Molecular Psychiatry (1999) 4, 129–144

IMMEDIATE COMMUNICATION

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

RE Straub[1,3], PF Sullivan[1,3], Y Ma[1,3], MV Myakishev[1,3], C Harris-Kerr[1,3], B Wormley[1,3], B Kadambi[1,3], H Sadek[1,3], MA Silverman[1,3], BT Webb[1,3], MC Neale[1,3], CM Bulik[1,3], PR Joyce[4] and KS Kendler[1,2,3]

Departments of [1]Psychiatry; [2]Human Genetics; [3]Virginia Institute for Psychiatric and Behavioral Genetics, Medical College of Virginia, Virginia Commonwealth University, Richmond, VA 23298, USA; [4]Department of Psychological Medicine, Christchurch School of Medicine, Christchurch, New Zealand

Cigarette smoking is associated with considerable morbidity, mortality, and public health costs. Genetic factors influence both smoking initiation and nicotine dependence, but none of the genes involved have been identified. A genome scan using 451 markers was conducted to identify chromosomal regions linked to nicotine dependence in a collection of 130 families containing 343 genotyped individuals (308 nicotine-dependent) from Christchurch, New Zealand. By pairwise analysis, the best result was with marker D2S1326 which gave a lod score under heterogeneity (H-LOD) of 2.63 ($P = 0.0012$) and a nonparametric linkage (NPL, Z_{all}) score of 2.65 ($P = 0.0011$). To identify regions that warranted further study, rather than comparing the pairwise scores from the scan to theoretical thresholds, we compared them to an empirical baseline, found here to be H-LOD scores of 0.5 and Z_{all} scores of 1.0. We also found a number of large (31–88 cM) regions where many (8–16) consecutive markers yielded small but positive Z_{all} scores. Selected regions of chromosomes 2, 4, 10, 16, 17 and 18 were investigated further by additional genotyping of the Christchurch sample and an independent sample from Richmond, Virginia (91 families with 264 genotyped individuals, 211 nicotine-dependent). Multipoint nonparametric analysis showed the following maximums for the Christchurch sample: Chr. 2 ($Z_{lr} = 2.61$, $P = 0.005$), Chr. 4 ($Z_{lr} = 1.36$, $P = 0.09$), Chr. 10 ($Z_{lr} = 2.43$, $P = 0.008$), Chr. 16 ($Z_{lr} = 0.85$, $P = 0.19$), Chr. 17 ($Z_{lr} = 1.64$, $P = 0.05$), Chr. 18 ($Z_{lr} = 1.54$, $P = 0.06$). Analysis of the Richmond sample showed the following maximums: Chr. 2 ($Z_{lr} = 1.00$, $P = 0.15$), Chr. 4 ($Z_{lr} = 0.39$, $P = 0.34$), Chr. 10 ($Z_{lr} = 1.21$, $P = 0.11$), Chr. 16 ($Z_{lr} = 1.11$, $P = 0.13$), Chr. 17 ($Z_{lr} = 1.60$, $P = 0.05$), Chr. 18 ($Z_{lr} = 1.33$, $P = 0.09$). It is probable that the small samples used here provided only limited power to detect linkage. It may have been difficult therefore to detect genes of small effect, or those that are influencing risk in only a small proportion of the families. When simply judged against the usual standards of linkage significance, none of the individual regions yielded strong evidence in either sample. Some or all of the most positive results in the genome scan of the Christchurch sample, therefore, could be due to chance. However, the presence in the Christchurch scan of multiple large regions containing many consecutive positive markers, coupled with the relatively positive results in these same regions in the Richmond sample, suggests that some of these regions may contain genes influencing nicotine dependence and therefore deserve further study.

Keywords: nicotine dependence; smoking; genetics; linkage study

Introduction

Despite substantial reductions in the prevalence of current smoking in the last 25 years, approximately one third of adults in the United States continue to smoke.[1]

Correspondence: RE Straub, Director, Molecular Genetics Laboratory, Department of Psychiatry, Virginia Commonwealth University, Virginia Institute for Psychiatric and Behavioral Genetics, 800 East Leigh St, Suite 110, Richmond, VA 23219-1534, USA. E-mail: rstraub@hsc.vcu.edu
Received 16 December 1998; revised and accepted 30 December 1998

The tremendous impact of cigarette smoking on mortality and morbidity has been well chronicled.[2–6] In 1990, smoking was responsible for 418 690 deaths in the US compared to 125 000 for alcohol and 8100 for cocaine.[7,8] It has now been conclusively demonstrated that: (i) a large proportion of cigarette smokers are nicotine dependent; and (ii) nicotine is more addictive than most illicit drugs.[9–11] The 1988 Surgeon General's report on cigarette smoking[10] concluded that: '(1) cigarettes and other forms of tobacco are addicting; (2) nicotine is the drug in tobacco that causes addiction; and (3) the pharmacologic and behavioral processes that

determine tobacco addiction are similar to those that determine addiction to drugs such as heroin and cocaine'. The majority of smokers want to quit, but few are successful — each year 34% try, but only 3% succeed.[12] The sheer magnitude of the health problems and costs created by smoking mandates that tremendous effort should be devoted to basic research into the biological and environmental factors involved. The insight gained would then hopefully be translated quickly into practical benefits such as earlier identification of individuals at risk, and development of more effective methods for helping people to quit smoking.

Data from family, adoption, and twin studies have consistently supported a substantial genetic component to many aspects of smoking behavior. Most convincingly, a number of large-scale population-based twin studies have shown that genetic factors contribute strongly to individual differences in smoking behavior.[13–18] Work using inbred strains of mice, notably the elegant set of experiments by Marks, Collins, and coworkers,[19–23] has clearly demonstrated genetic influence on inter-individual differences of a variety of related phenotypes. These include sensitivity to nicotine, nicotinic receptor density, the ability to develop tolerance, the extent of nicotine self-administration, and cross-tolerance with other substances such as alcohol.

Liability to nicotine dependence is likely to be primarily a function of specific anatomic and functional characteristics of the brain. A recent real-time functional MRI study[24] of smokers showed that intravenous injection of nicotine induced a dose-dependent increase in neuronal activity that was most pronounced in regions that comprise specific functional systems. These were primarily the cingulate cortex and frontal regions (involved in working memory, attention, motivation, mood and emotion), and the locus ceruleus (modulation of behavioral arousal) — activation of these anatomical regions is functionally consistent with the mood-elevating and anxiolytic effects often reported by cigarette smokers. In addition, limbic subcortical regions, including the nucleus accumbens, amygdala, and hypothalamus were also strongly activated. Moreover, several independent lines of evidence have demonstrated that nicotine, like other drugs of abuse such as cocaine, amphetamines and opiates, stimulates this so-called 'central dopamine reward pathway', ie activation of mesolimbic dopaminergic neurons resulting in increased energy metabolism and dopamine release in the shell of the nucleus accumbens.[25–27]

It is presumed that most of the behavioral effects of nicotine are due to activation of central nicotinic acetylcholine receptors.[28] Different pentameric combinations of receptor subunits (α_2 through α_7, α_9, and β_2 through β_4 are found in brain) generate numerous receptor subtypes, each with different regional distributions and functional characteristics that remain to be well characterized.[28] Few polymorphisms in these families of genes have been described, and the roles of the receptor subtypes in mediating the effects of nicotine on neurotransmitter release are virtually unknown.[29,30] It does appear though, that the CHRNB2 gene (located near marker D1S305 in region 1q21.3) is essential, since knockout mice[31] without the β_2 subunit lack high affinity nicotine binding in brain (levels of α_2 through α_7, β_3, and β_4 subunits are normal). They also lack sensitivity[31] to nicotine-elicited increases in electrophysiological discharge frequency and dopamine release in striatal dopaminergic neurons, and exhibit altered responses in a behavioral paradigm thought to reflect self-administration of reinforcing substances.

Inactivation of nicotine in the liver and lung, predominantly to (S)-cotinine, is catalyzed principally by microsomal cytochrome P450 2A6 (CYP2A6).[32] There is large inter-individual variation in cotinine formation in human microsomal preparations, but it is not clear how much is due *in vivo* to variations in CYP2A6 levels and activity compared to other factors.[33] The gene is located in region 19q13.2, and three mutant alleles, each of which produce an inactive enzyme, have been described:[34,35] a substitution (CYP2A6*2), a gene conversion (CYP2A6*3), and a deletion. The *2 and *3 alleles have been suggested to protect against nicotine dependence,[36] but the accuracy of the genotyping methods used has been questioned.[37] The frequency of the *2 allele was found to be 1–3% in Swedish, Spanish, and Finnish populations,[37] and therefore is probably too rare in most Caucasian populations to be a major factor in a trait as common as nicotine dependence. No other polymorphisms known to affect CYP2A6 function have been reported.

It is reasonable, therefore, to investigate whether polymorphisms in the genes involved in the regulation of neurotransmission by nicotinic acetylcholine receptors, dopaminergic transmission (eg the D4 receptor[38]), or metabolic pathways, may contribute to individual differences in the susceptibility to smoking initiation, nicotine dependence and withdrawal symptoms. In addition, the genes that regulate the expression of such genes, or that serve to modulate the function of the protein products, are also functional candidates. A number of other candidate genes besides these also deserve consideration.[39,40] The strategy of our research program on smoking and nicotine is to use the positional and candidate gene approaches synergistically, and to collaborate with other groups that have complementary resources and methods. The goal is to identify those genes that contribute most to susceptibility, to use this knowledge as a tool to identify and characterize in greater detail the important environmental variables, and then to explore the interactions between genes and environment. The long-term goal is to contribute to a better understanding of the variety of mechanisms, from molecular to interpersonal, that mediate smoking behavior and nicotine dependence. We report here preliminary results from one approach in the first phase — a complete genome scan of an affected sib pair sample collected in Christchurch, New Zealand, and followup of the six most positive regions in a second, independent sample from Richmond, Virginia, USA.

Methods

Subjects

The subjects for this investigation were recruited and studied at two sites using nearly identical protocols. Table 1 shows some of the characteristics of the samples. The protocols were reviewed and approved by the appropriate ethical committees at each site. The first sample was recruited in Christchurch, New Zealand and was used for the complete genome scan. The second sample was recruited in Richmond, Virginia (USA) and was used for followup of regions selected because they were positive in the scan. Ascertainment was via convenience sampling and included advertisements in a wide variety of media, smoking cessation clinics, and via word-of-mouth. Inclusion criteria for a sibling pair included: the presence of lifetime nicotine dependence, current age 18 years or older, and the provision of informed consent.

Affection status

The symptoms of nicotine dependence during the period of lifetime heaviest tobacco use were assessed with an adapted version of the Fagerstrom Tolerance Questionnaire (FTQ)[41,42] that had previously been administered to over 7000 individuals in other studies conducted by our group.[17] The FTQ is widely used in smoking research to assess nicotine dependence and consists of eight brief questions (eg how soon after you wake up do you smoke your first cigarette? and do you smoke if you are so ill that you are in bed most of the day?). Total FTQ scores range from 0–11. A substantial body of research suggests that FTQ scores are positively correlated with plasma nicotine and cotinine levels and predict unsuccessful attempts at smoking cessation.[42] Following the cutoff used by others[43–45] we defined nicotine dependence as an FTQ score greater than or equal to 7. All subjects must have reported FTQ scores of 7 or more on two separate occasions, during an initial telephone screening and then again during the study interview.

Table 1 Nicotine dependence was assessed using an adapted version of the Fagerstrom Tolerance Questionnaire (FTQ),[41,42] and individuals with FTQ scores of 7 or greater were classified as affected

	Selected characteristics of the samples		
	Christchurch	*Richmond*	*Total*
No. families	130	91	221
No. genotyped — total	343	264	607
No. genotyped — affected	308	211	519
% Caucasian	90%	85%	
% Female	63%	61%	
Age, mean (sd)	40.7 (11.5)	2.7 (11.5)	
FTQ score, mean (sd)	8.9 (1.2)	8.5 (1.2)	
Affected sibships — size	No.	No.	Total
2	112	72	184
3	13	10	23
4	5	3	8
5	2	4	6
6	0	2	2
Affected sib pairs — all	201	190	391

Isolation of DNA from cytology brushes

DNA was obtained during direct interview by having the subjects collect buccal epithelial cells (from inside of the cheek) using standard cytology brushes[46] (Fisher Scientific, Pittsburgh, PA, USA). DNA preparation was done from four brushes per subject using the 'Instagene Purification Matrix' from Biorad Laboratories (Hercules, CA, USA) as per the protocol provided. The procedure simultaneously lyses the buccal cells, lyses the nuclei, and the majority of the non-DNA contaminants are adsorbed to the matrix, which is pelleted, leaving the chromosomal DNA free in solution. We prepare two tubes of DNA per subject, 600 μl each, and made working stocks at a dilution of 1:50. We have done extensive work with DNA from brushes, and have not found the banding pattern, error rates, or 'readability' of microsatellite markers to be significantly different between DNA from brushes and DNA isolated from blood or cell lines. The percentage of samples for which most markers amplify well is about 95%. From our twin zygosity studies, MZ twins tested with 15 fluorescent microsatellite markers provided an estimate of 0.7% for the genotyping error rate.

Markers, maps and genotyping

The markers were predominantly tri- and tetra-nucleotide repeat microsatellites generated by the Cooperative Human Linkage Center (CHLC; http://www.chlc.org/), and many are from the Weber screening set, version 8.0 (http://www.marshmed.org/genetics/). Allele information, marker order, sex-averaged distances[47] were copied from the Marshfield Medical Research Center web site (http://www.marshmed.org/genetics/). Before production work, all markers were tested to confirm their amplification in multiplex PCR, lack of spurious alleles, and ease of scoring. Genomic DNA (3.5 μl of a 1:50 dilution of the stock sample) per well was loaded on the bottom of a 96-well Greiner (Lake Mary, FL, USA) Thermoquick PCR plate with use of a Tecan (Research Triangle Park, NC, USA) Genesis 200 Robotic Sample Processor (RSP) and dried. The plates were stored at room temperature for up to 2 months — dried DNA produces essentially the same results as a newly prepared sample. PCR mix (3.5 μl) was distributed on the side of the wells of the PCR plate by the Tecan RSP. Plates were manually tapped, put on an oscillating shaker for 10 min (380 rpm) and the wells covered with mineral oil. The final concentrations of the components of PCR reaction were: 1–5 ng μl^{-1} human genomic DNA; 0.24 μM primers; 10 mM Tris-HCl (pH 8.3); 50 mM KCl; 2.15 mM $MgCl_2$; 200 μM (each) dNTPs; and 0.036 Units μl^{-1} (0.072 Units μl^{-1} was used for some markers) ABI AmpliTaq Gold. From one to four

primer pairs were combined in a single reaction — most primer multiplexes were tested on a limited number of samples prior to full-scale PCR. After storage in the dark for up to 12 h at 4°C, premixed PCR plates were placed on Hybaid (Holbrook, NY, USA) Omnigene Thermocyclers. For most markers, a 'touchdown' PCR program[48] was used: 10 min at 95°C, then (1 min–2 min–1 min) cycles at temperatures (°C) (94–60–72), (94–59–72), (94–58–72), (94–57–72), (94–56–72), (94–55–72), (94–54–72), (94–53–72), (94–52–72), (94–51–72), then 30 cycles (°C): (94–50–72), then 6 min at 72°C. Samples from up to 10 PCR plates were pooled together and mixed with formamide and dye using the Tecan RSP in the following proportions: 33% pooled PCR products, 13% ABI Tamra GS500, 53.7% formamide, 0.3% Dextran Blue, 0.025% Bromphenol Blue. The pooling and multiplexing schemes were designed in accordance with the size and color of the PCR product, and we generally loaded 10 markers per lane. The mix was denatured for 10 min at 95°C and approximately 0.4 μl was stagger loaded with a Kloehn 8-syringe loader onto a 30-cm gel (6% Long Ranger acrylamide) with 96-well sharktooth comb. Gels were run on PE-ABI (PE Biosystems, Foster City, CA, USA) Model 377 DNA sequencers at 100 W (3000 V) for 2.5 h. DNA from CEPH individuals 1331-01 and 1331-02 was amplified and loaded in parallel to confirm allele sizes. Gel data files were automatically tracked and then sometimes manually adjusted using ABI GENESCAN™ software (v3.0) and scored blind to phenotype using the ABI GENOTYPER™ software (v2.0). Mendelian inheritance was checked using the program Geno (v1.1 from Marshfield Center for Medical Genetics), and sex was determined. Incompatibilities were either resolved unambiguously or the data for the entire family were discarded. We used the program RELPAIR[49] to estimate empirically the degree of genetic relationship based on multipoint analysis of 30 and then again of 95 markers in the Christchurch sample and of 45 markers in the Richmond sample.

Linkage analysis

We are unaware of any prior estimates of single-major locus models for nicotine dependence. We therefore began by assuming a penetrance in males of 60% for the highest risk and 5% for the lowest risk genotype. From a large epidemiologic sample of twins from the Virginia Twin Registry that included the FTQ, the lifetime prevalence for nicotine dependence, as defined in this report, was 24% in males and 11% in females. However, among regular smokers, the prevalence of nicotine dependence was 36% in males and 26% in females. We decided to utilize population prevalence rates, but set the penetrance in females to be 70% of that in males.

	Penetrance			Frequency of
	AA	Aa	aa	risk allele
MALES:				
Dominant	0.60	0.60	0.05	0.19
Recessive	0.60	0.05	0.05	0.59

Prevalence = 0.24; Sporadic risk = 0.050; Phenocopy proportion = 0.137)

FEMALES:				
Dominant	0.42	0.042	0.035	0.19
Recessive	0.42	0.035	0.035	0.59

Prevalence = 0.11; Sporadic risk = 0.035; Phenocopy proportion = 0.209)

For analysis of X chromosome markers, the parameters for females were the same: for males, the dominant model was: A = 0.60, a = 0.05, risk allele frequency = 0.19 and the recessive model was A = 0.60, a = 0.05, risk allele frequency = 0.59.

Using the GENEHUNTER program[50] (v1.0) we generated pairwise lod scores under the assumption of heterogeneity (H-LODS) and nonparametric linkage (NPL Z_{all}) scores and their associated *P* values. We report only the maximum H-LOD score for each marker — estimates of the recombination fractions (θ) and proportions of families segregating the disease locus (α) at which that maximum occurred are available on request. We report the NPL Z_{all} statistic, which is calculated based on group-wise S_{all} allele sharing among affecteds in sibships — for many markers in this study, the values of Z_{all} and $Z_{pairwise}$ were very similar (data not shown). GENEHUNTER (v 2.0) was also used to generate multipoint H-LODs and NPL (Z_{all}) scores over entire chromosomes for the Christchurch sample and over the regions tested in the Richmond sample. We also applied the extensions to GENEHUNTER proposed by Kong and Cox,[51] which test the significance of observed multipoint allele sharing using the efficient score statistic distribution, as implemented in the GENEHUNTER-PLUS program. This program generates a different statistic termed Z_{lr}, which is probably more accurate than Z_{all}, in which the perfect data approximation can lead to overly conservative *P* values when there is missing information.[51,52] However, Z_{lr} has the property that its minimum is zero, which limits its utility in evaluating the evidence against linkage provided by the negative Z_{all} values given by GENEHUNTER. We therefore used both programs for multipoint analysis and present both results. Allele frequencies were estimated from all independent (unrelated) individuals.

Results

Sample characteristics

A genome scan using 451 markers was conducted to identify chromosomal regions linked to nicotine dependence in a collection of 130 families containing 343 genotyped individuals (308 affected) from Christchurch, New Zealand. Six of the most positive regions were then further investigated by additional genotyping of the Christchurch sample and of an independent sample from Richmond, Virginia (91 families containing 264 genotyped individuals, 211 affected). Some descriptive statistics for the two samples are shown in Table 1. They are similar in general composition, phenotypic characteristics, and size distribution of affected sibships. As might be expected since both samples are predominantly Caucasian (mixed European) in origin, the allele frequencies for markers tested on the Richmond sample were extremely similar to those found in Christchurch.

Maps and final marker distribution
We used the sex-averaged maps[47] from the Marshfield Medical Research Center (3655 cM total). Including estimated distances to the telomeres, the final mean intermarker distance was 7.7 cM (SD = 3.8 cM; range 0–22 cM). Over 90% of the markers contained tetranucleotide repeats, and the mean marker heterozygosity was 0.74 (SD 0.08; range 0.45–0.95).

Distribution of scores
We used the program GENEHUNTER[50] to calculate heterogeneity lod scores (H-LOD) under dominant and recessive models, and to calculate the NPL statistic Z_{all}. The vast majority of markers must be unlinked, but each will still produce a maximum H-LOD near or at zero at large recombination fractions. The distribution of scores above 0.5, given as score (number observed) was as follows. Dominant: >2 (1), 1–2 (7), 0.5–1.0 (20); for recessive: >2 (1), 1–2 (7), 0.5–1.0 (19). Z_{all}: >2 (2), 1.5–2 (4), 1.0–1.5 (18), 0.5–1.0 (61). The percentage of markers with scores >0 was 38, 36, and 42 for dominant, recessive and Z_{all} respectively. Across the genome, the mean (SD) for all 451 markers was 0.096 (0.260) for dominant, 0.091 (0.251) for recessive, and –0.181 (0.778) for Z_{all}. The Pearson correlation coefficients between the results from the three tests, given as tests compared (data from all 451 markers; data from only the 85 markers with Z_{all} >0.5), was as follows. Dominant/recessive (0.83; 0.70), dominant/Z_{all} (0.66; 0.90), and recessive/Z_{all} (0.64; 0.79). As expected, about 6% of markers gave H-LODs greater than 0.5 and 5% of Z_{all} scores were greater than 1.0. We used these two values as the working baseline.

Most positive scores
Figure 1 shows the pairwise H-LODs, under dominant (top panel) and recessive (middle panel) models, for all 451 markers in the genome scan of the Christchurch sample (hereafter called Christchurch). Maximum H-LODs (ie the largest lod score obtained at any recombination fraction and proportion of families linked) are plotted at the marker locations. The bottom panel of Figure 1 shows the pairwise nonparametric score Z_{all}. The largest H-LOD was 2.63 ($P = 0.0012$) with marker D2S1326 under a dominant model, which also produced a Z_{all} score of 2.65 ($P = 0.0011$). In addition to three markers on chromosome 10 (D10S2469, D10S1239, D10S677), markers D16S422, D17S2059, and D18S869 also produced Z_{all} scores greater than 1.5.

Choice of followup regions
To identify chromosomal regions that warranted follow up, we ranked the results from each of the three tests and found that multiple markers on chromosomes 2, 10, 16, 17, and 18 gave scores that were above baseline (Figure 1). Flanking markers in all five regions provided additional support for linkage. Since at this stage false negatives are highly undesirable, and the highest scores in the genome scan will not necessarily identify all regions containing susceptibility genes, we examined the data more carefully.

Clustering of positive scores
Application of sib-pair methods to simulated data showed that the region around a true susceptibility locus contained a long string of consecutive positive ($P < 0.05$) markers.[53] In addition, by simulation it has been shown[54] that for two regions containing equal maximum peak heights, the distance over which the positive pairwise scores are observed in the presence of a gene will be longer than in the absence of one, and that longer positive regions are more likely to contain the gene of interest than shorter ones. This suggests that the longer positive regions should be given priority, but tests based on detecting this differential do not serve to *definitively* distinguish true signals from false.[53,55] Our scan sample consisted primarily of affected sib pairs (132 unique, 201 total), and we expect that for false positives, the correlation between adjacent markers will extend over only relatively 'short' distances. We have not developed a good method to accurately predict the characteristics of this effect in our samples, but the scan data suggest that IBD sharing due to random fluctuation extends over large (>20 cM) regions for only a percentage of the (presumably false) positives. We found that often Z_{all} scores > 1.0 are flanked by negative Z_{all} scores. Examples of this include Chr. 16: (–1.03/6.4 cM/+1.08/6.2 cM/–1.22); X Chr.: (–0.74/8.3 cM/+1.24), Chr. 7 (–0.75/7.4 cM/+1.44/13.0 cM/+0.49/5.2 cM/–0.69) and Chr. 11: (–0.44/13.1 cM/+1.10/7.9 cM/–0.842).

Overall in the scan, if we define a 'cluster' as three or more markers (which would cover on average only about 15 cM) each with Z_{all} scores >0, we found few. There were eight cluster of three markers (two were on Chrs. 2 and 4), four of 4 (two were on Chr. 17, 24 and 22 cM), and clusters of 6 (Chr. 2, 25 cM), 9 (Chr. 4, 51 cM), 12 (Chr. 10, 63 cM) and 14 (Chr. 18, 82 cM) markers each. Only 84 out of the 451 scan markers were contained in clusters. When compared to this baseline, chromosomes 2, 10, 18 were again noteworthy. In addition, chromosome 4 appeared to be of interest, with nine consecutive positive Z_{all} scores (including scores of 1.30, 1.16, and 1.23) over 51 cM. Based on the combined criteria of most positive Z_{all} scores and eight clusters of scores greater than zero across the genome, five regions, on chromosomes 2, 4, 10, 17 and 18, were followed up by further genotyping of the Christchurch and Richmond samples. Chromosome 16 was added to for comparison, to observe followup results from a region containing negative results: (over 57 cM: Z_{all} = 0.60, 0.95, 0.80, –0.39, –1.03, 1.08, –1.22, 1.50).

Followup results from the Christchurch and Richmond samples
The pairwise H-LOD and Z_{all} scores obtained for markers on these six chromosomes (which includes some followup markers to fill gaps), are shown for the Christchurch and Richmond samples in Table 2a–f. The corresponding nonparametric multipoint results, the Z_{all} score from GENEHUNTER and the test statistic Z_{lr} from GENEHUNTER-PLUS, are shown in Figure 2. Below

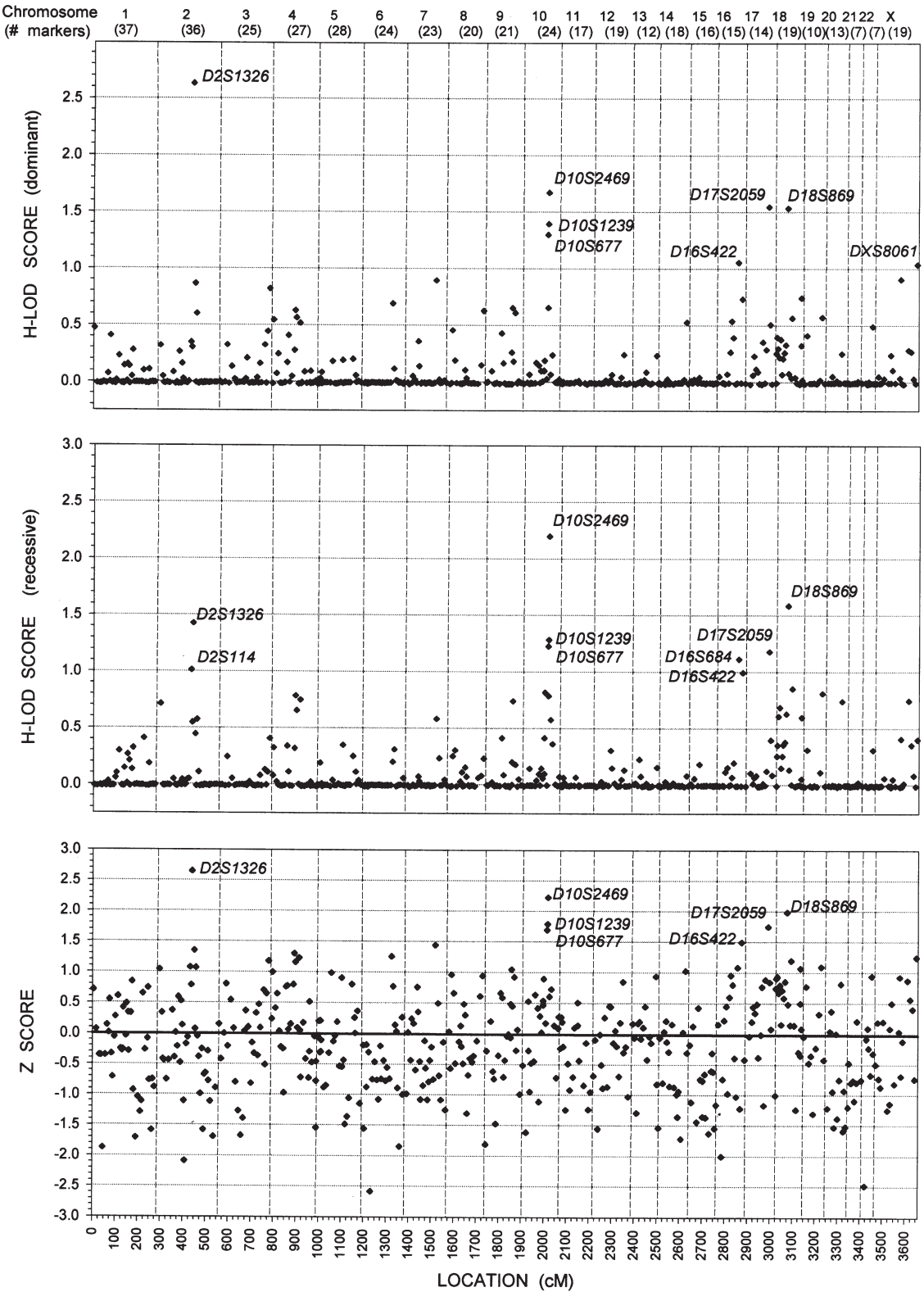

Figure 1 Pairwise heterogeneity lod (H-LOD) and nonparametric linkage (NPL, Z_{all}) scores for all 451 markers in the genome scan of the Christchurch sample. The program GENEHUNTER[50] (v 1.0) was used to calculate dominant model H-LOD (top panel), recessive model H-LOD (middle panel), and Z_{all} scores (bottom panel). The parameters used for the H-LOD calculations are given in the Methods section. The maximum value achieved for each marker is plotted at the marker location. Chromosome number and number of markers per chromosome are shown at the top. Marker locations and estimated distances to the telomeres were taken from the sex-averaged maps[47] available from the Marshfield Medical Research Center.

Table 2 Pairwise heterogeneity lod (H-LOD) scores and nonparametric (NPL, Z_{all}) scores for all markers tested on chromosomes 2, 4, 10, 16, 17 and 18 on the Christchurch and Richmond samples

MARKER	*LOCUS*	*Christchurch*					*Richmond*			
		H-LOD			*NPL*		*H-LOD*		*NPL*	
		cM	*Dom*	*Rec*	Z_{all}	*p val.*	*Dom*	*Rec*	Z_{all}	*p val.*
(a) Pairwise results with chromosome 2 markers										
p_tel		0.00								
GATA72G11	D2S1780	7.30	0.32	0.71	1.04	0.15				
GATA116B01	D2S2952	17.80	0.05	–0.01	0.34	0.37				
GAAT1A5	D2S423	22.10	–0.01	–0.01	–0.42	0.66				
GGAA20G10	D2S1400	27.60	–0.01	–0.01	–0.44	0.67				
GATA11H10	D2S1360	38.33	–0.01	–0.01	–0.75	0.77				
GATA8F07	D2S405	47.97	–0.01	–0.01	–0.44	0.67				
ATA4F03	D2S1356	64.29	0.09	0.05	0.37	0.35				
ATA27D04	D2S1352	73.61	–0.01	–0.01	–0.39	0.65				
GATA9A11	D2S1337	79.00	–0.01	–0.01	0.01	0.49				
GATA72A05	D2S441	86.82	–0.02	0.00	–0.21	0.58				
GATA69E12	D2S1394	90.82	0.26	0.00	0.59	0.28				
GATA71G04	D2S1777	99.41	–0.02	–0.01	–0.48	0.68				
GATA88G05	D2S1790	103.16	0.16	0.04	0.53	0.30				
GATA112E03		107.99	0.03	0.01	0.13	0.45				
ATA19E11	D2S1343	115.49	0.00	0.00	–1.11	0.87				
GATA5G02	D2S436	118.16	0.00	0.00	–2.09	0.98	–0.01	0.03	–0.02	0.51
GATA4E11	D2S410	125.18	–0.01	0.04	–0.17	0.57	0.00	–0.01	0.04	0.48
GATA27A12	D2S1328	132.58	0.00	0.05	–0.06	0.52	0.13	0.53	0.62	0.26
AFM052xf8	D2S114	142.83	0.35	1.01	1.08	0.14	0.06	0.64	0.42	0.33
GATA8H05	D2S442	147.40	0.30	0.55	0.79	0.22	0.24	1.37	1.05	0.14
GATA26B04	D2S1326	149.89	2.63	1.43	2.65	0.004	0.06	0.43	0.39	0.34
GGAA20G04	D2S1399	152.04	0.18	0.16	0.69	0.24	0.19	0.52	0.72	0.22
AFMb331yf1	D2S2299	157.55	0.88	0.40	1.23	0.11	0.09	0.12	0.22	0.41
GATA113F01		161.26	0.87	0.45	1.35	0.09	0.04	0.06	0.15	0.43
ATA27H09	D2S1353	164.51	0.00	–0.01	0.07	0.47	–0.01	0.04	–0.01	0.50
AFM164za5	D2S2195	169.41					0.37	0.53	0.85	0.19
GATA71D01	D2S1776	173.00	0.60	0.58	1.07	0.14	0.00	–0.01	–0.58	0.73
GATA194A05	D2S2981	180.79	–0.01	0.11	–0.04	0.52	0.07	0.16	0.39	0.34
AFM263xe1	D2S324	184.04	–0.02	–0.01	–0.18	0.57	–0.01	–0.02	–0.21	0.59
GATA65C03	D2S1391	186.21	–0.01	–0.02	–0.38	0.65	–0.01	–0.01	–0.42	0.67
CATA11H04	D2S425	193.26	0.00	–0.01	–0.99	0.84				
GATA52A04	D2S1384	200.43	–0.01	–0.01	–0.25	0.60	0.11	0.26	0.59	0.27
GATA29A06	D2S1369	206.13	0.00	0.00	–1.58	0.94				
GATA30E06	D1S1649	210.43	0.00	–0.01	0.03	0.49				
GATA26D05	D2S1327	210.43	–0.01	0.00	–0.67	0.75				
GATA4G12	D2S434	215.78	0.00	–0.01	–0.65	0.74				
GATA23D03	D2S1363	227.00	–0.01	–0.01	–0.29	0.61				
GATA12H10	D2S427	236.70	0.00	–0.01	–1.11	0.87				
AFM276ZF5	D2S338	250.54	0.00	0.00	–1.69	0.96				
AFM112yd4	D2S125	260.63	–0.01	0.00	–0.89	0.81				
q_tel		269.07								

Continued

we summarize the results by chromosome. The pairwise H-LODs are not discussed, as they are highly concordant with the Z_{all} score.

Chromosome 2 pairwise In Christchurch, over 19 cM there are six positive Z_{all} scores in a row, four of six greater than 1.0, and marker D2S1326 (dominant H-LOD = 2.63, $P = 0.0012$ and $Z_{all} = 2.65$, $P = 0.0011$) is roughly in the middle. In the Richmond sample (hereafter called Richmond), there is a cluster of seven markers — marker D2S442 gave a recessive H-LOD of 1.37 and a Z_{all} of 1.05, and all seven flanking markers have Z_{all} scores >0. D2S442 is about 2 cM from the D2S1326 which yielded the best result in the genome scan of the Christchurch sample. *Chromosome 2 multipoint*: In Christchurch, there is a broad peak which extends from 130–180 cM, and the sharp peak ($Z_{lr} = 2.63$, $P = 0.0012$) occurs at 149 cM at marker

Table 2 Continued

MARKER	*LOCUS*	*Christchurch*					*Richmond*			
		H-LOD			*NPL*		*H-LOD*		*NPL*	
		cM	*Dom*	*Rec*	Z_{all}	*p val.*	*Dom*	*Rec*	Z_{all}	*p val.*
(b) Pairwise results with chromosome 10 markers										
p_tel		0.00								
GATA88F09	D10S1435	4.32	–0.01	–0.01	–0.52	0.70				
AFM063xf4	D10S189	19.00	0.00	0.00	–1.61	0.95				
ATA31G11	D10S1412	28.31	0.07	0.14	0.53	0.30				
GGAA8G02	D10S1216	30.00	–0.01	–0.01	–0.27	0.61				
GAAT5F06	D10S2325	32.80	–0.01	–0.01	–0.23	0.59				
UT538	D10S506	35.41	0.00	0.00	–0.95	0.83				
GATA6E06	D10S674	41.79	–0.01	–0.01	–0.48	0.68				
GATA70E11	D10S1423	46.23	0.03	–0.01	0.05	0.48				
UT1357	D10S525	54.23	–0.01	–0.01	–0.45	0.67				
GATA73E11	D10S1426	59.03	0.17	0.04	0.64	0.26				
ATA20B07	D10S1220	70.23	0.15	0.10	0.42	0.34	0.62	0.14	0.54	0.28
ATA25D11	D10S1227	75.57	0.00	0.00	–1.11	0.87	0.13	0.09	0.56	0.28
ATA24F10	D10S1225	80.77	0.10	0.15	0.29	0.38				
GATA164F07	D10S2480	86.81	0.01	0.06	0.02	0.49	0.20	0.21	0.44	0.32
GATA87G01	D10S1432	93.92	0.10	0.10	0.47	0.32	0.64	0.86	1.25	0.09
GATA196C10		98.41	0.11	0.42	0.54	0.29	0.01	0.04	0.17	0.43
AFMa120xc5	D10S605	98.42					0.00	0.12	0.42	0.33
GGAT1A4	D10S2327	100.92					–0.01	–0.01	–0.30	0.62
GATA134F03	D10S2475	103.43	0.19	0.82	0.89	0.19	–0.02	0.00	–0.28	0.61
GATA115E01	D10S2470	112.58	0.03	0.00	0.18	0.43	0.00	0.00	–1.13	0.88
AFM269xg9	D10S571	117.42	0.07	0.28	0.47	0.32				
GGAA2F11	D10S677	117.42	1.30	1.23	1.69	0.05	–0.02	–0.01	–0.21	0.59
UT5126	D10S1147	122.84					–0.01	–0.01	–0.58	0.73
AFM094tc9	D10S192	124.27	0.66	0.79	1.46	0.07				
GATA64A09	D10S1239	125.41	1.40	1.29	1.79	0.04				
GATA114H09	D10S2469	127.11	1.67	2.20	2.22	0.01	–0.01	–0.01	–0.54	0.71
GATA48G07	D10S1237	134.70	0.07	0.58	0.61	0.27	0.00	0.00	–1.13	0.88
GATA127H01	D10S2473	138.47	0.02	0.06	0.26	0.40				
ATA29C03	D10S1230	142.78	0.24	0.36	0.73	0.23	–0.01	–0.01	–0.35	0.64
ATA24A05	D10S1223	156.27	0.00	0.01	0.14	0.44				
GGAA23C05	D10S1248	165.27	–0.01	–0.01	–0.72	0.76				
MFD187	D10S169	173.13	0.03	0.07	0.09	0.46				
q_tel		173.13								
(c) Pairwise results with chromosome 17 markers										
p_tel		0.00								
GAAT2C03	D17S1298	10.70	–0.02	0.11	–0.04	0.51				
GATA8C04	D17S974	22.20	0.07	0.01	0.20	0.42	–0.01	–0.01	–0.18	0.57
GATA10H07	D17S969	27.80	0.23	0.07	0.44	0.33				
AFM191XH12	D17S921	36.10	0.12	0.08	0.34	0.36	–0.01	0.00	–0.30	0.62
GATA185H04	D17S2196	44.60	0.09	–0.01	0.49	0.31	–0.01	0.00	–0.66	0.76
GGAA9D03	D17S1294	50.70	–0.01	–0.01	–0.38	0.65	–0.01	0.00	–0.69	0.77
GATA25A04	D17S1299	62.00	–0.02	–0.01	–0.01	0.50	0.30	–0.01	0.57	0.27
ATC6A06	D17S2180	66.90	0.36	0.01	0.77	0.22	0.83	0.58	1.55	0.05
AFM095tc5	D17S787	75.00	0.00	0.00	–1.17	0.88	0.18	0.16	0.63	0.25
GATA49C09	D17S1290	82.00	0.29	0.12	0.88	0.19	0.06	–0.01	0.51	0.29
GATA31B11	D17S2059	93.30	1.55	1.18	1.75	0.04	–0.01	0.00	–0.61	0.74
GATA28D11	D17S1301	100.00	0.51	0.40	0.84	0.20	–0.01	0.00	–0.79	0.80
ATA58A02	D17S2195	106.80	–0.01	0.09	0.08	0.47	–0.01	–0.01	–0.52	0.71
AFM044xg3	D17S784	116.90	0.28	0.47	1.07	0.14				
AFM217yd10	D17S928	126.50	0.00	–0.01	–1.00	0.84				
q_tel		126.50								

Continued

Table 2 Continued

MARKER	*LOCUS*	*Christchurch*					*Richmond*			
		H-LOD			*NPL*		*H-LOD*		*NPL*	
		cM	*Dom*	*Rec*	Z_{all}	*p val.*	*Dom*	*Rec*	Z_{all}	*p val.*
(d) Pairwise results with chromosome 16 markers										
p_tel		0.00								
GATA73G05	D16S2622	8.20	0.06	–0.01	0.16	0.44				
ATA3A07	D16S748	22.70	0.00	0.00	–2.00	0.98				
GATA62G05	D16S2619	28.30	–0.01	–0.01	–0.74	0.77				
AFM113xa9	D16S501	34.20	–0.01	0.12	0.23	0.41				
AFM049xd2	D16S403	43.90	0.08	0.16	0.45	0.32				
GATA71H05	D16S769	50.60	–0.01	0.00	–0.98	0.84				
GGAA3G05	D16S753	57.80	0.27	0.00	0.60	0.27	0.00	–0.01	–0.58	0.73
ATA55A11	D16S3396	63.80	0.54	0.07	0.95	0.17	–0.01	0.00	–0.25	0.60
GATA22F09	D16S3253	71.80	0.40	0.20	0.80	0.21	0.41	0.76	1.10	0.12
GATA67G11	D16S2620	81.20	–0.01	–0.01	–0.39	0.65	0.00	–0.01	–0.29	0.62
GATA81D12	D16S2624	87.60	0.00	–0.01	–1.03	0.85	–0.01	0.00	0.24	0.40
UT7223	D16S684	93.80	1.06	1.12	1.08	0.14	–0.02	–0.01	–0.11	0.54
UT7370	D16S686	105.20	0.00	–0.01	–1.22	0.89	0.00	–0.01	–0.99	0.85
AFM249xc5	D16S422	111.10	0.74	0.99	1.50	0.07	–0.01	0.14	0.36	0.35
GATA11C06	D16S539	124.70	–0.01	–0.01	–0.42	0.66				
q_tel		134.10								
(e) Pairwise results with chromosome 18 markers										
p_tel		0.00								
AFM178xc3	D18S59	0.00	0.26	0.26	0.75	0.23				
GATA178F11		2.80	0.40	0.35	0.91	0.18	0.04	0.18	0.27	0.39
ATA45G06	D18S1370	6.90	0.30	0.61	0.93	0.17	0.01	0.11	0.12	0.45
GATA88A12	D18S976	12.80	0.22	0.69	0.67	0.25	0.06	0.04	–0.14	0.56
UT1302	D18S391	18.70	0.38	0.15	0.79	0.21	0.09	0.24	0.31	0.37
AFMa101xf9	D18S1163	24.10	0.07	0.26	0.16	0.44	0.11	0.06	0.21	0.41
ACT1A01	D18S843	28.10	0.21	0.36	0.71	0.24	0.03	–0.01	0.24	0.40
AFMa054yg9	D18S1153	35.50	0.26	0.38	0.59	0.28	0.37	0.41	1.03	0.14
GATA11A06	D18S542	41.20	0.33	0.63	0.85	0.20	0.00	0.00	–1.41	0.93
D18S40	D18S40	47.00	0.76	0.94	1.28	0.10	0.00	0.00	–1.08	0.87
GATA41G05	D18S869	49.60	1.54	1.59	1.99	0.02	0.04	–0.01	0.19	0.42
GATA64G04	D18S877	54.40	0.08	0.14	0.48	0.31	0.21	0.22	0.53	0.29
GATA183H03		58.50	0.00	0.02	0.01	0.50	–0.01	–0.01	–0.41	0.67
GATA13	D18S535	64.50	0.04	0.01	0.15	0.44	–0.01	–0.01	–0.52	0.71
AFM292wg1	D18S473	71.30					0.05	0.14	0.28	0.38
GATA6D09	D18S851	74.90	0.57	0.85	1.19	0.12	0.04	0.04	0.43	0.33
ATA7D07	D18S1357	88.60	0.04	–0.01	0.14	0.44	–0.01	–0.01	0.06	0.47
GATA26C03	D18S1270	96.50	0.00	0.00	–1.24	0.89				
GATA114F07	D18S1358	108.00	–0.01	0.00	–0.30	0.62				
GATA177C03	D18S1371	115.90	0.32	0.06	0.50	0.31	0.03	–0.01	0.05	0.48
ATA1H06	D18S844	116.40	–0.01	–0.01	–0.35	0.64				
ATA38G09		122.60	0.75	0.60	1.07	0.14	0.38	0.18	0.64	0.25
q_tel		126.00								

Continued

D2S1326. In Richmond, the positive region is from 125 to 154 cM, with a maximum of Z_{lr} of 1.00 at 132 cM.

Chromosome 4 pairwise In Christchurch, from 109–129 cM, three of five markers have Z_{all} scores > 1 and H-LODs >0.5. From 72–114 cM, nine markers in a row have Z_{all} scores >0. In the Richmond sample, only 9/23 are positive and markers in the cluster of four all have very small Z_{all}. *Chromosome 4 multipoint*: In Christchurch, there are two peaks, $Z_{lr} = 1.36$ at 33 cM and $Z_{lr} = 1.34$ at 129 cM. In Richmond, consistent with the pairwise results, Z_{lr} is not positive and Z_{all} is somewhat negative.

Chromosome 10 pairwise In Christchurch, from 93–142 cM, twelve markers in a row have Z_{all} scores >0, with four of them from 117–127 cM greater than > 1.5. Marker D10S2469 gave a recessive H-LOD of 2.20

Table 2 Continued

MARKER	*LOCUS*	*Christchurch*					*Richmond*			
		H-LOD			*NPL*		*H-LOD*		*NPL*	
		cM	*Dom*	*Rec*	Z_{all}	*p val.*	*Dom*	*Rec*	Z_{all}	*p val.*
(f) Pairwise results with chromosome 4 markers										
p_tel		0.00								
UT5936	D4S2285	9.90	–0.01	0.08	0.24	0.40	–0.01	–0.01	–0.69	0.77
GATA22G05	D4S2366	12.93	0.55	0.33	1.00	0.16	–0.01	–0.01	–0.06	0.52
AFM157xg3	D4S403	25.90	0.08	0.02	0.41	0.34	0.07	0.08	0.46	0.31
GATA90B10	D4S2639	33.42	0.25	–0.01	0.65	0.26	–0.01	0.03	–0.17	0.57
ATA27C07	D4S2397	42.74	–0.01	–0.01	0.04	0.48	0.00	–0.01	–0.79	0.80
GATA170H04	D4S3353	50.53	–0.01	–0.01	–0.22	0.59				
GATA7D01	D4S1627	60.16	–0.01	–0.01	–0.25	0.60	0.00	0.00	–1.26	0.91
GATA61B02	D4S3254	63.58	0.00	–0.01	–0.97	0.83				
GATA28F03	D4S3248	72.52	0.17	0.34	0.76	0.22	–0.01	–0.01	–0.10	0.54
GATA24H01	D4S2367	78.43	0.41	0.11	0.77	0.22	–0.01	–0.01	–0.39	0.66
AFMa062yc1	D4S3042	83.29	0.00	–0.01	0.07	0.47	0.00	–0.01	–1.14	0.89
GATA10G07	D4S3243	88.35	0.24	0.09	0.43	0.33	0.04	0.15	0.44	0.32
ATA2A03	D4S2361	93.48	0.05	–0.01	0.16	0.44	–0.01	0.03	0.11	0.45
GGAA22C12	D4S2433	100.06	0.00	0.11	0.47	0.32	0.00	0.30	0.50	0.30
GATA2F11	D4S1647	104.94	0.28	0.32	0.79	0.21	0.01	–0.01	0.20	0.42
AFM186xa7	D4S411	109.02	0.63	0.79	1.30	0.10	–0.01	–0.01	0.06	0.47
GATA62A12	D4S2623	114.04	0.57	0.66	1.16	0.12	–0.01	–0.01	–0.18	0.57
AFM319yg9	D4S1611	121.61	–0.01	–0.01	–0.01	0.50	–0.01	0.08	0.16	0.43
UT1337	D4S1522	123.13	–0.02	0.00	0.10	0.46	0.09	0.05	0.69	0.23
ATA26B08	D4S2394	129.92	0.52	0.75	1.23	0.11	–0.01	–0.01	–0.28	0.61
GATA150B10		134.74	–0.01	–0.02	0.05	0.48	–0.01	–0.01	–0.51	0.70
GATA11E09	D4S1644	143.31					0.00	0.00	–1.06	0.87
GATA107	D4S1625	145.98	–0.01	–0.01	–0.72	0.76	0.11	0.07	0.66	0.24
GATA8A05	D4S1629	157.99	0.09	–0.02	0.18	0.43				
GATA27G03	D4S2368	167.55	–0.01	–0.01	–0.41	0.66	–0.01	–0.01	–0.05	0.52
GATA42H02	D4S2417	181.93	0.00	–0.01	–0.73	0.77				
UT1950	D4S1530	185.28	0.10	0.02	0.52	0.30	–0.01	–0.01	–0.05	0.52
GATA136E02	D4S3335	195.06	–0.01	–0.01	–0.30	0.62				
AFM165XC11	D4S408	195.06	–0.01	–0.01	–0.47	0.68				
GATA5B02	D4S1652	208.07	0.01	–0.02	–0.04	0.52				
q_tel		211.65								

($P = 0.003$) and Z_{all} of 2.22 ($P = 0.01$). In Richmond, there is no evidence for linkage from 100–142 cM, but a cluster of six markers in a row with Z_{all} scores >0 proximally, from 70–98 cM. Two of the four markers with Z_{all} scores > 1.5 in Christchurch were tested in Richmond and were negative. Thus there is little overlap in the signals from the two samples. *Chromosome 10 multipoint*: In Christchurch, the positive region extends from 85–149 cM, with the larger peak of $Z_{lr} = 2.43$ ($P = 0.008$) at 125 cM and a secondary peak at 105 cM. In Richmond, the maximum is $Z_{lr} = 1.21$ at 75 cM, and the curve remains positive through 93 cM. Z_{all} is strongly negative from 100–150 cM, the location of the larger peak in the Christchurch sample. More markers will be run across the length of chromosome 10 to clarify the Richmond peak location and maximum.

Chromosome 16 pairwise In Christchurch, from 57–111 cM there are some positive Z_{all} scores intermixed with negative scores (0.60, 0.95, 0.80, –0.39, –1.03, 1.08, –1.22, 1.50) — the 1.50 is from the scan marker D16S422. In Richmond, for D16S3253, $Z_{all} = 1.10$, but the remaining scores are small or negative. *Chromosome 16 multipoint*: In Christchurch, the maximum is $Z_{lr} = 0.85$ at 67 cM and Z_{all} is strongly negative towards the p telomere. In Richmond the maximum is $Z_{lr} = 1.11$ at 71 cM, which from the pairwise results appears to be due to D16S3253 alone.

Chromosome 17 pairwise From 66–166 cM in Christchurch, five of seven Z_{all} scores are positive, including a 1.75 and a 1.07. With the exception of marker D17S2059, the H-LODs are quite small. In Richmond, there is only a cluster of four including a Z_{all} of 1.55 at 66.90 cM. *Chromosome 17 multipoint*: The maximum in Christchurch is $Z_{lr} = 1.64$ at 93 cM, with a broad peak between 80 and 120 cM. The maximum in Richmond is $Z_{lr} = 1.6$ at 66 cM, the peak extending

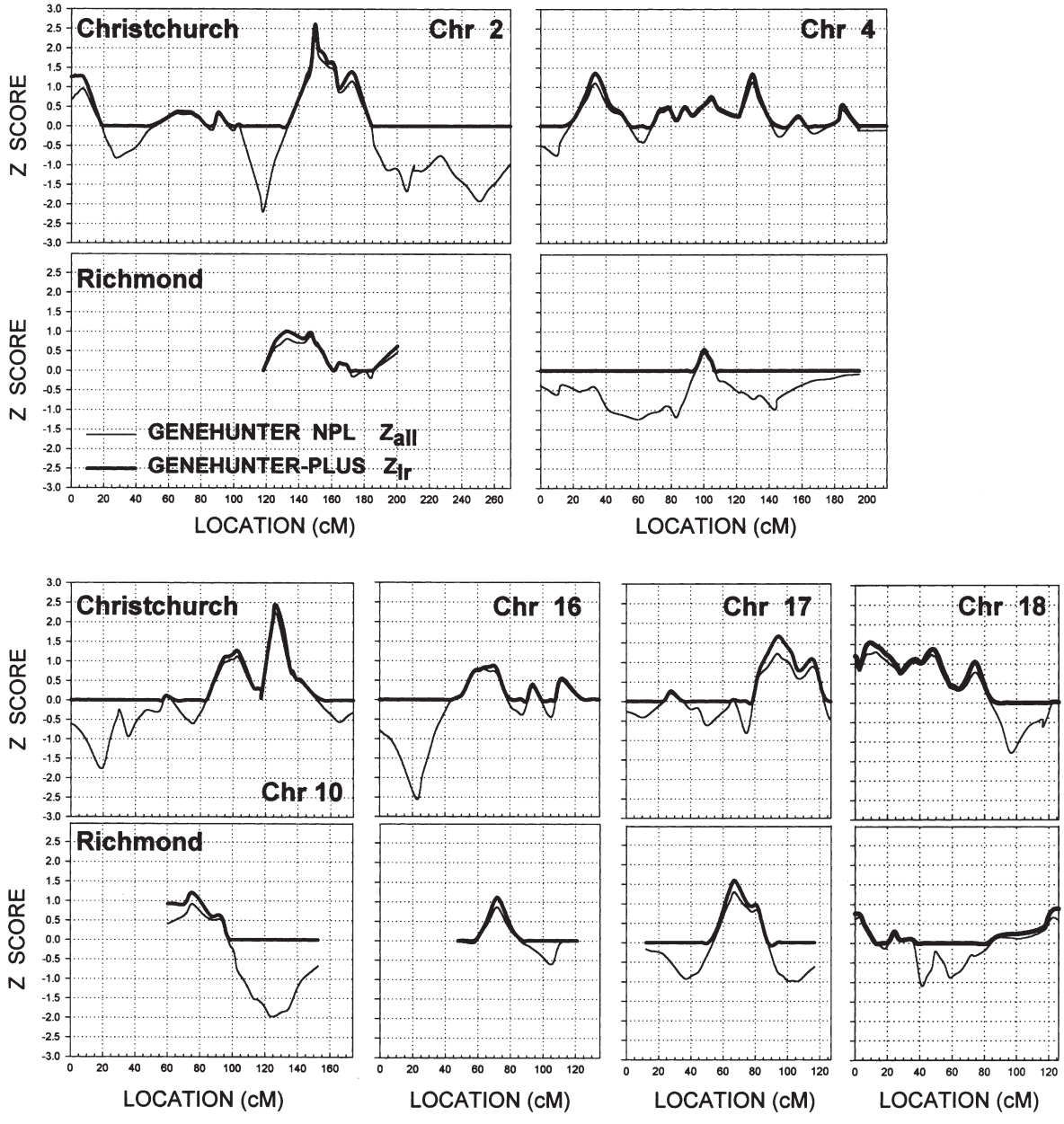

Figure 2 Multipoint nonparametric linkage results for chromosomes 2, 4, 10, 16, 17, and 18 for the samples from Christchurch (top panels) and Richmond (bottom panels). Z_{all} (thin line) was computed by GENEHUNTER (v 2.0). Z_{lr} (thick line) was computed by a modification of GENEHUNTER called GENEHUNTER-PLUS.[51] The markers and map used are shown in Table 2a–f.

from 56–82 cM. Thus there is little overlap in positive regions between the two samples.

Chromosome 18 pairwise Extending from 0–88 cM, there is a cluster of 16 markers in a row with Z_{all} scores >0, including the scan signal D18S869 ($Z_{all} = 1.99$, $P = 0.02$). In the Richmond sample, overall 13/18 markers are positive with two clusters of four positives each. *Chromosome 18 multipoint*: In Christchurch, the maximum is $Z_{lr} = 1.54$ at 9 cM, but the curve stays >0 from 0–80 cM. Z_{all} becomes negative at 80 cM. The Richmond sample maximizes at the p ($Z_{lr} = 0.73$) and q ($Z_{lr} = 0.90$) telomeres. Z_{all} becomes negative from 40–80 cM.

Richmond multipoint Z_{all} results — centimorgans positive vs negative The two samples are very similar in ascertainment methods, phenotyping, markers used, and methods of analysis. We may (with caution) use the Christchurch scan data as a benchmark to see if the Richmond results *differ overall* from what would be expected if there were no nicotine dependence genes present in Richmond in *any* of the six regions. From the multipoint graphs of Z_{all}, we calculated the per-

centage of the total distance for which Z_{all} is positive. Overall, the Christchurch genome was 28% positive (1023 of 3655 cM). The six Christchurch (whole) chromosomes followed up were 54% positive and the seventeen that were not followed up were 18% positive. As a whole, the regions followed up (677 cM total) were 39% positive in Richmond. If one omits chromosome 4 which is 211 cM long and is clearly negative, the Richmond sample was 51% positive (244 of 481 cM). By this measure, the Richmond multipoint results resemble more closely the six followup chromosomes (39 or 51% *vs* 54%) than the genome as a whole (39 or 51% *vs* 28%) and were quite unlike the seventeen chromosomes not followed up (39 or 51% *vs* 18%). The Richmond sample is smaller (about 70% in both number of families and number of affecteds), so if the gene has identical effects in each, the maximum Z_{all} should be smaller for Richmond than for Christchurch. If so, the reduced peak height should result[54] in a *reduced* percentage positive cM in Richmond. This not unreasonable possibility makes the relatively positive results observed in Richmond even more interesting.

Results from the Christchurch genome scan — other chromosomes Multipoint Z scores (Z_{all} and Z_{lr}) for the seventeen chromosomes which were not followed up are shown in Figure 3. Five chromosomes (1, 3, 8, 13, and X) have Z_{lr} peaks > 1.0. Four of these five peaks occur at the most telomeric marker — the significance of this is unclear. The results for these five chromosomes — given as *chromosome*. test: marker (location in cM; Z_{lr} or Z_{all}) — are as follows. *Chromosome 1*. Multipoint: (4.2; 1.08). Pairwise: D1S468 (4.20: 0.71), D1S1612 (16.20: 0.061), D1S1597 (29.90: –0.356). *Chromosome 3*. Multipoint: (224.88; 1.26). Pairwise: D3S3053 (181.87: 0.47), D3S2427 (188.29: 0.09), D3S2436 (203.28: 0.70), D3S2398 (209.41: –0.51), D3S2418 (215.84: 0.64), D3S1311 (224.88: 1.18). *Chromosome 8*. Multipoint: (159.24; 1.21). Pairwise: D8S1128 (139.53: –0.12), D8S1100 (154.02: 0.61), D8S373 (164.47: 0.94). *Chromosome 13*. Multipoint: (110.6; 1.04). Pairwise: D13S1809 (90.30: –0.15) and D13S285 (110.60: 0.94). *X Chromosome*. Multipoint: (142; 1.09). Pairwise: DXS6799 (107.42: 0.92), DXS6797 (112.89: –0.69), DXS6805 (120.53: –0.12), HPRT ((142.00: 0.87), D3S2390 (154.29: 0.56), DXS9908 (165.11: 0.40), DXS1200 (173.40: –0.74), DXS8061 (181.08: 1.2).

Discussion

To our knowledge, this is the first published report of a complete genome scan designed to detect genes that influence the risk of nicotine dependence. The goal of the scan was to identify regions that are likely to contain genes, and the goal of the followup was to distinguish between true and false signals. As we expected from the outset, we have made some progress towards these goals, but they have not been achieved. For traits in which both the phenotype and the genetic architecture are complex, it is now apparent that the relationship between the absolute values of linkage statistics (or their associated *P* values) and the true likelihood of a susceptibility gene being present is not straightforward.[56–62] Genome scan results from a wide variety of complex traits have shown that both highly 'significant' linkages and rapid, convincing 'replications' are uncommon. The presence of genetic heterogeneity and multiple, interacting genes each of moderate effect size, are likely to be the primary causes of the difficulty. With the sample sizes and genomic information currently available, low density genome scans for complex traits remain a crude and incomplete tool for gene identification — as a rule producing large potential target regions of debatable significance. This recurring roadblock requires that a more powerful and integrated set of tools be applied. These will include DNA microarrays and other emerging technologies, utilization of functionally annotated sequence databases, and the exploitation of the tremendous biological manipulations possible with model organisms—especially in the development of candidate gene assays and disease models. Better planning and greater attention to sample characteristics is essential, including trait-specific strategies for ascertainment—especially including access to populations with particular advantageous characteristics, gaining access to much larger sample sizes while keeping undesired increases in genetic and etiologic heterogeneity to a minimum. Phenotyping that is more standardized, detailed, and comprehensive, with the identification and validation of reliable endophenotypes will also be required. Finally, a new generation of analytical tools would be welcome, including more flexible and comprehensive tests of linkage and linkage disequilibrium that integrate multivariate, multilocus, and quantitative approaches and fully exploit their advantages, and as well as the refinement of routines for haplotype analysis capable of capitalizing on the upcoming wealth of information on human evolution.

Although the definition of nicotine dependence we used is fairly well justified, it is only one component of a larger set of interrelated phenotypes, which probably include elements of personality[17,63] and psychopathology.[17,64,65] This composite phenotype is likely to be influenced by multiple, interacting genes, some of which interact bidirectionally with environmental variables. Under such a scenario, studies of polygenic traits[66] that use only a single sample of moderate power are likely to detect only genes (if they exist at all) of relatively large effect. Individual studies will probably not provide an unambiguous, coherent, and reasonably complete picture of the genetic architecture. However, careful comparison of the results from independent genome scans may enable detection and confirmation of at least some genes whose signals were not obvious previously. To facilitate this type of comparison, we have chosen to present our dataset as a work in progress, and explain in detail why in the absence of 'significant' linkage results, we still think that these regions deserve some priority in other studies. We have presented as much of the data as possible to allow for

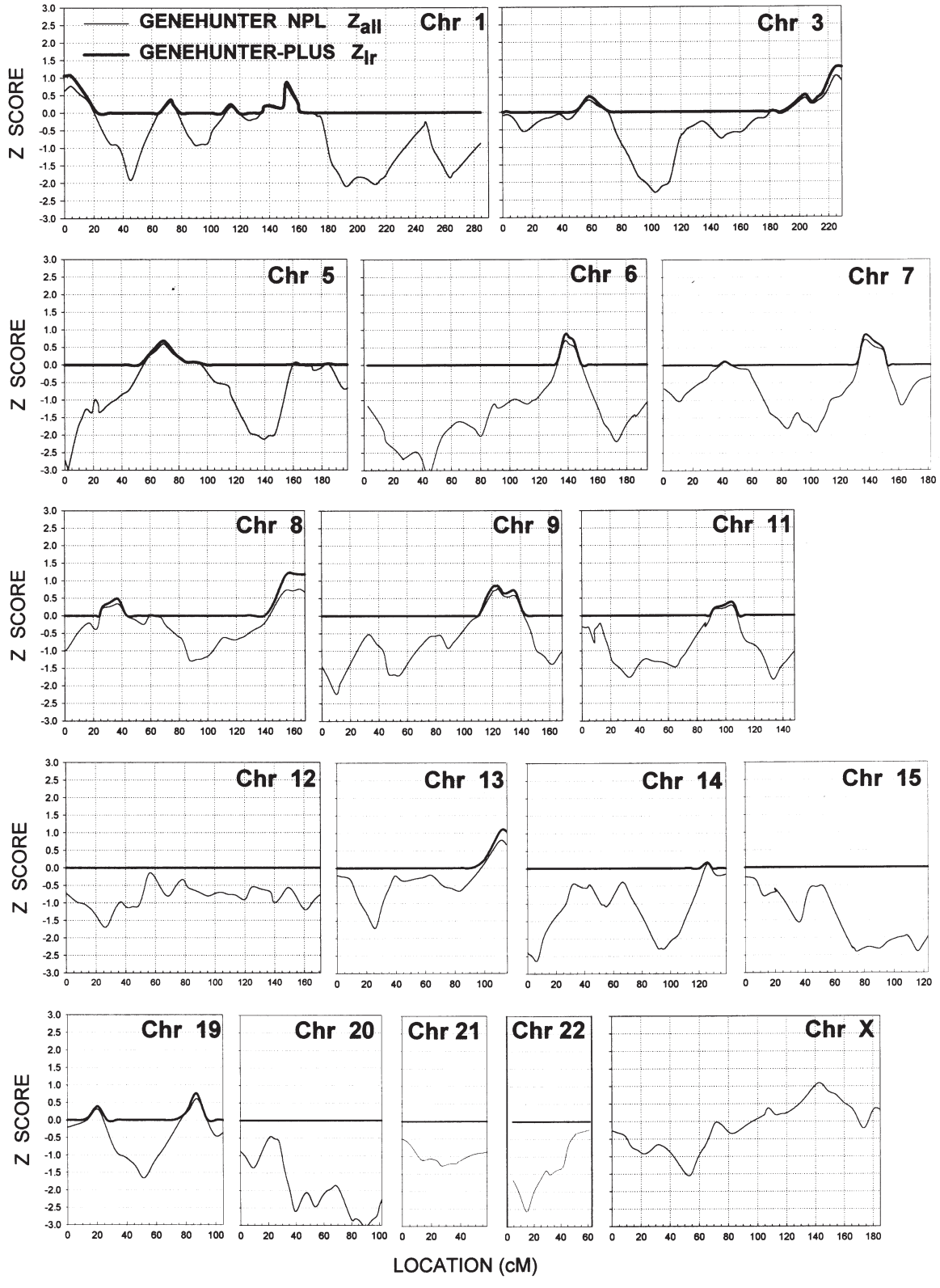

Figure 3 Multipoint nonparametric results for the other chromosomes in the genome scan of the Christchurch sample. Z_{all} and Z_{lr} were calculated as described in the Figure 2 legend. The version of GENEHUNTER-PLUS we used does not process X chromosome markers, so only Z_{all} is shown.

independent evaluation of strengths and weaknesses, and will provide further information upon request.

If only the most simple evaluation (ie comparison to the traditional or proposed thresholds of significance) of the data is made, then none of the putative loci are 'significant' in either sample. It would not be unexpected, therefore, if most of the signals are false. The pertinent question then becomes in the *absence* of *any* detectable susceptibility genes in *either* sample, how likely is it that the observed (and interrelated) patterns in the data are due to chance? For example, we would like to be able to calculate accurately how likely it is that the clusters (the pattern of consecutive Z_{all} scores > 0) found in the Christchurch sample on chromosomes 4 (nine markers over 62 cM), 18 (16 markers over 88 cM), and 10 (15 markers over 76 cM) have *all* occurred by chance? In the Christchurch scan, chromosomes 2 and 10 had the two highest pairwise H-LOD scores (2.63 and 2.20 respectively) and Z_{all} scores (2.65 and 2.22 respectively) and the greatest multipoint Z_{lr} maximums (2.63 and 2.43 respectively). In the Richmond followup of chromosomes 2 and 10 (154 cM total), we found clusters of eight and six markers respectively. Since clusters of size six or greater occurred only three times in the 25-fold larger scan, it seems rather unlikely that the Richmond results on these two chromosomes were due to chance. Finally, using the Christchurch scan data as a benchmark, a number of different (but not independent) comparisons suggest that, when considered as a whole, the results from the Richmond sample in the six regions tested (677 cM total) are notably more positive than would be expected if the markers or the regions had been chosen at random.

On chromosome 2, markers over 82 cM were tested in both samples, and the most positive pairwise results in each are from markers that are only 2 cM apart. On chromosome 10, however, the Richmond multipoint values are quite negative at the location of the larger Christchurch peak, although there is some overlap (from 85 to 93 cM) with the secondary peak. On chromosome 16, the peaks in the two samples are located at the same position, but on chromosome 17 they are offset by about 25 cM. The poor overlap between the multipoint peaks on chromosomes 10 and 17 in the two samples is not surprising, as this has been predicted by several simulation studies. It has been shown that the size of the confidence intervals for gene location increases quite rapidly as the relative risk or the proportion of alleles shared by affected relatives decreases.[67] Even with large sample sizes, and a gene conferring a sixfold increased risk to sibs, if one confines the search to the region of maximal sharing, the gene will be missed over 60% of the time.[68] Hovatta and colleagues noted[69] considerable error in the position of the maximum lod scores, especially for sample sizes of 100 or 200 sib pairs and for the less significant findings. We have also examined this issue by simulation.[70] For example, with 200 families (each with two parents and two affected offspring), assuming 25% of families were linked, and setting an intermarker distance of 2 cM, analysis of 500 replicates showed that the GENEHUNTER multipoint peak varied markedly. The standard deviation was 25 cM, and the 95% confidence interval was 100 cM. More importantly, for a number of genes whose locations are now known (for example the MHC region and the insulin gene in type 1 diabetes[71]), the magnitude and position of the maximum signal was routinely quite variable across studies.

The data presented are preliminary in a number of ways. First, there are more potential positives from Christchurch (chromosomes 1, 3, 8, 13 and X) to follow up. Second, we are working on higher density followups as well as completing the scan on the Richmond sample using the same marker sets. Third, we present here only the results of standard parametric and nonparametric analyses using GENEHUNTER to analyze a categorical phenotype, where individuals with FTQ scores greater than or equal to 7 are considered affected. We have done a few preliminary analyses of chromosomes 2 and 10 using the QUANTITATIVE (SIBS) options of GENEHUNTER (v 2.0), treating the FTQ scores and two other phenotypic measures as individual quantitative traits. The results were not an improvement over the standard methods in terms of significance levels or gene localization (data not shown). Exploratory analyses that may provide more power are planned, including multipoint quantitative,[72] multilocus[73,74] and multivariate[75] approaches.

In parallel to these two genome scans, we have begun a systematic evaluation of functional candidate genes in the nicotinic cholinergic, dopaminergic, GABAergic, and cytochrome systems, among others. We are identifying new single nucleotide polymorphisms (SNPs) by sequence alignment, direct sequencing, and other methods, and are testing multiple SNPs per gene to analyze haplotypes. Candidate gene testing is being done in parallel in the families from Christchurch and Richmond, and in a case-control study of an independent set of individuals from the Virginia Twin Study.[17] In this study, we are attempting to make some headway on the fundamental question of the probable overlap between genes contributing to smoking initiation and those contributing to nicotine dependence by contrasting three groups: lifelong non-smokers ($n = 338$), non-dependent regular smokers (FTQ scores 0–3; $n = 258$), and nicotine-dependent smokers (FTQ scores 7–11; $n = 337$). We are also actively accessing new families containing concordant and/or discordant sib pairs and families with available parents to be used for tests of linkage disequilibrium by TDT, and by Sib-TDT.[76,77]

Even though the linkage results presented here were not 'significant' by standard statistical criteria, they are nevertheless encouraging, especially that the Richmond sample appears to be more positive than one would expect if no susceptibility gene(s) were present. After completion of the scan of the Richmond sample, we will analyze that scan independently. We will then combine the genotype data with those from Christchurch and apply more powerful analytical tools to the

combined sample. These linkage studies will be integrated with case-control and family-based association studies of functional (and eventually positional), candidate genes. We hope that this report serves to stimulate and enhance the search in independent samples for genes contributing to nicotine dependence.

Acknowledgements

We thank Alison Pickering, Lisa Halberstadt, Stacey Garnett, and Barbara Brooke for their excellent assistance in completing this project. We are indebted to the subjects who volunteered for this study. This project was supported by grant DA-10228 to KSK from the United States National Institute of Drug Abuse (NIDA).

References

1 World Health Organization. *Tobacco or Health. A Global Status Report*. WHO: Geneva, 1997, pp 1–49.
2 Peto R, Lopez AD, Boreham J, Thun M, Heath C, Doll R. Mortality from smoking worldwide. *Br Med Bull* 1996; **52**: 12–21.
3 US Department of Health E and Welfare. *Smoking and Health: A Report of the Surgeon General*. US Department of Health, Education and Welfare: Washington, DC, 1979.
4 Schoenborn CA, Boyd G. Smoking and other tobacco use: United States, 1987. National Center for Health Statistics. *Vital Health Statistics* 1989; **10**: 169.
5 US Department of Health and Human Services. *The Health Consequences of Smoking: Cancer. A Report of the Surgeon General*. US Department of Health and Human Services, Public Health Service, Office on Smoking: Washington, DC, 1982.
6 US Department of Health and Human Services. *The Health Consequences of Smoking: Cardiovascular Disease. A Report of the Surgeon General*. US Department of Health and Human Services, Public Health Service, Office on Smoking: Washington, DC, 1983.
7 Mathias R. NIDA, the FDA, and nicotine regulation. *NIDA Notes* 1994; **9**: 5–7.
8 Centers for Disease Control and Prevention. Cigarette smoking — attributable mortality and years of potential life lost — United States, 1988–1991. *MMWR* 1993; **42**: 37–39.
9 Schelling TC. Addictive drugs: the cigarette experience. *Science* 1992; **255**: 430–433.
10 US Department of Health and Human Services. *The Health Consequences of Smoking: Nicotine Addiction. A Report of the Surgeon General, 1988*. Department of Health and Human Services: Washington, DC, 1988.
11 Hughes JR, Gust SW, Pechacek TF. Prevalence of tobacco dependence and withdrawal. *Am J Psychiatry* 1987; **144**: 205–208.
12 Rose JE. Nicotine addiction and treatment. *Annu Rev Med* 1996; **47**: 493–507.
13 Boomsma DI, Koopmans JR, Lorenz JP, Doornen V, Orlebeke JF. Genetic and social influences on starting to smoke: a study of Dutch adolescent twins and their parents. *Addiction* 1994; **89**: 219–226.
14 Heath AC, Madden PAF. Genetic influences on smoking behavior. In: Turner JR, Cardon LR, Hewitt JK (eds). *Behavior Genetic Approaches in Behavioral Medicine*. Plenum Press: New York, 1995, pp 45–66.
15 True WR, Heath AC, Phil D, Scherrer JF, Waterman B, Goldberg J *et al*. Genetic and environmental contributions to smoking. *Addiction* 1997; **92**: 1277–1287.
16 Swan GE, Carmelli D. Behavior genetic investigation of cigarette smoking and related issues in twins. In: Blum K, Noble EP (eds). *Handbook of Psychiatric Genetics*. CRC Press: New York, 1997, pp 387–406.
17 Kendler KS, Neale MC, Sullivan PF, Gardner CO, Prescott CA. A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol Med* 1999; **29**: (in press).
18 Sullivan PF, Kendler KS. The genetic epidemiology of smoking. *Nicotine and Tobacco Research* 1999; (in press).
19 Marks MJ, Stitzel JA, Collins AC. Genetic influences on nicotine responses. *Pharmacol Biochem Behav* 1989; **33**: 667.
20 Marks MJ, Romm E, Campbell SM, Collins AC. Variation of nicotinic binding sites among inbred strains. *Pharmacol Biochem Behav* 1989; **33**: 679.
21 Collins AC, Marks MJ. Chronic nicotine exposure and brain nicotinic receptors — influence of genetic factors. *Prog Brain Res* 1989; **79**: 137.
22 Collins AC, Marks MJ. Progress towards the development of animal models of smoking-related behaviors. *J Addict Dis* 1991; **10**: 109–126.
23 Collins AC, Marks MJ. Progress towards the development of animal models of smoking-related behaviors. *Behavioral and Biochemical Issues in Substance Abuse*. The Haworth Press: 1991, pp 109–126.
24 Stein EA, Pankiewicz MD, Harsch HH, Cho J-K, Fuller SA, Hoffman RG *et al*. Nicotine-induced limbic cortical activation in the human brain: a functional MRI study. *Am J Psychiatry* 1998; **155**: 1009–1015.
25 Pontieri FE, Tanda G, Orzi F, Di Chiara G. Effects of nicotine on the nucleus accumbens and similarity to those of addictive drugs. *Nature* 1996; **382**: 255–257.
26 Pontieri FE, Passarelli F, Calò L, Caronti B. Functional correlates of nicotine administration: similarity with drugs of abuse. *J Mol Med* 1998; **76**: 193–201.
27 Picciotto MR. Common aspects of the action of nicotine and other drugs of abuse. *Drug Alcohol Depend* 1998; **51**: 165–172.
28 Gotti C, Fornasari D, Clementi F. Human neuronal nicotinic receptors. *Prog Neurobiol* 1997; **53**: 199–237.
29 Gray R, Rajan AS, Radcliffe KA, Yakehiro M, Dani JA. Hippocampal synaptic transmission enhanced by low concentrations of nicotine. *Nature* 1996; **383**: 713–716.
30 McGehee DS, Role LW. Memories of nicotine. *Nature* 1996; **383**: 670–671.
31 Picciotto MR, Zoli M, Rimondini R, Léna C, Marubio LM, Pich EM *et al*. Acetylcholine receptors containing the β2 subunit are involved in the reinforcing properties of nicotine. *Nature* 1998; **391**: 173–177.
32 Nakajima M, Yamamoto Y, Nunoya K-I, Yokoi T, Nagashima K, Funae Y *et al*. Role of human cytochrome P4502A6 in C-oxidation of nicotine. *Drug Metab Dispos* 1998; **24**: 1212–1217.
33 Berkman CE, Park SB, Wrighton SA, Cashman JR. *In vitro-in vivo* correlations of human (S)-nicotine metabolism. *Biochem Pharmacol* 1995; **50**: 565–570.
34 Fernandez-Salguero P, Hoffman SM, Cholerton S, Mohrenweiser H, Raunio H, Rautio A *et al*. A genetic polymorphism in coumarin 7-hydroxylation: sequence of the human CYP2A genes and identification of variant CYP2A6 alleles. *Am J Hum Genet* 1995; **57**: 653–660.
35 Yokoi T, Kamataki T. Genetic polymorphism of drug metabolizing enzymes: new mutations in CYP2D6 and CYP2A6 genes in Japanese. *Pharm Res* 1998; **15**: 517–524.
36 Pianezza ML, Sellers EM, Tyndale RF. Nicotine metabolism defect reduces smoking. *Nature* 1998; **393**: 750.
37 Oscarson M, Gullsten H, Rautio A, Bernal ML, Sinues B, Dahl M-L *et al*. Genotyping of human cytochrome P450 2A6 (CYP2A6) a nicotine C-oxidase. *Fed Eur Biochem Soc* 1998; **438**: 201–205.
38 Lerman C, Caporaso N, Main D, Audrain J, Boyd NR, Bowman ED *et al*. Depression and self-medication with nicotine: the modifying influence of the dopamine D4 receptor gene. *Health Psychol* 1998; **17**: 56–62.
39 Lerman C, Shields PG, Audrain J, Main D, Cobb B, Boyd NR *et al*. The role of the serotonin transporter gene in cigarette smoking. *Cancer Epidemiol Biomarkers Prev* 1998; **7**: 253–255.
40 Rossing MA. Genetic influences on smoking: candidate genes. *Environ Health Perspect* 1998; **106**: 231–238.
41 Fagerstrom KO. Measuring degree of physical dependence to tobacco smoking with reference to individualization of treatment. *Addict Behav* 1978; **3**: 235–241.
42 Fagerstrom KO, Schneider NG. Measuring nicotine dependence: a review of the Fagerstrom Tolerance Questionnaire. [Review]. *J Behav Med* 1989; **12**: 159–182.
43 Jarvik ME, Schneider NG. Degree of addiction and effectiveness of nicotine gum therapy for smoking. *Am J Psychiatry* 1984; **141**: 790–791.

44 Hjalmarson A. Effect of nicotine chewing gum in smoking cessation. *JAMA* 1984; **252**: 2835–2838.
45 Hughes JR, Hatsukami DK. Short-term effects of nicotine gum. In: Grabowski J, Hall SM (eds). *Pharmacological Adjuncts in Smoking Cessation*. Dept of Health and Human Services: Washington, DC, 1985.
46 Richards B, Skoletsky J, Shuber AP, Balfour R, Stern RC, Dorkin HL *et al*. Multiplex PCR amplification from the CFTR gene using DNA prepared from buccal brushes/swabs. *Hum Mol Genet* 1993; **2**: 159–163.
47 Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998; **63**: 861–869.
48 Hecker KH, Roux KH. High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. *Biotechniques* 1996; **20**: 478–480.
49 Boehnke M, Cox NJ. Accurate inference of relationships in sib-pair linkage studies. *Am J Hum Genet* 1997; **61**: 423–429.
50 Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES. Parametric and nonparametric linkage analysis: a unified multipoint approach. *Am J Hum Genet* 1996; **58**: 1347–1363.
51 Kong A, Cox NJ. Allele-sharing models: LOD scores and accurate linkage tests. *Am J Hum Genet* 1997; **61**: 1179–1188.
52 Davis S, Weeks DE. Comparison of nonparametric statistics for detection of linkage in nuclear families: single-marker evaluation. *Am J Hum Genet* 1997; **61**: 1431–1444.
53 Goldin LR, Chase GA, King TM, Badner JA, Gershon ES. Use of exact and adjusted liability scores to detect genes affecting common traits. *Genet Epidemiol* 1995; **12**: 765–769.
54 Terwilliger JD, Shannon WD, Lathrop GM, Nolan JP, Goldin LR, Chase GA *et al*. True and false positive peaks in genomewide scans: applications of length-biased sampling to linkage mapping. *Am J Hum Genet* 1997; **61**: 430–438.
55 Goldin LR, Chase GA. Improvement of the power to detect complex disease genes by regional inference procedures. *Genet Epidemiol* 1997; **14**: 785–789.
56 Eaves LJ. Effect of genetic architecture on the power of human linkage studies to resolve the contribution of quantitative trait loci. *Hered (Edinburgh)* 1994; **72**: 175–192.
57 Morton NE. Significance levels in complex inheritance. *Am J Hum Genet* 1998; **62**: 690–697.
58 Vieland VJ. Statistical genetics '98: Bayesian linkage analysis, or: How I learned to stop worrying and love the posterior probability of linkage. *Am J Hum Genet* 1998; **63**: 947–954.
59 Vieland VJ, Hodge SE. Statistical evidence: a likelihood paradigm (book review). *Am J Hum Genet* 1998; **63**: 283–289.
60 Greenberg DA, Abreu P, Hodge SE. The power to detect linkage in complex disease by means of simple LOD-score analyses. *Am J Hum Genet* 1998; **63**: 870–879.
61 Elston RC. 1996 William Allan Award Address. Algorithms and inferences: the challenge of multifactorial diseases. *Am J Hum Genet* 1997; **60**: 255–262.
62 Elston RC. Methods of linkage analysis — and the assumptions underlying them. *Am J Hum Genet* 1998; **63**: 931–934.
63 Heath AC. Personality and the inheritance of smoking behavior: a genetic perspective. *Behav Genet* 1995; **25**: 103–117.
64 Kendler KS, Neale MC, MacLean CJ, Heath AC, Eaves LJ, Kessler RC. Smoking and major depression: a causal analysis. *Arch Gen Psychiatry* 1993; **50**: 36–43.
65 Freedman R, Coon H, Myles-Worsley M, Orr-Urtreger A, Olincy A, Davis A *et al*. Linkage of a neurophysiological deficit in schizophrenia to a chromosome 15 locus. *Proc Natl Acad Sci USA* 1997; **94**: 587–592.
66 Paterson AH. Molecular dissection of quantitative traits: progress and prospects. *PCR Meth Appl* 1995; **5**: 321–333.
67 Kruglyak L, Lander ES. High-resolution genetic mapping of complex traits. *Am J Hum Genet* 1995; **56**: 1212–1223.
68 Kruglyak L, Lander ES. Limits on fine mapping of complex traits. *Am J Hum Genet* 1996; **58**: 1092–1093.
69 Hovatta I, Lichtermann D, Juvonen H, Suvisaari J, Terwilliger JD, Arajarvi R *et al*. Linkage analysis of putative schizophrenia gene candidate regions on chromosomes 3p, 5q, 6p, 8p, 20p, and 22q in a population-based sampled Finnish family set. *Mol Psychiatry* 1998; **3**: 452–457.
70 Roberts SB, MacLean CJ, Neale MC, Eaves LJ, Kendler KS. Replication of linkage studies of complex traits: an examination of the variation in position of the peak LOD score. *Am J Hum Genet* (submitted).
71 Mein CA, Esposito L, Dunn MG, Johnson GCL, Timms AE, Goy JV *et al*. A search for type 1 diabetes susceptibility genes in families from the United Kingdom. *Nature Genet* 1998; **19**: 297–300.
72 Almasy L, Blangero J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am J Hum Genet* 1998; **62**: 1198–1211.
73 Borecki IB, Blangero J, Rice T, Pérusse L, Bouchard C, Rao DC. Evidence for at least two major loci influencing human fatness. *Am J Hum Genet* 1998; **63**: 831–838.
74 Cloninger CR, Van Eerdewegh P, Goate A, Edenberg HJ, Blangero J, Hesselbrock V *et al*. Anxiety proneness linked to epistatic loci in genome scan of human personality traits. *Am J Med Genet* 1998; **81**: 313–317.
75 Allison DB, Thiel B, St Jean P, Elston RC, Infante MC, Schork NJ. Multiple phenotype modeling in gene-mapping studies of quantitative traits: power advantages. *Am J Hum Genet* 1998; **63**: 1190–1201.
76 Curtis D. Use of siblings as controls in case-control association studies. *Ann Hum Genet* 1997; **61**: 319–333.
77 Spielman RS, Ewens WJ. A sibship test for linkage in the presence of association: the sib transmission/disequilibrium test. *Am J Hum Genet* 1998; **62**: 450–458.