A Possible Smoking Susceptibility Locus on Chromosome 11p12: Evidence from Sex-limitation Linkage Analyses in a Sample of Australian Twin Families

Katherine I Morley,[1,2,5] Sarah E Medland,[1,3] Manuel AR Ferreira,[1] Michael T Lynskey,[1,4] Grant W Montgomery,[1] Andrew C Heath,[4] Pamela AF Madden,[4] and Nicholas G Martin[1]

Many twin studies have identified sex differences in the influence of genetic and environmental factors on smoking behaviors. We explore the evidence for sex differences for smoking initiation and cigarette consumption in a sample of Australian twin families, and extend these models to incorporate sex differences in linkage analyses for these traits. We further examine the impact of including or excluding non-smokers in genetic analyses of tobacco consumption. Accounting for sex differences improved linkage results in some instances. We identified one region suggestive of linkage on chromosome 11p12. This locus, as well as another region identified on chromosome 6p12, replicates regions identified in previous studies.

KEY WORDS: Linkage; sex differences; smoking; twin study.

INTRODUCTION

Tobacco use is one of the major causes of global disease burden, increasing the risk of various cancers, chronic obstructive pulmonary disease, respiratory diseases, and vascular diseases (Ezzati *et al.*, 2002). Ezzati and Lopez (2003; 2004) have estimated that for the year 2000, 4.83 million premature deaths were caused by smoking-related diseases world-wide.

Despite these known health risks, approximately 30% of the global population continues to smoke, and this figure may increase by up to 35% by 2020 (Corrao *et al.*, 2000). Clearly better methods for preventing smoking initiation and increasing successful smoking cessation are needed, and this can be aided by a better understanding of the genetic and environmental factors that influence smoking behavior.

Twin studies suggest that genes exert a moderate to strong influence on various aspects of smoking behavior measured in adults such as smoking initiation and cigarette consumption, with heritability estimates ranging from 0.33 to 0.86 (Sullivan and Kendler, 1999). Sex differences in genetic and environmental influences have been identified for some phenotypes, particularly smoking initiation and consumption. Analyses of Australian, Swedish and Finnish adult twin samples have found differences in the magnitude of genetic and environmental influences on smoking initiation for males and females (Heath and Martin, 1993; Madden *et al.*, 1999, 2004). Hettema and colleagues (1999) fitted a general

[1] Genetic Epidemiology, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane, Queensland, Australia.
[2] Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia.
[3] Department of Psychology, The University of Queensland, Brisbane, Queensland, Australia.
[4] Department of Psychiatry, Washington University St. Louis, St Louis, MO, USA.
[5] To whom correspondence should be addressed at Genetic Epidemiology, Queensland Institute of Medical Research, PO Royal Brisbane Hospital, Brisbane, Queensland, 4029, Australia. Tel.: +61-7-3362-0228; Fax: +61-7-3362-0101; e-mail: kateM@qimr.edu.au

sex-limitation model to data on tobacco consumption in adult twins from the Virginia Twin Registry and found evidence of unique additive genetic effects for women, accounting for approximately 30% of the variance. Using a sample of Swedish twins reared together and apart, Kendler *et al.* (2000) found that regular tobacco use in men was influenced primarily by additive genetic factors. In contrast, correlated environment (that which influences both twins reared apart and twins reared together e.g. social influences on tobacco use) was found to have the strongest influence in women. The results of these studies suggest that the etiology of tobacco use may differ in males and females. A meta-analysis of twin studies of adult smoking behavior found heritability estimates for smoking initiation of 0.37 for men, 0.55 in women, and estimates for smoking persistence (which included consumption), of 0.59 for men and 0.46 for women (Li *et al.*, 2003a). The authors also found that additive genetic and shared environmental influences differed significantly for males and females, but that unique environmental influences did not.

However, two recent studies of adult twins found no evidence to support sex differences for smoking behavior. Examining initiation, regular tobacco use, and nicotine dependence in adult twins from the Virginia Twin Registry, Maes and colleagues (2004) found no differences between the sexes in genetic or environmental influences on initiation, regular tobacco use or nicotine dependence. Vink *et al.* (2004), analyzing smoking initiation and maximum number of cigarettes smoked in a sample of Dutch adult twins, found that models in which variance component estimates were constrained across sexes provided the best fit for both phenotypes. Although sex differences in genetic and environmental influences have not been found in all samples there is some evidence that smoking initiation and consumption may follow slightly different paths in males and females. These differences may have relevance not only for twin studies, but for linkage analyses as well.

A small but growing number of linkage analyses have been conducted for smoking behavior. As yet, only one genome scan for nicotine dependence has been published (Straub *et al.*, 1999; Sullivan *et al.*, 2004). Other studies have focused on more easily obtained measures of smoking behavior such as smoking initiation, and tobacco consumption, which can be viewed as a proxy measure for nicotine

dependence (Sullivan and Kendler, 1999). Linkage results have been found on almost all chromosomes, but only a handful have reported LOD scores above or close to genome-wide significance (Lander and Kruglyak, 1995). Regions on chromosome 2 (Bierut *et al.*, 2004; Goode *et al.*, 2003), chromosome 6 (Duggirala *et al.*, 1999; Vink *et al.*, 2004), chromosome 9 (Bierut *et al.*, 2004; Gelernter *et al.*, 2004; Li *et al.*, 2003b) and chromosome 11 (Bierut *et al.*, 2004; Goode *et al.*, 2003; Li *et al.*, 2003b) appear to be the most promising, having been identified in more than one study, and providing significant or suggestive evidence for linkage. However, no studies have yet assessed whether accounting for the sex-differences in genetic influences observed in twin studies may have an impact upon the results of linkage analyses.

In light of previous work on smoking behavior, the current study had a number of aims. Firstly, we aimed to investigate the genetic and environmental influences on smoking initiation and cigarette consumption, and to examine the evidence for a difference in the magnitude of these influences between the sexes. Secondly, we aimed to investigate the possibility that quantitative trait loci (QTL) involved in smoking behavior may exert stronger influences on one sex over the other. Thirdly, we explored whether including and excluding non-smokers in cigarette consumption analyses had any impact on the results of linkage analyses.

We expected to find that the strength of genetic and environmental influences on smoking initiation and cigarette consumption varied between males and females. Thus we also expected that including separate QTL parameters for males and females in our linkage model would provide different results to a model constraining the QTL to have the same effects in males and females. Saccone *et al.* (2003) investigated the impact on linkage results of including or excluding non-smokers as the lowest category of a consumption measure, finding that the two measures yielded completely different linkage results. We expected that including non-smokers in cigarette consumption analyses would provide somewhat different (although not entirely different) results to excluding non-smokers, as the underlying liability distributions for these two variables are related. Thus we have chosen to present univariate models for smoking initiation, cigarette consumption including non-smokers, and cigarette consumption excluding non-smokers to allow for comparison.

METHODS

Subjects

Phenotypic measures of smoking behavior were taken from four complementary self-report Health and Lifestyle questionnaires that were administered to adult twins and their families recruited from the Australian Twin Registry. These questionnaires were mailed to all twins as well as the parents, siblings, spouses and children of the twins, between 1989 and 1993. The total sample consists of 5321 twin pairs, 986 single twins, 3715 siblings, 1660 children of twins, 3570 parents of twins, and 2444 spouses of twins. Zygosity was determined using standard questionnaire methods, validated by genotyping subsets of twins. More detailed descriptions of these studies can be found elsewhere (Heath and Martin, 1994; Kirk et al., 2000; Lake et al., 2000).

Phenotypic measures

The self-report questionnaires contained a number of questions about the smoking behavior of the individual participant and their family members, along with other health and lifestyle, and personality questions. Participants were asked to classify themselves and other family members as non-smokers, ex-smokers, or current smokers. Participants were also asked to report their average number of cigarettes consumed per day, either currently or when they previously smoked if an ex-smoker (scaled as never smoker, 1–4 cigarettes, 5–10, 11–20, 21–40, 40 +); the age at which they began smoking, age at which they had successfully quit smoking (if they were an ex-smoker), and the total number of years for which they had been a smoker. Responses to the different questions were used to verify participants' answers and ensure internal consistency.

Three measures of smoking behavior were calculated from the questionnaire data obtained from twins and their siblings. Smoking initiation (SI) was defined as ever versus never having been a regular smoker. Individuals who reported themselves as ex- or current smokers and/or reported consuming an average of 1 or more cigarettes per day were considered regular smokers. Two different measures of cigarette consumption were analyzed, to further examine how different classifications of cigarette consumption may influence results. The first consumption measure incorporated all six of the categories for cigarette consumption included in the questionnaire (cigarette consumption including non-smokers; CCI). The second measure excluded those individuals who reported themselves to be non-smokers (cigarette consumption excluding non-smokers; CCE).

Assumption testing

All three categorizations of smoking behavior were analyzed as ordinal phenotypes, assuming an underlying normal distribution of liability (Falconer and Mackay, 1996; Heath and Martin, 1993). Prior to univariate model fitting, some basic assumptions regarding the data were formally tested. For each phenotype, a basic model incorporating thresholds and calculation of tetrachoric or polychoric correlations was fitted to the data by the method of maximum likelihood as implemented in Mx (Neale et al., 2003). Correlations were allowed to vary for each zygosity group, as were thresholds, which were also permitted to vary between siblings. The threshold model included regression terms for sex, year of birth, and an interaction between the two, to allow for different relationships between the smoking variables and year of birth for each sex. Hypotheses regarding the homogeneity of thresholds and correlations within and across zygosity groups were tested by progressively equating model parameters until the fit of the model worsened significantly, as measured by the likelihood-ratio chi-squared test (Neale and Cardon, 1992). Due to the large number of tests being conducted, a significance level of $\alpha = 0.01$ was used. The appropriateness of including regression terms in the threshold model was also tested by successively dropping each term from the model and examining the difference in overall model fit.

Biometrical model fitting

Biometrical modeling involved estimating the components of variance in liability to each trait. Under the standard assumptions of the classical twin design, variance in liability to a particular trait may be decomposed into: additive genetic influences (σ_A^2), which are correlated at 1 for MZ twin pairs and 0.5 for DZ twin pairs; non-additive (or dominant) genetic effects (σ_D^2), correlated 1 for MZ pairs and 0.25 for DZ pairs; environmental influences shared by members of a family (σ_C^2), which are correlated at 1 for members of both MZ and DZ pairs; and environmental influences unique to each family member (σ_E^2) which are by definition uncorrelated. However, σ_C^2 and σ_D^2 are confounded in analyses consisting of only

twins reared together, and thus only one of these parameters can be estimated from the model (Grayson, 1989; Hewitt, 1989).

A general sex-limitation model including an additional additive genetic component for males was fitted to the data as a saturated model for all phenotypes (Medland, 2004; Neale and Cardon, 1992; Neale *et al.*, 1994). Under this model, the magnitude of genetic and environmental influences is permitted to differ for males and females, and an additional additive genetic (or non-additive genetic, or shared environmental) parameter specific to one sex ($\sigma_{A'}^2$) is estimated. A nested sub-model, the common effects sex-limitation model, can be fitted by dropping the $\sigma_{A'}^2$ parameter. Comparing the difference in fit between these two models provides an indirect test of whether the correlation for additive genetic effects between males and females is less than 0.5.

The standard twin design can be extended to include information from additional non-twin siblings, and non-twin sib-ships are parameterized as for DZ twins (Posthuma *et al.*, 2003). Thus the variance–covariance matrix may be extended to incorporate as many siblings as necessary. When siblings are included, the presence of a twin-specific environmental influence on a trait can be assessed by including an additional environmental variance component that is shared only by twin pairs (σ_T^2) (Koeppen-Schomerus *et al.*, 2003). Including at least one additional sibling in analyses provides increased statistical power (Posthuma and Boomsma, 2000). As our genotyped sample contained few sib-ships with smoking data that were larger than a trio, there were some phenotype categories for which there were not enough data available to include a second additional sibling. Consequently, genetic model fitting and linkage analyses were conducted with a maximum of one extra sibling.

Univariate models were fitted to the data using Mx. The saturated general sex-limitation model, including a male-specific additive genetic parameter and a female twin-specific environmental parameter, was applied to each phenotype. Hypotheses concerning genetic and environmental influences were tested by fitting sub-models that involved removing or equating model parameters. A general sex-limitation model incorporating a male-specific common environmental parameter was also fitted for each phenotype. The relative fit of the different models was evaluated using the likelihood-ratio chi-square test and the Akaike information criterion (AIC), a measure of model fit relative to model parsimony.

For each phenotype, the best fitting model was identified as the model with a non-significant likelihood-ratio chi-squared test and a negative AIC.

Linkage analyses

For all phenotypes, univariate variance components linkage analyses were performed using Mx. Fitting a variance components linkage model to data from DZ twins and their siblings, the variance between DZ twins and their siblings can be partitioned into the effects of the putative QTL (σ_Q^2), residual shared factors (σ_S^2) and unique environmental factors (σ_E^2). To allow for the partitioning of σ_S^2 into the residual additive genetic component (σ_A^2) and common environmental component (σ_C^2), we included 2018 ungenotyped MZ pairs in our model. For this model, the correlation for σ_Q^2 for MZ pairs is fixed at 1, while the correlation between DZ twins and siblings is specified as $\hat{\pi}$ an estimate of the proportion of alleles shared identical by descent (IBD) at a given marker locus (calculated as $\hat{\pi} = 1/2\, p_{IBD1} + p_{IBD2}$, where p_{IBD1} and p_{IBD2} are the probabilities that two relatives share 1 and 2 alleles IBD respectively) (Neale, 1998; Sham, 1998). MERLIN version 0.10.1 (Abecasis *et al.*, 2002) was used to calculate multipoint estimates of the probabilities of sib-pairs sharing 0, 1 or 2 alleles IBD using a 5 cM grid to maximize the inheritance information available for analysis.

The biometrical models previously fitted for each phenotype were extended to incorporate separate QTL parameters for males and females (Medland, 2005). For all three phenotypes, the QTL parameters were incorporated in the full common-effects sex-limitation models. Female twin-specific environmental parameters were included for SI and CCI, as this parameter could not be removed from the model without a significant loss of fit for these phenotypes. A sub-model in which the QTL parameters for males and females were constrained to be equal was also fitted.

Both models were compared to a model in which male and female QTL parameters were set to zero. The difference in fit of the models was tested using the likelihood-ratio chi-squared test (Neale, 1998). For the model with separate QTL parameters for males and females, the distribution of the test statistic is approximately $1/4\chi_0^2 : 1/2\chi_1^2 : 1/4\chi_2^2\, 1/2\chi_2^2$, while the distribution for the model in which the QTL parameters are equated across the sexes is approximately $1/2\chi_0^2 : 1/2\chi_1^2$ (Medland, 2005). Thus the results

for these models are reported as *p*-values to facilitate comparison between the two. Significant linkage was measured by a *p*-value of 2.2×10^{-5}, suggestive as a *p*-value of 7.4×10^{-4}, and replication of an existing finding as a *p*-value of 0.01, in keeping with Lander and Kruglyak (1995).

Genotyping and data cleaning

The genotypic data used for this research represents a compilation of data from four genome scans undertaken for particular phenotypic studies at the Queensland Institute of Medical Research. Recruitment of participants for these studies was primarily based upon participant involvement in previous phenotype collection studies. As a result, genotypic data for a subset of the individuals who participated in the Health and Lifestyle questionnaires were available. Details of genotyping and data cleaning are discussed briefly below. For more detailed information regarding the individual studies, DNA collection, genotyping and data cleaning methods see Cornes *et al.* (2005).

Pedigree structures for each scan were examined using Graphic Representation of Relationships (GRR)(Abecasis *et al.*, 2001) and RELPAIR version 2.0 (Epstein *et al.*, 2000; Duren *et al.*, 2003) to identify inconsistencies between the genotypic data and self-reported pedigree relationships. Once discrepancies were resolved, data for the four scans were merged, then checked again for pedigree errors using GRR, and for Mendelian inconsistencies using SIB-PAIR version 0.99.9 (Duffy, 2002). The combined genome scan data included 458 unique markers that were typed in two or more scans, which were included separately on the genetic map for the scan, separated by 0.001 cM. The consistency of genotype information from different scans for these 458 markers was checked via comparison of the allele calls between different scans. Additionally the number of Mendelian errors produced by each set of marker genotypes, and whether the marker was in Hardy–Weinberg equilibrium, were also assessed. Markers with genotypic data inconsistent between different genome scans were removed from further analysis.

Unlikely genotypes were identified and removed using MERLIN. GENEHUNTER version 2.1 r5 beta (Kruglyak *et al.*, 1996) and MENDEL (Lange *et al.*, 1988) were used to identify and examine potential map errors. Map positions were in Kosambi cM, estimated via locally weighted linear regression from the NCBI Build 34.3 physical map positions and published Decode and Marshfield genetic map positions (Kong *et al.*, 2004). This map was converted to Haldane cM prior to MERLIN analyses. Where the results suggested inconsistencies between genetic map distances and recombination fractions, the primer sequences for all markers in the region were blasted against the entire human genome sequence (http://www.ensembl.org, NCBI Build 34.3). The genetic map was then revised to include the updated physical positions of all markers in problematic regions, with new map distances interpolated as above. Using the revised map, the original genotype data (with inconsistent duplicated markers removed) were cleaned of unlikely genotypes using MERLIN, and map errors were resolved using GENEHUNTER.

The cleaned genome scan data includes 1770 autosomal markers, of which 394 are duplicates, leaving a total of 1376 unique markers. The mean intermarker distance for all sib-pairs in the sample was 7.1 cM. The combined genome scan includes a total of 5160 individuals from 1587 families. Linkage analyses were performed on individuals with at least 198 markers. This represents 90% of the number of markers in the smallest scan, and is close to the median for that scan (196.5). The average intermarker distance for these sib-pairs was 6.1 cM, and the mean information content was 0.58. For SI, there were a total of 1646 individuals from 772 families with at least two siblings available for analysis. There were 559 individuals from 246 families available for CCE, and 1603 individuals from 751 families available for CCI.

RESULTS

Sample demographics

Smoking measures were available for individuals from 6257 families from the entire sample. The number of twins and siblings are shown in Table I.

Participant year of birth ranged from 1902 to 1974, with a mean age at survey of 34 years (18–87 years; S.D. = 16.7 years). Females made up 60.8% of the analyzed sample. 55.2% of women reported themselves to be non-smokers, compared to 49% of men. Proportions were similar for the linkage sub-sample; 62.1% of participants were female, with a slightly higher proportion of female non-smokers (56.9%) and a slightly lower proportion of male non-smokers (47%). Women were generally lighter smokers than men. Of the female current and ex-smokers, 50.3% reported smoking an average of 1 to 10 cigarettes per day; 33.0% reported smoking 11 to

Table I. Composition of Families in the Analyzed Sample

Zygosity	Family Structure						Total
	●●	●●♀	●●♂	●○	●○♀	●○♂	
MZ Female	1102	292	189	168	12	13	1776
MZ Male	587	129	105	156	4	6	987
DZ Female	731	158	138	160	11	9	1207
DZ Male	418	78	68	159	12	11	746
DZ OS							
Female	466	118	99	174	14	17	888
Male	397	71	63	136	9	7	653

Structures for the 6257 families included in the dataset are shown for monozygotic (MZ) and dizygotic (DZ) families with same-sex and opposite-sex (OS) twins, as well as the additional sibling. A black circle (●) indicates the twin was included in the sample, a white circle (○) indicates the twin was not included or did not participate. Male (♂) or female (♀) symbols indicate the presence and gender of the additional non-twin sibling.

20 cigarettes; and 16.7% reported smoking 21 or more cigarettes per day. In contrast, only 36.7% of male current and ex-smokers reported smoking 1 to 10 cigarettes per day; 36.5% reported smoking 11 to 20 cigarettes; and 26.8% reported smoking 21 or more.

Assumption testing

Thresholds for all measures could be equated across same-sex zygosity groups without a significant reduction in model fit (SI:$\chi_6^2 = 10.7$, $p = 0.1$; CCE:$\chi_{24}^2 = 36.92$, $p = 0.05$; CCI:$\chi_{30}^2 = 3.93$, $p = 0.02$). For SI and CCI, the thresholds for same-sex female twins could be equated to those of the female DZ opposite-sex twins (SI:$\chi_1^2 = 0.98$, $p = 0.32$; CCI:$\chi_5^2 = 6.40$, $p = 0.27$), but this was not the case for the male DZ opposite-sex twins (SI:$\chi_1^2 = 18.91$, $p < 0.001$; CCI:$\chi_5^2 = 29.26$, $p < 0.001$). This was due to the higher prevalence of smoking amongst the DZ opposite-sex twin males. CCE thresholds for all twins could be equated for each sex ($\chi_8^2 = 10.36$, $p = 0.24$), but the sex-specific thresholds could not be equated without worsening model fit significantly ($\chi_4^2 = 15.06$, $p < 0.01$). Sex-specific sibling thresholds could be equated to sex-specific twin thresholds without a significant loss of model fit for SI and CCE (SI:$\chi_2^2 = 5.04$, $p = 0.08$; CCE:$\chi_8^2 = 19.76$, $p = 0.01$), but not for CCI ($\chi_{10}^2 = 25.85$, $p < 0.01$).

For both SI and CCI, the interaction term for year of birth and sex could not be dropped from the model without a significant loss of fit (SI:$\chi_1^2 = 116.09$, $p < 0.001$; CCI:$\chi_1^2 = 93.30$, $p < 0.001$), but could be dropped from the model for CCE ($\chi_1^2 = 0.77$, $p = 0.38$). However, the year of birth term could not subsequently be removed from the CCE model without a significant loss of fit ($\chi_1^2 = 50.91$, $p < 0.001$).

For all variables, twin correlations for the MZ, and same-sex DZ pairs could be equated, but the correlations of the opposite-sex DZ twins were significantly lower than those of the same-sex DZ twins (SI:$\chi_1^2 = 7.18$, $p < 0.01$; CCE:$\chi_1^2 = 6.98$, $p < 0.01$; CCI:$\chi_1^2 = 7.04$, $p < 0.01$). This difference in correlations between the DZ same-sex and opposite-sex groups suggests the presence of non-scalar sex-limitation for these phenotypes. Twin-sibling correlations were not influenced by the zygosity of the twin. The same-sex male and opposite-sex twin-sibling correlations could be equated to the respective DZ twin-twin correlations without a significant loss of model fit for all phenotypes. However, the same-sex female twin-sibling correlation could not be equated to the DZ female correlation without a significant loss of fit (SI:$\chi_1^2 = 29.98$, $p < 0.001$; CCE:$\chi_1^2 = 6.78$, $p < 0.01$; CCI:$\chi_1^2 = 16.93$, $p < 0.001$). This suggests a stronger environmental correlation for female twin-twin pairs as opposed to female twin-sibling pairs, or the presence of gene × age or environment × age effects.

Biometrical model fitting

Based upon the assumption testing results presented above, general sex-limitation models estimating unique additive genetic factors for males and a female twin-specific environmental factor were fitted for each phenotype. Thresholds were specified separately for males and females, with separate thresholds for the male DZ opposite-sex twins included for SI and CCI. Separate twin and sibling thresholds were used for CCI. A general sex-limitation model including a unique common environmental parameter for males was also fitted for each phenotype. However, dropping the sex-specific common

environmental parameter did not result in a significant loss of fit for any of the phenotypes (SI:$\chi_1^2 = 2.01$, $p = 0.16$; CCE:$\chi_1^2 = 0.47$, $p = 0.49$; CCI:$\chi_1^2 = 0.66$, $p = 0.42$). Consequently, all model fitting was conducted using the general sex-limitation model including a sex-specific additive genetic component as the saturated model. Model fitting statistics for each phenotype are shown in Table II.

For all phenotypes, unique additive genetic effects for males only accounted for a small proportion of the total variance and could be dropped from the model without a significant loss of fit. The female twin-specific environmental component could not be dropped from the common effects sex-limitation models for SI or CCI without a significant loss of fit (SI:$\chi_1^2 = 12.35$, $p < 0.01$; CCI:$\chi_1^2 = 26.93$, $p < 0.01$). However, dropping this component from the CCE model did not significantly worsen model fit.

The male common environmental parameter could be dropped from the models for all three phenotypes without a significant loss of fit. The female common environmental parameter could also be dropped from the models for SI and CCI without a significant worsening of fit, but dropping the parameter from the CCE model significantly worsened model fit ($\chi_1^2 = 5.53$, $p = 0.02$). The unique environmental parameters for males and females

could be equated without worsening the fit of the model for all phenotypes. A model in which the additive genetic parameters were equated instead of the unique environmental parameters did not provide as good a fit to the data for any of the phenotypes (SI:$\chi_1^2 = 8.29$, $p < 0.05$; CCE:$\chi_1^2 = 1.89$, $p = 0.17$ CCI:$\chi_1^2 = 30.98$, $p < 0.01$).

For all phenotypes the best-fitting model was one in which σ_E^2 was equated across the sexes. For SI the best-fitting model included σ_A^2, σ_T^2 and σ_E^2 parameters for females and σ_A^2 and σ_E^2 parameters for males. For CCE the best-fitting model incorporated σ_A^2, σ_C^2 and σ_E^2 parameters for females, σ_A^2 and σ_E^2 parameters for males. The best-fitting model for CCI included σ_A^2, σ_T^2 and σ_E^2 parameters for females, σ_A^2 and σ_E^2 parameters for males. The proportion of variance explained by genetic and environmental factors for the general sex-limitation, common-effects sex-limitation and best-fitting models for each phenotype are shown in Table III.

Linkage analyses

Variance component estimates and phenotype prevalences for the genotyped sample were comparable to those obtained for the full sample. Genome-

Table II. Model Fitting Statistics for each Phenotype

Phenotype	Model number	Model	Model tested against	χ^2	Δ d.f.	p value	AIC
Smoking initiation	1	Saturated		*16488.25*	*13023*		
	2	Drop $\sigma_{A'm}^2$	1	0	1	1	−2.00
	3	Drop σ_{Tf}^2	2	12.35	1	<0.01	10.35
	4	Drop σ_{Cf}^2	2	0.05	1	0.83	−1.95
	5	Drop σ_{Cm}^2	4	0.96	1	0.33	−1.04
	6	**Equate σ_E^2**	**4**	**0.78**	**1**	**0.38**	**−1.22**
	7	Equate σ_A^2	5	8.29	1	<0.05	6.29
Cigarette consumption excluding non-smokers	1	Saturated		*16382.69*	*5896*		
	2	Drop $\sigma_{A'm}^2$	1	0.42	1	0.52	−1.58
	3	Drop σ_{Tf}^2	2	1.44	1	0.23	−0.57
	4	Drop σ_{Cm}^2	3	0.08	1	0.77	−1.92
	5	Drop σ_{Cf}^2	4	5.53	1	<0.05	3.53
	6	**Equate σ_E^2**	**5**	**0.55**	**1**	**0.46**	**−1.45**
	7	Equate σ_A^2	5	1.89	1	0.17	−0.11
Cigarette consumption including non-smokers	1	Saturated		*32742.55*	*12783*		
	2	Drop $\sigma_{A'm}^2$	1	0.25	1	0.62	−1.75
	3	Drop σ_{Tf}^2	1	26.93	1	<0.01	24.93
	4	Drop σ_{Cm}^2	2	0.07	1	0.79	−1.93
	5	Drop σ_{Cf}^2	4	0	1	1	−2.00
	6	**Equate σ_E^2**	**5**	**2.93**	**1**	**0.09**	**0.93**
	7	Equate σ_A^2	5	30.98	1	<0.01	28.98

Variance components for females are denoted by $\sigma_{Af}^2, \sigma_{Cf}^2, \sigma_{Ef}^2, \sigma_{Tf}^2$ and for males by $\sigma_{Am}^2, \sigma_{Cm}^2, \sigma_{Em}^2, \sigma_{A'm}^2$. The −2 log likelihood and degrees of freedom of the saturated models for each phenotype are indicated in italics. The best-fitting sub-models are indicated in bold.

Table III. Parameter Estimates for the Two Sex-limitation Models and the Best-fitting Univariate Models for Each Phenotype

Phenotype	Model	σ^2_{Af}	σ^2_{Cf}	σ^2_{Ef}	σ^2_{Tf}	σ^2_{Am}	σ^2_{Cm}	σ^2_{Em}	$\sigma^2_{A'm}$
Smoking initiation	General sex-limitation	0.62 (0.46–0.75)	0.0 (0.0–0.10)	0.17 (0.14–0.20)	0.22 (0.10–0.33)	0.72 (0.28–0.85)	0.09 (0.0–0.26)	0.20 (0.15–0.26)	0.0 (0.0–0.38)
	Common effects sex-limitation	0.62 (0.47–0.75)	0.0 (0.0–0.07)	0.17 (0.14–0.20)	0.22 (0.09–0.34)	0.72 (0.52–0.85)	0.09 (0.0–0.26)	0.20 (0.15–0.25)	–
	Best-fitting model[a]	0.59 (0.48–0.71)	–	0.18 (0.15–0.20)		0.83 (0.80–0.85)	–	0.18 (0.15–0.20)	–
Cigarette consumption excluding non-smokers	General sex-limitation	0.28 (0.06–0.52)	0.18 (0.0–0.36)	0.46 (0.39–0.53)	0.23 (0.12–0.34)	0.11 (0.0–0.58)	0.09 (0.0–0.33)	0.51 (0.42–0.62)	0.28 (0.0–0.46)
	Common effects sex-limitation	0.34 (0.12–0.53)	0.20 (0.03–0.38)	0.46 (0.39–0.53)	0.08 (0.0–0.20)	0.50 (0.37–0.59)	0.0 (0.0–0.07)	0.50 (0.41–0.60)	–
	Best-fitting model[a]	0.33 (0.19–0.49)	0.20 (0.05–0.33)	0.47 (0.42–0.52)		0.53 (0.48–0.58)	–	0.47 (0.42–0.52)	–
Cigarette consumption including non-smokers	General sex-limitation	0.52 (0.43–0.59)	0.0 (0.0–0.06)	0.24 (0.22–0.28)	0.24 (0.15–0.32)	0.65 (0.46–0.75)	0.0 (0.0–0.14)	0.29 (0.25–0.33)	0.07 (0.0–0.28)
	Common effects sex-limitation	0.51 (0.42–0.60)	0.0 (0.0–0.0)	0.25 (0.22–0.28)	0.24 (0.15–0.32)	0.69 (0.55–0.75)	0.02 (0.0–0.13)	0.29 (0.25–0.34)	–
	Best-fitting model[a]	0.49 (0.41–0.58)	–	0.26 (0.24–0.29)		0.74 (0.71–0.76)	–	0.26 (0.24–0.29)	–

Variance components for females are denoted by σ^2_{Af}, σ^2_{Cf}, σ^2_{Ef}, σ^2_{Tf} and for males by σ^2_{Am}, σ^2_{Cm}, σ^2_{Em}, $\sigma^2_{A'm}$. 95% confidence intervals are shown in brackets. Dashes indicate the parameter was set to zero.
[a] σ^2_{Ef} constrained to equal σ^2_{Em}.

wide linkage results for all three phenotypes are shown in Figure 1.

For SI, the highest peak ($p = 3.59 \times 10^{-3}$) was found on chromosome 20p13. One additional peak above the threshold for replication ($p = 0.01$) was identified on chromosome 16p13. The highest peak for CCE ($p = 3.99 \times 10^{-3}$) was found on chromosome 11q23. Two additional peaks above the threshold for replication were identified on chromosomes 4q35 and 6p21. The highest peak for CCI was found on chromosome 11p12 ($p = 5.68 \times 10^{-5}$), which is above the threshold for suggestive linkage. Results above the threshold for replication were also found on 6p12, 7q34, 11q23-24 and 16p13. The two most significant results were found on chromosomes 6 and 11, and both chromosomes are shown in detail in Figure 2.

Table IV lists the most significant linkage results for each phenotype with the *p*-values for both models, as well as the σ^2_Q estimates.

For most peaks, the QTL effects were stronger in one sex. For example, the highest result on chromosome 11 for CCI appears to have a much stronger effect in males than in females, accounting for 68% of the variance in males as compared to 22% of variance in females. The reverse is true of the QTLs identified on chromosomes 6 and 7, which appear to have a greater influence in females.

Although analyses of CCI and CCE did produce quite different results, they are not entirely dissimilar. Overlapping peaks for the two phenotypes can be found on chromosomes 6 and 11. In both instances results above the significance level for replication were obtained from markers in the same region. The difference between the results for these two phenotypes may be due in part to the fact that the CCE analysis was conducted on a significantly smaller sample and thus had less power to detect linkage. There are also some instances where the CCE results provide more evidence for linkage than the CCI results, such as chromosome 4, although this could also indicate that the CCE result is in fact spurious.

DISCUSSION

Overall, the results from this study lend support to the presence of sex differences in the magnitude of genetic and environmental influences on smoking initiation and cigarette consumption. Heritability estimates for all phenotypes were higher for males, explaining 53% to 74% of the variance in trait liability, compared to 33% to 59% in females. These estimates for SI and CCI in women are similar to
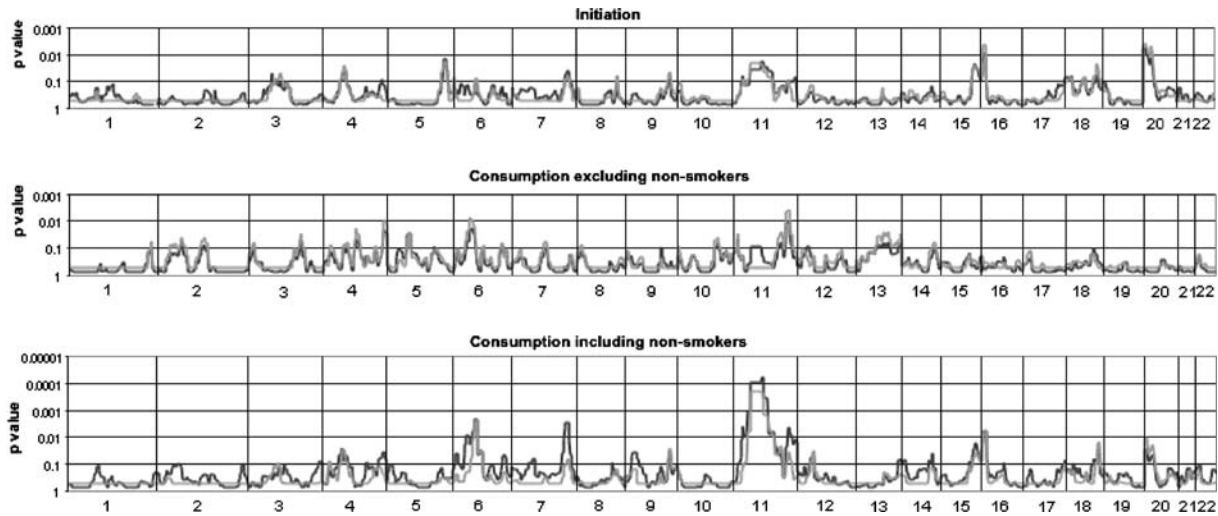
Fig. 1 Whole genome results for each phenotype. Chromosome numbers are shown across the bottom of each graph. For each phenotype, results from the model in which $\sigma^2_{Qf} \neq \sigma^2_{Qm}$ are shown in black. Results from the model in which $\sigma^2_{Qf} = \sigma^2_{Qm}$ are shown in gray.

those reported in a recent meta-analysis of twin studies of smoking behavior, while those for men are somewhat higher (Li *et al.*, 2003a). For CCE, the heritability estimate for males is similar to that found in the meta-analysis, while the estimate for females is lower. However, none of these results are incompatible with heritability estimates previously reported for this, or other samples (Heath and Martin, 1993; Kendler *et al.*, 1999; Koopmans *et al.*, 1999).

Shared environmental influences were only important for CCE in women, accounting for 20% of the variance. For males, shared environmental influences were not important for any phenotype. Unique environmental factors were found to have a strong influence on CCE (47% of the variance in both sexes). This influence was smaller for the other two phenotypes, but as has been found in other studies of smoking behavior, there was no significant difference between the unique environmental influences for SI, CCI or CCE in males and females (Li *et al.*, 2003a). In this sample we found evidence for shared environmental influences specific to female twin pairs for SI and CCI. For both phenotypes, these influences accounted for approximately 25% of the variance in trait liability. Other studies have also found evidence to suggest that smoking initiation may be influenced by environmental factors shared by twins, particularly in females (Kendler and Gardner, 1998; Pergadia *et al.*, submitted). While we have modeled this effect as a twin-specific environmental factor, it could also be the result of environment × age or gene × age effects (Eaves *et al.*, 1978), and deserves further investigation.

In light of the sex difference in additive genetic effects identified for these phenotypes in this sample, we performed linkage analyses that allowed the magnitude of the QTL effect to differ between males and females. We found that doing so increased the significance of our results as compared to a model in which the QTL effects were constrained across the sexes in some instances. Using this type of analysis also provided an indication of how strongly particular QTLs may influence traits in males and females. In this sample, a number of the QTLs identified appeared to have a stronger influence in females. However, for the most significant linkage result, on chromosome 11, the QTL identified accounted for a much greater proportion of the variance in males. This type of information may be of benefit in designing and conducting association studies of candidate genes for smoking behavior.

We found that excluding non-smokers from linkage analyses for consumption produced very different results from those produced by including non-smokers. However, unlike Saccone *et al.* (2003), we found that there was some overlap in results for the two measures. The difference in results is not surprising, given the difference in sample size and that including or excluding non-smokers will alter the underlying liability distribution. If liability to initiate smoking is not entirely independent of liability for the amount of cigarettes smoked, then excluding non-smokers from analyses will result in an artificial truncation of the liability distribution, biasing results through the systematic removal of the most
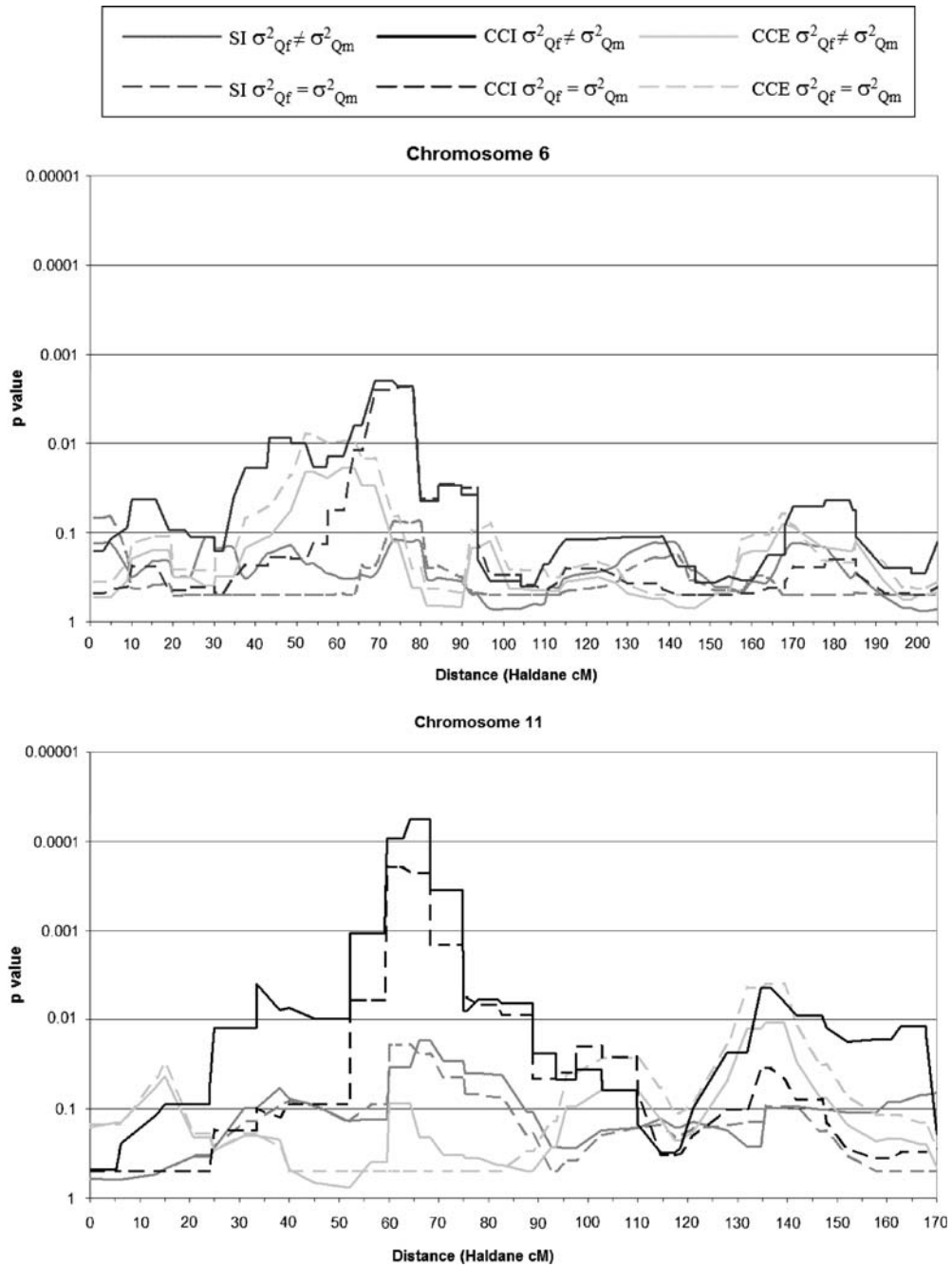
Fig. 2 Linkage results for chromosomes 6 and 11. Results from the model in which $\sigma^2_{Qf} \neq \sigma^2_{Qm}$ and in which $\sigma^2_{Qf} = \sigma^2_{Qm}$ are shown by the dashed and solid lines respectively. Legend as shown.

discordant pairs (Heath and Martin, 1993; Heath et al., 1998; Koopmans et al., 1999). Previous studies have found that a combined liability model, in which individuals may become a non-smoker either because of genetic and environmental factors that influence

initiation, or because their liability to cigarette consumption is so low that they are effectively non-smokers, provides the best fit to data on smoking initiation and cigarette consumption (Koopmans et al., 1999; Vink et al., 2004). Thus the most

Table IV. Most Significant Linkage Regions for Each of the Three Phenotypes

Phenotype	Chromosome	Markers	$\sigma^2_{Qf} = \sigma^2_{Qm}$		$\sigma^2_{Qf} \neq \sigma^2_{Qm}$		
			p-value	σ^2_Q	p-value	σ^2_{Qf}	σ^2_{Qm}
Smoking initiation	16	D16S423-D16S2616	0.00423	0.37	0.00960	0.29	0.60
	20	D20S103-GATA51D03	0.00359	0.43	0.00541	0.34	0.68
Cigarette consumption excluding non-smokers	4	GATA129-D4S2930	0.00982	0.41	0.0226	0.36	0.48
	6	ATA12005-D6S271	0.00774	0.40	0.0208	0.42	0.34
	11	D11S1998-D11S1345	0.00399	0.45	0.0107	0.50	0.38
Cigarette consumption including non-smokers	6	D6S1017-GATA29C09	0.00253	0.30	0.00197	0.46	0.12
	7	D7S684-D7S794	0.103	0.14	0.00270	0.41	0.001
	11	GATA73B08-D11S2371	0.00023	0.37	0.0000568	0.22	0.68
		AAT228-TAA003	0.0350	0.21	0.00440	0.42	0.013
	16	D16S423-D16S616	0.00556	0.32	0.00610	0.45	0.17

Estimates for the QTL parameter under the saturated and constrained models are shown for each region. Female and male parameter estimates are denoted σ^2_{Qf} and σ^2_{Qm} respectively, and the constrained estimate as σ^2_Q.

appropriate model for linkage analyses may actually be to fit a bivariate model including initiation as the first dimension and cigarette consumption as the second.

To further explore this, it would also be beneficial to differentiate those individuals who had tried cigarettes but not progressed to regular smoking from substance naïve individuals, as this may give a better approximation of the underlying liability distribution. Unfortunately this was not possible in this sample as participants were asked to classify themselves as current, ex-, or never smokers, which may have led individuals who had experimented with cigarettes to report the same classification as individuals who had never smoked a cigarette, thus confounding cigarette experimentation and non-smoking (Madden et al., 1999).

Our linkage findings replicate results from a number of other linkage analyses of smoking behavior. Our results for CCI on chromosome 6 ($p = 1.97 \times 10^{-3}$) replicate those of two previous studies. Vink et al. (2004) identified a region on chromosome 6p12 (at approximately 75 cM on our map) for smoking initiation in Dutch sibling pairs with a LOD score of 3.05. This region encompasses a locus previously identified by Duggirala et al. (1999) with LOD score 1.10 for cigarette consumption in sample of families from the Collaborative Study on the Genetics of Alcoholism (COGA). There are no known candidate genes for smoking located in this area, but these results suggest that it may merit further investigation.

Our most significant result, on chromosome 11p12 ($p = 5.68 \times 10^{-5}$), also overlaps with regions identified in previous studies. Many linkage analyses

of smoking behavior have found linkage to chromosome 11. The most significant linkage result for smoking behavior reported to date has been a LOD score of 3.95 on chromosome 11q12 (at approximately 70 cM on our map), identified by Li (2003b) for cigarette consumption in a Framingham Heart Study (FHS) sample. Goode et al. (2003), also analyzing a FHS sample, found a peak of LOD score 1.10 at 11p11 (at approximately 75 cM) for cigarette consumption. Bierut et al. (2004) also identified linkage to chromosome 11, at 11q14 (at approximately 88 cM), for habitual smoking, and habitual smoking with co-morbid alcohol dependence, with LOD scores of 1.64 and 1.05 respectively.

A recent review of candidate genes for smoking behavior identified a number of genes on chromosome 11 (Munafo et al., 2004), but none of these are located under either of the peaks we identified on this chromosome. One potential candidate gene located in this region is a glutamate transporter gene, SLC1A2. The glutamatergic system appears to mediate drug dependence, possibly through its moderation of reward processes (Cryan et al., 2003). Consequently, genes involved in this system have been suggested as targets for smoking cessation pharmacotherapies (Cryan et al., 2003; George and O'Malley, 2004). Thus, given previous studies produced LOD scores of 3.95 and 1.10, and our results were strongly suggestive, this region still appears to warrant further investigation. These results were obtained from an unselected sample, and consequently we plan to investigate the regions identified here in a sample selected for smoking behavior.

ACKNOWLEDGMENTS

REFERENCES

Abecasis, G. R., Cherny, S. S., Cookson, W. O., and Cardon, L. R. (2001). GRR: graphical representation of relationship errors. *Bioinformatics* **17**(8):742–743.

Abecasis, G. R., Cherny, S. S., Cookson, W. O., and Cardon, L. R. (2002). Merlin – rapid analysis of dense genetic maps using sparse gene flow trees. *Nat. Genet.* **30**:97–101.

Bierut, L. J., Rice, J. P., Goate, A., Hinrichs, A. L., Saccone, N. L., Foroud, T., Edenberg, H. J., Cloninger, C. R., Begleiter, H., Conneally, P. M., Crowe, R. R., Hesselbrock, V., Li, T.-K., Nurngerger, J. I., Porjesz, B., Schuckit, M. A., and Reich, T. (2004). A genomic scan for habitual smoking in families of alcoholics: common and specific genetic factors in substance dependence. *Am. J. Med. Genet.* **124**:19–27.

Cornes, B. K., Medland, S. E., Ferreira, M. A. R., Morley, K. I., Duffy, D. L., Heijmans, B. T., Montgomery, G. W., and Martin, N. G. (2005). Sex-limited genome-wide linkage scan for body mass index in an unselected sample of 933 Australian twin families. *Twin Res. Hum. Genet.* **8**(6):616–632.

Corrao, M. A., Guindon, G. E., Cokkinides, V., and Sharma, N. (2000). Building the evidence base for global tobacco control. *Bull. World Health Organ.* **78**(7):884–890.

Cryan, J. F., Gasparini, F., van Heeke, G., and Markou, A. (2003). Non-nicotinic neuropharmacological strategies for nicotine dependence: beyond bupropion. *Drug Discov. Today* **8**(22):1025–1034.

Duffy, D. L. (2002). *Sib-pair version 0.99.9 [Computer program].* Brisbane, Australia: Queensland Institute of Medical Research.

Duggirala, R., Almasy, L., and Blangero, J. (1999). Smoking behavior is under the influence of a major quantitative trait locus on human chromosome 5q. *Genet. Epidemiol.* **17**(Suppl 1): S139–S144.

Duren, W. L., Epstein, M. P., Li, M., and Boehnke, M. (2003). RELPAIR: A Program that Infers the Relationships of Pairs of Individuals Based on Marker Data. Version 2.0

Eaves, L. J., Last, K. A., Young, P. A., and Martin, N. G. (1978). Model-fitting approaches to the analysis of human behaviour. *Heredity* **41**(3):249–320.

Epstein, M. P., Duren, W. L., and Boehnke, M. (2000). Improved inference of relationship for pairs of individuals. *Am. J. Hum. Genet.* **67**(5):1219–1213.

Ezzati, M., and Lopez, A. D. (2003). Estimates of global mortality attributable to smoking in 2000. *Lancet* **362**:847–852.

Ezzati, M., and Lopez, A. D. (2004). Regional, disease specific patterns of smoking-attributable mortality in 2000. *Tob. Control* **13**:388–395.

Ezzati, M., Lopez, A. D., Rodgers, A., Van der Hoorn, S., and Murray, C. J. L.Comparative Risk Assessment Collaborating Group (2002). Selected major risk factors and global and regional burden of disease. *Lancet* **360**:1347–1360.

Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to Quantitative Genetics.* Essex: Longman.

Gelernter, J., Liu, X., Hesselbrock, V., Page, G. P., Goddard, A., and Zhang, H. (2004). Results of a genomewide linkage scan: support for chromosomes 9 and 11 loci increasing risk for cigarette smoking. *Am. J. Med. Genet.* **128**:94–101.

George, T. P., and O'Malley, S. S. (2004). Current pharmacological treatments for nicotine dependence. *Trends Pharmacol. Sci.* **25**(1):42–48.

Goode, E. L., Badzioch, M. D., Kim, H., Gagnon, F., Rozek, L. S., Edwards, K. L., and Jarvik, G. P. (2003). Multiple genome-wide analyses of smoking behavior in the Framingham Heart Study. *BMC Genet.* **4**(Suppl 1):S102 .

Grayson, D.A. (1989). Twins reares together: minimizing shared environmental effects. *Behav. Genet.* **19**(4):593–604.

Heath, A. C., Madden, P. A., and Martin, N. G. (1998). Statistical methods in genetic research on smoking. *Stat. Methods. Med. Res.* **7**(2):165–186.

Heath, A. C., and Martin, N. G. (1993). Genetic models for the natural history of smoking: evidence for a genetic influence on smoking persistence. *Addict. Behav.* **18**(1):19–34.

Heath, A. C., and Martin, N. G. (1994). Genetic influences on alcohol consumption patterns and problem drinking: results from the Australian NH & MRC twin panel follow-up survey. *Ann. NY Acad. Sci.* **780**:72–85.

Hettema, J. M., Corey, L. A., and Kendler, K. S. (1999). A multivariate genetic analysis of the use of tobacco, alcohol, and caffeine in a population based sample of male and female twins. *Drug Alcohol Depend.* **57**(1):69–78.

Hewitt, J. K. (1989). Of biases and more in the study of twins reared together: a reply to Grayson. *Behav. Genet.* **19**(4):593–604.

Kendler, K. S., and Gardner, C. O. Jr. (1998). Twin studies of adult psychiatric and substance dependence disorders: are they biased by differences in the environmental experiences of monozygotic and dizygotic twins in childhood and adolescence? *Psychol. Med.* **28**(3):625–633.

Kendler, K. S., Neale, M. C., Sullivan, P., Corey, L. A., Gardner, C. O., and Prescott, C. A. (1999). A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol. Med.* **29**(2):299–308.

Kendler, K. S., Thornton, L. M., and Pedersen, N. L. (2000). Tobacco consumption in Swedish twins reared apart and reared together. *Arch. Gen. Psychiatry.* **57**(9):886–892.

Kirk, K. M., Birley, A. J., Statham, D. J., Haddon, B., Lake, R. I. E., Andrews, J. G., and Martin, N. G. (2000). Anxiety and depression in twin and sib pairs extremely discordant and concordant for neuroticism: prodromus to a linkage study. *Twin Res.* **3**:299–309.

Koeppen-Schomerus, G., Spinath, F. M., and Plomin, R. (2003). Twins and non-twin siblings: different estimates of shared

environmental influence in early childhood. *Twin Res.* **6**(2):97–105.

Kong, X., Murphy, K., Raj, T., He, C., White, P. S., and Matise, T. C. (2004). A combined linkage-physical map of the human genome. *Am. J. Hum. Genet.* **75**(6):1143–1148.

Koopmans, J. R., Slutske, W. S., Heath, A. C., Neale, M. C., and Boomsma, D. I. (1999). The genetics of smoking initiation and quantity smoked in Dutch adolescent and young adult twins. *Behav. Genet.* **29**(6):383–393.

Kruglyak, L., Daly, M. J., Reeve-Daly, M. P., and Lander, E. S. (1996). Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am. J. Hum. Genet* **58**(6):1347–1363.

Lake, R. I., Eaves, L. J., Maes, H. H., Heath, A. C., and Martin, N. G. (2000). Further evidence against the environmental transmission of individual differences in neuroticism from a collaborative study of 45,850 twins and relatives on two continents. *Behav. Genet.* **30**(3):223–233.

Lander, E., and Kruglyak, L. (1995). Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**:241–247.

Lange, K., Weeks, D., and Boehnke, M. (1988). Programs for pedigree analysis: MENDEL, FISHER, and dGENE. *Genet. Epidemiol.* **5**(6):471–472.

Li, M. D., Cheng, R., Ma, J. Z., and Swan, G. E. (2003a). A meta-analysis of estimated genetic and environmental effects on smoking behavior in male and female adult twins. *Addiction* **98**(1):23–31.

Li, M. D., Ma, J. Z., Cheng, R., Dupont, R. T., Williams, N. J., Crews, K. M., Payne, T. J., and Elston, R. C. (2003b). A genome-wide scan to identify loci for smoking rate in the Framingham Heart Study population. *BMC Genet.* **4**(Suppl 1):S103 .

Madden, P. A., Heath, A. C., Pedersen, N. L., Kaprio, J., Koskenvuo, M. J., and Martin, N. G. (1999). The genetics of smoking persistence in men and women: a multicultural study. *Behav. Genet.* **29**(6):423–431.

Madden, P. A., Pedersen, N. L., Kaprio, J., Koskenvuo, M. J., and Martin, N. G. (2004). The epidemiology and genetics of smoking initiation and persistence: crosscultural comparisons of twin study results. *Twin Res.* **7**(1):82–97.

Maes, H. H., Sullivan, P. F., Bulik, C. M., Neale, M. C., Prescott, C. A., Eaves, L. J., and Kendler, K. S. (2004). A twin study of genetic and environmental influences on tobacco initiation, regular tobacco use and nicotine dependence. *Psychol. Med.* **34**:1–11.

Medland, S. E. (2004). Alternate parameterization for scalar and non-scalar sex-limitation models in Mx. *Twin Res.* **7**(3):299–305.

Medland, S. E. (2005). Parameterization of sex-limited autosomal linkage analysis for Mx. *Twin Res. Hum. Genet.* **8**(6):569–573.

Munafo, M. R., Clark, T. G., Johnstone, E. C., Murphy, M. F. G., and Walton, R. T. (2004). The genetic basis for smoking behavior: a systematic review and meta-analysis. *Nicotine Tob. Res.* **6**(4):583–597.

Neale, M. C. (1998). Twin analysis. In P. Armitage and T. Colton (eds.*Encyclopedia of Biostatistics.* Chinchester, New York: Wiley, pp. 4613–4626.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (2003). Mx: Statistical Modeling. VCU Box 900126, (6th Ed.), Richmond, VA 23298: Department of Psychiatry

Neale, M. C., and Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families.* Dordrecht, Boston, London: Kluwer Academic Publishers.

Neale, M. C., Walters, E. E., Eaves, L. J., Maes, H. H., and Kendler, K. S. (1994). Multivariate genetic analysis of twin-family data on fears: Mx models. *Behav. Genet.* **24**(2):119–139.

Posthuma, D., Beem, A. L., de Geus, E. J. C., van Baal, G. C. M., von Hjelmborg, J. B., Iachine, I., and Boomsma, D. I. (2003). Theory and practice in quantitative genetics. *Twin Res.* **6**(5):361–376.

Posthuma, D., and Boomsma, D. I. (2000). A note on the statistical power in extended twin designs. *Behav. Genet.* **30**:147–158.

Saccone, N. L., Neuman, R. J., Saccone, S. F., and Rice, J. P. (2003). Genetic analysis of maximum cigarette-use phenotypes. *BMC Genet.* **4**(Suppl 1):S105 .

Sham, P. (1998). *Statistics in Human Genetics.* London: Arnold.

Straub, R. E., Sullivan, P. F., Ma, Y., Myakishev, M. V., Harris-Kerr, C., Wormley, B., Kadambi, B., Sadek, H., Silverman, M. A., Webb, B. T., Neale, M. C., Bulik, C. M., Joyce, P. R., and Kendler, K. S. (1999). Susceptibility genes for nicotine dependence: a genome scan and followup in an independent sample suggest that regions on chromosomes 2, 4, 10, 16, 17 and 18 merit further study. *Mol. Psychiatry* **4**(2):129–144.

Sullivan, P. F., and Kendler, K. S. (1999). The genetic epidemiology of smoking. Nicotine Tob. Res. 1 Suppl 2: S51–S57; discussion S69–70

Sullivan, P. F., Neale, B. M., van den Oord, E., Miles, M. F., Neale, M. C., Bulik, C. M., Joyce, P. R., Straub, R. E., and Kendler, K. S. (2004). Candidate genes for nicotine dependence via linkage, epistasis, and bioinformatics. *Am. J. Med. Genet.* **126**(1):23–36.

Vink, J. M., Beem, A. L., Posthuma, D., Neale, M. C., Willemsen, G., Kendler, K. S., Slagboom, P. E., and Boomsma, D. I. (2004). Linkage analysis of smoking initiation and quantity in Dutch sibling pairs. *Pharmacogenomics J.* **4**(4):274–282.

Edited by Stacey Cherny