,257.44	2,244.33	2,222.14	2,226.93

*Definitions of abbreviations: q_A, gene frequency; μ, population mean; μ(AA), μ(AB), and μ(BB), genotypic means; σ^2, genotypic variance. For other definitions, see Table III. Assume transmission probabilities τ(AA) = 1.0, τ(AB) = 0.5, and τ(BB) = 0.0.
[a]Parameters are fixed and not estimated in the models.

had the lowest value of AIC although it showed only a modest improvement in AIC value, suggesting that a nontransmitted environmental factor being responsible for the mixture of distributions could not be ruled out.

Table VII shows the results of segregation analysis for MMFR. The Mendelian "mixed" model fits the data as well as the general model [LRT = 3.46, .177 (df = 2) $< P <$ .362 (df = 3)]. However, the Mendelian "mixed" model gave nonsignificant improvement in $\log_e$ likelihood compared to either the simple polygenic or the single-locus models (LRT = 0.06, df = 3, P = .996; and LRT = 0.80, df = 1, P =

TABLE VI. Transmission Parameter Estimates From Segregation Analysis of FEV_1 Under Class D Regressive Models*

	Mendelian "mixed"	Environmental	General
q_A	0.182	0.182	0.179
τ(AA)	[1.0][a]	= q_A	0.130
τ(AB)	[0.5][a]	= q_A	0.349
τ(BB)	[0.0][a]	= q_A	0.108
μ(AA)	−2.55	−2.51	−2.53
μ(AB)	0.32	0.09	0.40
μ(BB)	0.04	0.07	−0.06
σ^2	0.77	0.79	0.75
ρ	0.183	0.184	0.193
Parameters	6	6	9
$-2 \ln L$	2,210.14	2,207.47	2,206.98
AIC	2,222.14	2,219.47	2,224.98

*For definition of abbreviations, see Tables III and V.
[a]Parameters are fixed and not estimated in the models.

TABLE VII. Parameter Estimates From Segregation Analysis of MMFR Under Mendelian Transmission–Class D Regressive Models*

	Sporadic	Polygenic	Single locus	"Mixed"	General
q_A	[1.0][a]	[1.0][a]	0.570	0.155	0.115
μ(AA)	0.00	0.00	−0.66	−0.68	−0.64
μ(AB)			0.07	0.17	0.12
μ(BB)			0.96	−0.04	−0.02
σ^2	1.00	1.00	0.68	0.98	0.99
ρ_{fm}	[0][a]	[0][a]	[0][a]	[0][a]	0.108
ρ_{mo}	[0][a]	0.171	[0][a]	0.174	0.154
ρ_{fo}	[0][a]	= ρ_{mo}	[0][a]	= ρ_{mo}	0.224
ρ_{sib}	[0][a]	= ρ_{mo}	[0][a]	= ρ_{mo}	0.194
Parameters	2	3	5	6	9
$-2 \ln L$	2,266.30	2,238.48	2,239.22	2,238.42	2,234.96
AIC	2,270.30	2,244.48	2,249.22	2,250.42	2,252.96

*For definition of abbreviations, see Tables III and V. Assume transmission probabilities τ(AA) = 1.0, τ(AB) = 0.5, and τ(BB) = 0.0.
[a]Parameters are fixed and not estimated in the models.

.371), indicating that the single-locus component is not significant in the Mendelian "mixed" model and the simple polygenic model fits as well as the Mendelian "mixed" model. The data suggest that the familial correlation for MMFR is likely controlled by many loci with no major gene effects and/or is due to common environmental factors.

DISCUSSION

The class D regressive model was first adopted to examine familial correlations for residual FEV_1 and MMFR without including major gene components. Our data showed no significant differences in father–mother correlations for FEV_1 and MMFR, which is consistent with the results from previous reports [Coultas et al., 1991; Schilling et al., 1977; Tager et al., 1976]. However, some other family studies demonstrated significant father–mother correlations [Higgins and Keller, 1975; Kauffman et al., 1989]. Kauffman et al. [1989] reported a father–mother correlation of 0.20 for residual FEV_1 and of 0.23 for MMFR. The reasons for the apparent discrepancy among these studies are not known.

In addition, our data also showed no significant differences in mother–offspring, father–offspring, and sibling–sibling correlations for both residual FEV_1 and MMFR. Both of them fit the "polygenic" models before the major gene component was included.

In previous family studies, by using path analysis [Cotch et al., 1990; Coultas et al., 1991; Lewitter et al., 1984] and variance component analysis [Astemborski et al., 1985; Beaty et al., 1987], the narrow-sense heritability for FEV_1 was estimated to range from 28% [Astemborski et al., 1985] to 47% [Lewitter et al., 1984]. Lebowitz et al. [1984], however, found that familial correlations of lung function measurements were dependent on familial aggregation of body habitus, and were no longer significant after taking the Ponderal Index into account. Although overadjustment was commented on the study by Coultas et al. [1991], our study shows that

the familial correlation was significant after adjusting for both height and weight and their quadratic and cubic terms at $\alpha = 0.10$ level. Kauffmann et al. [1989] demonstrated no influence of height on the magnitude of the parent–child correlations. There was also variation in the heritability estimate for FEV_1 from studies of twins. For example, in one study, Hubert et al. [1982] studied 127 monozygotic and 141 dizygotic male twin pairs 42 to 56 years of age, and the heritability was estimated as high as 77%. In another study, however, Ghio et al. [1989] studied 74 university student pairs of sample-sex twins with an average age of 20 years and found that the heritability was not significant after adjustment for height. Since heritability is a ratio of the genetic variance to the total variance, a population with a more homogenous environment will provide a relatively higher estimate [Khoury et al., 1993, pp. 203].

Although family aggregation and a moderate degree of heritability of lung function have been found in most previous studies, specific genetic mechanisms have not been clear. Based on the data of the Johns Hopkins study of COPD, Rybicki et al. [1990] carried out segregation analysis of lung function by using class A regressive models. They provided statistical evidence for single-locus genetic control or a cluster of genes working in unison in the determination of FEV_1 under Mendelian inheritance in the COPD families. However, no familial correlations were found in the families of nonpulmonary patients, suggesting substantial etiologic heterogeneity in the control of lung function between the COPD families and the families without COPD [Rybicki et al., 1990]. The reasons for the lack of familial correlations in the families of those nonpatients and its discrepancy with other studies were not discussed in the report [Rybicki et al., 1990].

We used the class D regressive model for segregation analysis, which is characterized by equal sibling–sibling correlations. The class A model is simpler and makes a further assumption that the siblings are correlated only because of common parentage. This restriction, in the absence of a major gene, may lead to false inference of a major gene [Demenais and Bonney, 1989]. However, the computation of the likelihood in the class D model is more time consuming [Demenais et al., 1990]. For a quantitative trait, Demenais and Bonney [1989] demonstrated that the class D regressive model (by using SAGE program) has been shown to be mathematically and numerically equivalent to the mixed model (by using POINTER program), which specifies major gene effects and which partitions the residual variance into polygenic and environmental components.

For FEV_1 in our study, the "mixed" model, which includes both a single locus and polygenic components, had a better fit than either the single-locus or polygenic model. When further examining the transmission, we found that both a Mendelian "mixed" model and a model which included a nontransmitted environmental factor and polygenic loci gave adequate description for FEV_1. The former has a moderately higher value of AIC than does the latter. For MMFR, the polygenic model had the best fit.

FEV_1 and MMFR are most commonly used as indicators of airway obstruction. If they share similar genetic mechanisms, lung function is more likely to be controlled mainly by multiple loci, namely many independent genes each contributing in a additive fashion, and/or common environmental factors. However, etiologic heterogeneity for FEV_1 could also be possible. There may exist two major forms: a major gene and a major environmental factor plus polygenic loci. Nevertheless, it is not

easy to group these nuclear families into biologically and pathogenetically meaningful entities in this study.

There is a discrepancy between genetic influences on lung function and genetic determinants of lung disease or dysfunction. As the determinants of a function so complicated as expired flow rates in the human being, as measured by FEV_1 and MMFR, likely consist of multiple factors including airway size, elastic recoil properties of the lung, and developmental perspectives, it seems reasonable that multiple genetic influences controlled by many loci could be involved. However, it is possible that the departure from "normal" lung function, which in the extreme form is recognized as a disease, could be mediated by a single or fewer genetic influences, depending upon the disease process. An example of the latter would be airway obstruction associated with α_1-antitrypsin deficiency [Cohen, 1980] or, possibly, some aspects of the emerging understanding of the development of asthma [Marsh et al., 1993].

Identical twin studies have suggested that genetic factors are important in determining susceptibility of airways to cigarette smoke [Webster et al., 1979; Hankins et al., 1982]. In our study, no effects of smoking on lung function were observed in the children. The smoking effects on parental lung function had been adjusted at the first stage of the analysis. Other environmental factors, including passive smoking [Kauffmann et al., 1989; Tager et al., 1976] and farming exposure [Chen et al., 1991a], could affect lung function. Regressive models for quantitative traits allow simultaneous estimation of the parameters of genetic components and covariate effects. When passive smoking and farming exposure were included as covariates, the major results had no significant changes.

The importance of statistical power to detect major gene effects requires consideration. Power is related to gene effect size and sample size [MacLean et al., 1975]. In the present study, the average family size is relatively small. Nuclear families with larger sibships are generally more informative [Khoury et al., 1993, pp. 279]. A recent study, however, indicates that power is driven to a larger extent by the total number of subjects rather than sibship size per se [Borecki et al., 1994]. Since there is no single nongenetic alternative model, relatively little is known about the power of segregation analysis [Khoury et al., 1993, pp. 279]. In the present study, it may be a concern from a statistical point of view that there is lack of power in discriminating the Mendelian "mixed" model from the environmental model for FEV_1, but clinically, it is reasonable to assume that FEV_1 shares similar mechanisms with MMFR for which the single polygenic model fits as well as the Mendelian "mixed" model.

In summary, our study suggests that there are family aggregations of lung function tests, including FEV_1 and MMFR, which are most likely to be controlled by many loci with no major effects or are due to common environmental factors. However, the heterogeneity merits further investigation.

ACKNOWLEDGMENTS

The authors thank D. Still (Mayor), W. Herman (Town Administrator), W. German, N. Trach, J. Hergott (Co-Chairpersons of the local organizing committee), M. Gillis-Cipywnyk, L. Lockinger, L. Hagel, and all of the volunteers who provided canvassing and other field work in the study.

This work was supported by grant from the Saskatchewan Health Services Utilization and Research Commission. This research was supported in part by the National Health Research and Development Program through a National Health Research Scholar award to Yue Chen. Some of the results were obtained by using the SAGE program package, which was supported by U.S. Public Health Service Resource grant RR03655 from the Division of Research Resources.

REFERENCES

Akaike H (1974): A new look at the statistical model identification. IEEE Trans Automat Control AC-19:719–723.

American Thoracic Society (1987): Standardization of spirometry—1987 update. Am Rev Respir Dis 136:1285–1298.

Astemborski JA, Beaty TH, Cohen BH (1985): Variance components analysis of forced expiration in families. Am J Med Genet 21:741–753.

Beaty TH, Liang KY, Seerey S, Cohon BH (1987): Robust inference for variance components models in families ascertained through probands. II. Analysis of spirometric measures. Genet Epidemiol 4:211–221.

Bonney GE (1984): On the statistical determination of major gene mechanisms in continuous human traits: regressive models. Am J Med Genet 18:731–749.

Borecki IB, Province MA, Rao DC (1994): Power of segregation analysis for detection of major gene effects on quantitative traits. Genet Epidemiol 11:409–418.

Chen Y, Horne SL, Dosman JA (1991a): Increased susceptibility to lung dysfunction in female smokers. Am Rev Respir Dis 143:1224–1230.

Chen Y, Horne SL, McDuffie HH, Dosman JA (1991b): Combined effect of grain farming and smoking on lung function and the prevalence of chronic bronchitis. Int J Epidemiol 20:416–423.

Cohen BH (1980): Chronic obstructive pulmonary disease: a challenge in genetic epidemiology. Am J Epidemiol 112:274–288.

Cotch MF, Beaty TH, Cohen BH (1990): Path analysis of familial resemblance of pulmonary function and cigarette smoking. Am Rev Respir Dis 142:1337–1343.

Coultas DB, Hanis CL, Howard CA, Skipper BJ, Samet JM (1991): Heritability of ventilatory function in smoking and nonsmoking New Mexico Hispanics. Am Rev Respir Dis 144:770–775.

Demenais FM, Bonney GE (1989): Equivalence of the mixed and regressive models for genetic analysis. I. Continuous traits. Genet Epidemiol 6:597–617.

Demenais FM, Murigande C, Bonney GE (1990): Search for faster methods of fitting the regressive models to quantitative traits. Gene Epidemiol 7:319–334.

Devor EJ, Crawford MH (1984): Family resemblance for normal pulmonary function. Ann Hum Biol 11:439–448.

Dosman JA, Cotton DH, Graham BL, Hall D, Li R, Froh F, Barnett GD (1981): Sensitivity and specificity of early diagnostic tests of lung function in smokers. Chest 79:6–11.

Ghio AJ, Crapo RO, Elliott CG, Adams TD, Hunt SC, Jensen RL, Fisher AG, Afman GH (1989): Heritability estimates of pulmonary function. Chest 96:743–746.

Hankins D, Drage C, Zamel N, Kronenberg R (1982): Pulmonary function in identical twins raised apart. Am Rev Respir Dis 125:119–121.

Higgins M, Keller J (1975): Familial occurrence of chronic respiratory disease and familial resemblance in ventilatory capacity. J Chronic Dis 28:239–251.

Horne SL, Chen Y, Cockcroft DW, Dosman JA (1992): Risk factors for reduced pulmonary function in women: a possible relationship between Pi phenotype, number of children, and pulmonary function. Chest 102:158–163.

Horne SL, Tennent R, Lovegrove, A, Dosman JA, Cockcroft DW (1984): Pi type MZ and an increased risk of Pneumonia. Clin Invest Med 7:85–88.

Hubert HB, Fabsitz RR, Feinleib M, Gwinn C (1982): Genetic and environmental influences on pulmonary function in adult twins. Am Rev Respir Dis 125:409–415.

Kauffmann F, Tager IB, Muñoz A, Speizer FE (1989): Familial factors related to lung function in children aged 6–10 years. Am J Epidemiol 129:1289–1299.

Kawakami Y, Shida A, Yamamoto H, Yoshikawa T (1985): Pattern of genetic influence on pulmonary function. Chest 87:507–511.

Khoury MJ, Beaty TH, Tockman MS, Self SG, Cohen BH (1985): Familial aggregation in chronic obstructive pulmonary disease: use of the log linear model to analyze intermediate environmental and genetic risk factors. Genet Epidemiol 2:155–166.

Khoury MJ, Beaty TH, Cohen BH (eds) (1993): "Fundamentals of Genetic Epidemiology." New York: Oxford University Press.

Lebowitz MD, Knudson RJ, Burrows B (1984): Family aggregation of pulmonary function measurements. Am Rev Respir Dis 129:8–11.

Lewitter FI, Tager IB, McGue M, Tishler PV, Speizer FE (1984): Genetic and environmental determinants of level of pulmonary function. Am J Epidemiol 120:518–530.

MacLean CJ, Morton NE, Lew R (1975): Analysis of family resemblance. IV. Operational characteristics of segregation analysis. Am J Hum Genet 27:365–384.

Man SFP, Zamel N (1976): Genetic influence on normal variability of maximum expiratory flow-volume curves. J Appl Physiol 41:874–877.

Marsh DG, Lockhart A, Holgate SJ (eds) (1993): "The Genetics of Asthma." Blackwell Scientific, Oxford.

Norusis MJ (1993): "SPSS for Windows: Base System User's Guide, Release 6.0." Chicago: SPSS Inc.

Redline S, Tishler PV, Lewitter FI, Tager IB, Muñoz A, Speizer FE (1987): Assessment of genetic and nongenetic influences on pulmonary function. Am Rev Respir Dis 135:217–222.

Redline S, Tishler PV, Rosner B, Lewitter FI, Vandenbrugh M, Weiss ST, Speizer FE (1989): Genotypic and phenotypic similarities in pulmonary function among family members of adult monozygotic and dizygotic twins. Am J Epidemiol 129:827–836.

Rybicki BA, Beaty TH, Cohen BH (1990): Major genetic mechanisms in pulmonary function. J Clin Epidemiol 43:667–675.

Schilling RSF, Letai AD, Hui SL, Beck GJ, Schoenberg JB, Bouhuys A (1977): Lung function, respiratory disease, and smoking in families. Am J Epidemiol 106:274–283.

Statistical Analysis for Genetic Epidemiology (SAGE) (1994): Release 2.2. Department of Biometry and Genetics, Louisiana State University Medical Center, New Orleans.

Tager IB, Rosner B, Tishler PV, Speizer FE, Kass EH (1976): Household aggregation of pulmonary function and chronic bronchitis. Am Rev Respir Dis 114:485–492.

U.S. Department of Health and Human Services (1984): "The Health Consequences of Smoking: Chronic Obstructive Lung Disease." Washington, DC: USGPO, DHHS (PHS) 84-50205.

Webster PM, Lorimer EG, Man SFP, Woolf CR, Zamel N (1979): Pulmonary function in identical twins: comparison of nonsmokers and smokers. Am Rev Respir Dis 119:223–228.