



Extended Intermarker Linkage Disequilibrium in the Afrikaners

Diana Hall, Ellen M. Wijsman, J. Louw Roos, et al.

Genome Res. 2002 12: 956-961

Access the most recent version at doi:[10.1101/gr.136202](https://doi.org/10.1101/gr.136202)

References This article cites 25 articles, 5 of which can be accessed free at:
<http://genome.cshlp.org/content/12/6/956.full.html#ref-list-1>

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).



To subscribe to *Genome Research* go to:
<https://genome.cshlp.org/subscriptions>

Cold Spring Harbor Laboratory Press

Extended Intermarker Linkage Disequilibrium in the Afrikaners

Diana Hall,¹ Ellen M. Wijsman,² J. Louw Roos,³ Joseph A. Gogos,⁴
and Maria Karayiorgou^{1,5}

¹The Rockefeller University, Human Neurogenetics Laboratory, New York, New York 10021, USA; ²University of Washington, Division of Medical Genetics, Department of Medicine; and Department of Biostatistics, Seattle, Washington 98195, USA;

³University of Pretoria Department of Psychiatry & Weskoppies Hospital, Pretoria, 0083, Republic of South Africa; ⁴Columbia University, College of Physicians & Surgeons, Department of Physiology and Cellular Biophysics and Center for Neurobiology & Behavior, New York, New York 10032, USA.

In this study we conducted an investigation of the background level of linkage disequilibrium (LD) in the Afrikaner population to evaluate the appropriateness of this genetic isolate for mapping complex traits. We analyzed intermarker LD in 62 nuclear families using microsatellite markers covering extended chromosomal regions. The markers were selected to allow the first direct comparison of long-range LD in the Afrikaners to LD in other demographic groups. Using several statistical measures, we find significant evidence for LD in the Afrikaners extending remarkably over a 6-cM range. In contrast, LD decays significantly beyond 3-cM distances in the other founder and outbred populations examined. This study strongly supports the appropriateness of the Afrikaner population for genome-wide scans that exploit LD to map common, multigenic disorders.

The positional cloning of genes underlying common diseases is one of the greatest challenges in biomedical research. However, traditional analysis of the cosegregation of disease and genetic markers within pedigrees (linkage analysis) shows limited power when complex inheritance patterns are present. A number of approaches that exploit linkage disequilibrium (LD) have been proposed to circumvent this problem, including population-based allelic association methods or joint analysis of linkage and allelic association (Jorde 1995; Risch and Merikangas 1996; Freimer et al. 1997). Use of families from founder populations is expected to facilitate both designs. The measured LD level between a marker and a disease allele is determined by several elements, including the distance between the marker and the disease gene, the age of the disease allele, the frequency of the linked marker allele, and other less predictable forces such as genetic drift, rate of population expansion, genetic bottlenecks, and admixture. As expected theoretically and observed empirically, LD decays as distance increases. However, several recent studies indicate that within a short distance range, the contribution of recombination to the level of LD is negligible compared with the other forces (Shifman and Darvasi 2001). As distance increases, however, the increased recombination frequency is expected to erode any association and therefore recombination and distance become the primary elements that determine the level of LD.

Patients from founder populations likely inherited, from a common ancestor, segments of chromosomes containing a disease gene and surrounding marker loci for which a shared haplotype may be identified. The relatively recent origins of a certain founder population make it likely that chromosomal regions identical by descent surrounding a disease allele will

be larger than in outbred populations. Therefore, optimally designed genome-wide scans that exploit LD and involve founder populations of recent ancestry should require significantly fewer markers or significantly smaller sample size than similar scans in outbred populations (Chapman and Wijsman 1998). Many questions still remain open regarding the optimal design of such genome-wide mapping studies, including the appropriate demographic history of the founder population, the appropriate spacing of genetic markers, and the appropriate thresholds for significance testing. Addressing these issues depends partly on understanding the forces governing the distribution of background LD (bLD) in the genome. In the present study, we examined the distribution of bLD in a set of Afrikaner families from Tshwane (formerly known as Pretoria), Limpopo (formerly Northern Transvaal), and Mpumalanga (formerly the Eastern Transvaal) that were collected for a genetic study of schizophrenia.

The Afrikaners, who total approximately 3 million, descended from a small number of original settlers (who soon became known as "Boere", or farmers). In 1652, the first immigrants (just >1000), primarily of Dutch origin, settled in the Cape. They later spread inland, among other places, to Tshwane and Limpopo. The communities that were eventually established were geographically isolated. Cultural considerations, including language differences (Afrikaans language derived from Dutch) and religious practices (most Afrikaners were members of the Dutch Reformed Church), further maintained the isolation. Consanguinity was common, especially in early generations. Although in more recent years some admixture has occurred, the population growth over 13–15 generations was almost entirely through reproduction, as immigration subsequent to the founding was minimal. The demographic history of this population is reflected in the unusually high frequency of certain rare Mendelian disorders, the unusually low diversity of the associated allelic variants, and the unusually large extent (8–11 cM) of conserved haplotypes

⁵Corresponding author.

E-MAIL karayim@mail.rockefeller.edu; FAX (212) 327-7329.

Article and publication are at <http://www.genome.org/cgi/doi/10.1101/gr.136202>.

Table 1. Features and Relative Locations of Microsatellite Markers

Marker	Distance to the next marker		Heterozygosity observed
	(Kosambi cM) ^a	(Mb) ^b	
D18S851	0.4	0.1	0.70
D18S484	0.6	n.r.	0.73
AFM318xd5	0.2	n.r.	0.40
D18S1156	1.8	1.1	0.59
D18S487	1.8	1.8	0.75
D18S35	0.4	0.4	0.68
D18S69	0.6	0.7	0.76
D18S39	0.7	n.r.	0.78
D18S41			0.58
D4S425	1.1	0.7	0.72
D4S404	5.6	2.6	0.77
D4S1091	0.9	1.2	0.41
D4S391	1.1	n.r.	0.82
D4S230	0.0	n.r.	0.87
D4S418	1.3	0.2	0.77
D4S1618	55	66	0.54
D4S423	0.0	0.7	0.86
D4S410	0.0	1.1	0.50
D4S414			0.87

^aData from Eaves et al. 2000 for chromosome 18 map; Marshfield database for chromosome 4.

^bHuman Genome Working Draft-Santa Cruz; April 1, 2001, freeze. n.r., not reported.

around disease genes (Hayden et al. 1980; Brink et al. 1987; Rosendorff et al. 1987; Leitersdorf et al. 1989; Torrington and Viljoen 1991; Brink et al. 1995; Pronk et al. 1995; Goldman et al. 1996; Warnich et al. 1996; Groenewald et al. 1998; Roby et al. 1999).

We wished to compare patterns of LD in the Afrikaners with other founder populations, as well as outbred populations. To assess the level of bLD in the Afrikaner population, we used a total of 62 families collected from Tshwane, Limpopo, and Mpumalanga as part of an ongoing genetic study of schizophrenia.

Table 3. Comparison of LD between Locus Pairs at Increasing Ranges of Genetic Distance across Populations

Population	Proportion of locus pairs with $D'_{m} > 0.30$					
	0–1 cM	1–2 cM	2–3 cM	3–4 cM	4–5 cM	5–6.5 cM
Afrikaners	3/7	3/7	0/4	1/4	0/7	2/7
Finns	1/7	0/7	0/4	0/4	0/7	0/7
Sardinians	0/7	0/7	0/4	0/4	0/7	0/7
British	0/7	0/7	0/4	0/4	0/7	0/7
USA	0/7	0/7	0/4	0/4	0/7	0/7
Population	Proportion of locus pairs with $G^2 > 0.05$					
	0–1 cM	1–2 cM	2–3 cM	3–4 cM	4–5 cM	5–6.5 cM
Afrikaners	7/7	7/7	4/4	4/4	7/7	7/7
Finns	6/7	4/7	2/4	0/4	0/7	0/7
Sardinians	3/7	1/7	0/4	0/4	0/7	0/7
British	4/7	0/7	0/4	0/4	0/7	0/7
USA	3/7	0/7	0/4	0/4	0/7	0/7

LD, linkage disequilibrium.

Table 2. Associations Observed Between Pairs of Loci

Locus 1	Locus 2	FET p value < .01	cM
AFM318xd5	D18S1156	.0001*	0.2
D18S851	D18S484	<.00001*	0.4
D18S69	D18S39	<.00001*	0.6
D18S39	D18S41	<.00001*	0.7
D18S851	D18S1156	<.00001*	1.2
D18S69	D18S41	<.00001*	1.3
D18S35	D18S41	.0022	1.7
D18S1156	D18S487	.0088	1.8
AFM318xd5	D18S487	.0090	2
D18S487	D18S69	.0094	2.2
D18S487	D18S39	<.00001*	2.8
D18S487	D18S41	<.00001*	3.5
D18S1156	D18S35	<.00001*	3.6
D18S1156	D18S69	<.00001*	4
D18S1156	D18S39	<.00001*	4.6
D18S851	D18S69	.0096	5.2
D18S1156	D18S41	<.00001*	5.3
AFM318xd5	D18S41	<.00001*	5.5

Shown are locus associations detected using the Fisher exact test (FET).

The reported distances were computed using genetic maps indicated in Table 1.

*Locus pairs significant at p value < .01 after Bonferroni correction.

RESULTS

We examined nine microsatellite markers from a region in chromosome 18, investigated by Eaves et al. (2000) in samples from two genetically isolated populations (Finland and Sardinia) and two outbred populations (Britain and USA). The markers we used and their genetic map distances are reported in Table 1. The choice of the marker set was dictated by the possibility to compare directly the pattern of bLD on chromosome 18 between the Afrikaners and the four other populations reported in Eaves et al. 2000 (See Web Site References for location of raw data). We first evaluated all pairs of loci on chromosome 18 in the Afrikaner sample for the presence of allelic associations using Fisher exact test (FET). To describe the extent of nonrandom allelic association between pairs of

loci we computed the tail probability (FET p value) using the *Arlequin* software (Schneider et al. 2000). A total of 36 locus pairs were considered for chromosome 18 with an average distance of 3.1 cM. A large number of pairs (50%) with a p value <.01 were identified (Table 2): 4/7 pairs within 1 cM, 4/7 between 1 and 2 cM, 3/4 between 2 and 3 cM, 2/4 between 3 and 4 cM, 2/7 between 4 and 5 cM, and 3/7 between 5 and 6.5 cM. Multiple testing is an issue with regard to the two locus tests. However, when the conservative Bonferroni correction is applied for the 36 tests performed, 14/18 (78%) tests reported in Table 2 still remain significant at p < .01. Under this stringent criterion, 85% (6/7) of the tests for marker pairs separated by more than 3 cM remain significant, including two marker pairs at a distance >5 cM.

We also quantified the strength of bLD rather than the significance of the association between all possible pairs of loci. We used two different methods to quantify LD. First, we used the widely employed multiallelic extension of

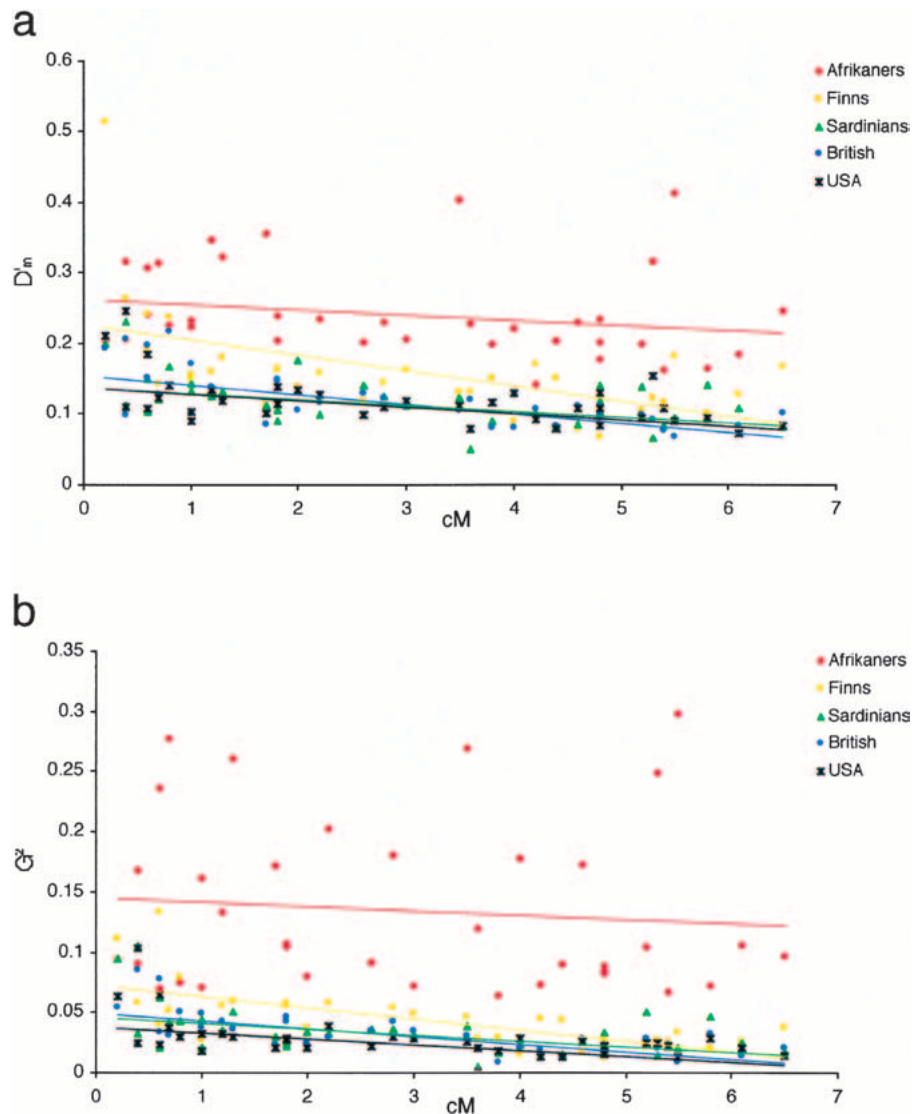


Figure 1 Relationship between multiallelic D'_m (a) and G^2 (b) values and genetic distance (cM) across populations. Solid lines are linear regression best fit lines.

Lewontin's standardized measure of disequilibrium (D' measure, Lewontin 1964; Hedrick 1987). This statistic may be affected by sample size (with smaller size samples demonstrating larger D'_m values) and allele frequencies (Mohlke et al. 2001) and therefore may not be appropriate for comparison of LD between studies that use different sample sizes and markers with differing numbers of alleles and differing allele frequencies. To facilitate comparison of bLD across populations, we also used the pairwise G^2 measure, originally proposed by Balakrishnan and Sangvi (1968) and then explored by Chapman and Wijsman (1998), as a measure of LD for loci with multiple alleles. G^2 values are not affected by changes in sample sizes.

Figure 1 reports the distribution of pairwise D'_m and G^2 values for all marker pairs tested for all five populations compared. We find that the extent of bLD in the Afrikaners, when either G^2 or D'_m values are considered across chromosome 18, is consistently higher than in the other populations tested. Over short intervals (0–3 cM), analysis of G^2 values distribu-

tion indicates high bLD in the Afrikaners and to a lesser extent in the Finns, with fewer differences among all the other populations. At this range, the Afrikaners show $G^2 > 0.05$ for 18/18 marker pairs compared with 12/18 in the Finnish sample, 4/18 in the Sardinian and British sample, and 3/18 in the USA sample (Fig. 1b, Table 3). When using the D'_m statistic, the Afrikaners show $D'_m > 0.30$ for 6/18 marker pairs compared with 1/18 in the Finns and 0/18 in all other populations tested (Fig. 1a, Table 3). Although the D'_m statistic may be affected by sample size (the Afrikaner sample is almost one third of all other demographic groups), the similar distribution pattern of the G^2 values strongly indicates that higher levels of bLD exist in the Afrikaner dataset in the entire 0–6.5-cM range at chromosome 18.

Because different pairs of loci in the genome will show variable amount of LD, we compared the distribution of bLD across two genomic regions within the Afrikaner sample by typing 10 additional markers on chromosome 4 (Table 1).

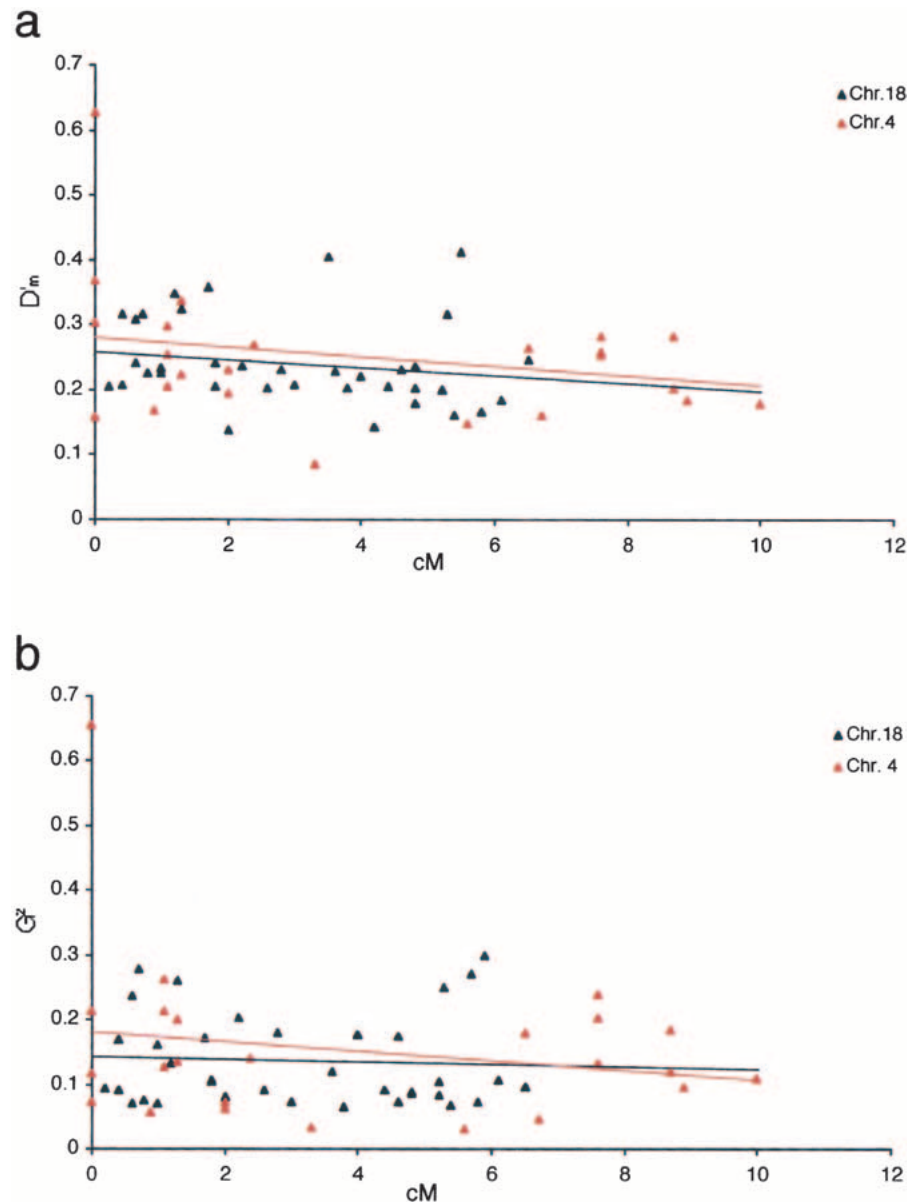


Figure 2 Comparison of the strength of linkage disequilibrium (LD) on chromosomes 18 and 4 in the Afrikaners. Multiallelic D'_m (a) and G^2 (b) are plotted against genetic distance (cM). Each point represents one pair of markers. Solid lines are linear regression best fit lines.

Analysis of the distribution of D'_m and G^2 values (Fig. 2) indicates that LD does not seem to vary significantly between the two genomic regions analyzed in this study. Marker pair comparisons within the common genetic range tested for both chromosomes (0–6.5 cM) indicate a similar proportion of pairs with $D'_m > 0.30$ or $G^2 > 0.05$, as well as similar D'_m and G^2 mean values (Table 4).

DISCUSSION

In the present study, we conducted an investigation of the level of bLD in the Afrikaner population to test the appropriateness of this genetic isolate for genome-wide linkage and linkage disequilibrium studies designed to identify genes for complex traits. We took care to avoid strategies previously

proven confusing in LD data interpretation. We used microsatellite markers as opposed to commonly used di-allelic single nucleotide polymorphisms (SNPs) for two reasons: (1) because of the large number of dense microsatellite marker panels already in use in genetic linkage studies and (2) because of the notion that di-allelic markers may have less power to detect LD than multiallelic markers (Chapman and Wijsman 1998). In addition, we did not collapse alleles to simulate di-allelic markers, as is often done to facilitate the statistical analysis but can give highly misleading results (Godard et al. 1999). Haplotypes were determined with precision using primarily family data rather than data inferred through statistical methods. Finally, alongside the most commonly used measures, such as D'_m (Lewontin 1964) and p values from pairwise significance tests, we also applied the G^2 statistic that fa-

Table 4. Comparison of LD between Two Genomic Regions in the Afrikaners

Chromosome	$D'_{m} > 0.30$ ratio (%)	D'_{m} average	$G^2 > 0.05$ ratio (%)	G^2 average
18	9/36 (25)	0.24	36/36 (100)	0.14
4	4/16 (25)	0.26	14/16 (87.5)	0.16

cilitated the comparison of data from different populations with different sample sizes (Chapman and Wijsman 1998).

One important finding of our study is the remarkable number of significant marker-to-marker associations at 3–6 cM of genetic distance observed in the Afrikaners (Table 2). This finding is in agreement with a previous, more limited study by Gordon et al. (2000) that found significant evidence for LD in the Afrikaners extending over a 5-cM region in other chromosomal regions. The validity of this finding is further supported in the current study by a direct bLD comparison of chromosome 18 between the Afrikaners and four other populations, including two genetic isolates and two mixed populations. This comparison indicates consistently higher mean LD levels for the Afrikaners at genetic distances over 3 cM when either the mean G^2 values or D'_{m} values are considered. The lower LD observed at this range in the Finns could be attributable to a larger number of generations of exponential growth compared with the Afrikaners who are characterized by a recent bottleneck event (13–15 generations) and have maintained the initial level of LD over larger genetic distances. Retention of bLD over large distances may also be related to a limited number of original founders for the families examined or to previously undetected admixture. Analysis of the population structure (Pritchard and Rosenberg 1999) and detailed genealogical research (M. Karayiorgou et al. in prep.) will help us understand the forces controlling bLD in this population.

Taken together, these results indicate that the Afrikaner population conforms to the standards desired in a population appropriate for a genome-wide scan requiring extensive LD. In comparison with other older founder populations with similar size (such as the Finns), significantly fewer markers may be sufficient for genome-wide LD analysis of complex disorders in the Afrikaners. This LD, coupled with the possibility of recruitment of large samples, could result in considerable power to detect LD and thus easier detection of genes with modest contribution to disease. In that respect, the Afrikaners, who total approximately 3 million, do not suffer from a common drawback encountered in founder populations with extended LD but small population size, namely, the availability of a limited number of cases even for common disorders (Zavattari et al. 2000). It is the combination of extended LD and availability of a large number of cases for ascertainment that makes the Afrikaners an ideal genetic isolate for mapping complex traits.

A common concern in using founder populations for mapping complex, multigenic traits is the relevance of the study's findings to other cosmopolitan populations. Although currently there is no empirical data arguing for or against this issue, the Northern European/Dutch origin of the Afrikaners is likely to render findings from studies in this population relevant to many other Western populations. It is of course likely that only a subset of susceptibility genes for any given complex disorder will segregate in a founder population. Nevertheless, identification of these genes is likely to provide im-

portant insight into the affected molecular and cellular pathways and inspire novel drug development approaches. Finally, an additional concern, applicable particularly in genetic mapping of phenotypes diagnosed as clinical syndromes rather than as quantitative traits, is the clinical similarity of a phenotype across populations. In that respect, we can mention that in our initial analyses comparing the clinical similarity between schizophrenia in the Afrikaner sample ($N = 159$) and our US sample ($N = 223$), we find that the basic sample descriptors and cardinal symptoms of the disease are equivalent (M. Karayiorgou et al. in prep.).

Our analysis used bLD estimates as an indication of disease-associated LD. It should be emphasized, however, that extended bLD may, in principle, confuse the interpretation of LD mapping unless it is accounted for (McPeck and Strahs 1999) and distinguished from disease-associated LD. It should be noted, however, that some of the limitations imposed by extended bLD can be overcome by following meiotic recombination events within linked families with multiple affected individuals, thus reducing the size of disease-associated haplotypes and facilitating fine mapping of the disease-causing gene. Alternatively, once LD is obtained and replicated, the fine mapping of the region could be pursued in the general Northern European/Dutch population of origin.

Finally, it is important to point out that the usefulness of founder populations in the dissection of complex traits also includes aspects other than increased bLD levels. These include a more uniform environment, good genealogical records, more intact families, and a phenotype definition that is easier to standardize.

METHODS

Sample

Of the 62 families used in the current study, 55 are three-member families composed of one child and both biological parents, and seven are four-member families composed of two children and both biological parents. All families studied are of Afrikaner heritage, as determined by our extensive research of genealogical records dating back to the early sixteenth century. DNA was extracted from blood samples from all study participants.

Our sampling approach is unlikely to make these families nonrepresentative of the general Afrikaner population for two reasons. First, separate analysis of transmitted and nontransmitted parental chromosomes provided identical results. Second, given the existing literature on genome-wide searches for schizophrenia susceptibility loci, any ascertainment bias introduced would have been minimal because the risk conferred by a putative schizophrenia susceptibility locus in the areas examined would be small.

The study was approved by Institutional Review Boards (IRBs) at both the Rockefeller University and University of Pretoria. Appropriate informed consent was obtained from all study participants.

Markers

All microsatellite markers we used are dinucleotide repeats except D18S851 that contains tetranucleotide repeats. Primer sequences are available from the Genome Database (GDB) with the exception of D18S1156, in which forward primer sequence was changed to CTTGCACCCTGCAAGTT, and D4S418, in which primer sequence was changed to GGATCA CAGAGTGAAG because of the presence of an undetermined nucleotide (N) in the primer sequence given at GDB. Observed heterozygosity values in the Afrikaner population range from 0.40 to 0.78 with an average value of 0.66.

Tests of Linkage Disequilibrium between Pairs of Markers

For each pair of loci, a contingency table of gamete frequencies was formed and 100,000 tables with the same marginal totals were generated on the basis of a Monte Carlo Markov-chain algorithm (Guo and Thompson 1992). The p value is the fraction of tables that were at least as extreme as the observed table.

D' is the standardized disequilibrium value that takes the usual disequilibrium coefficient $P(A_i B_j) - P(A_i) P(B_j)$ and divides it by its maximal possible value. Values range from 0 to 1, with 0 reflecting perfect independence between alleles at the two loci compared and 1 reflecting complete LD. We calculated a multiallelic extension of the normalized association measure D' (Lewontin 1964) as $D'_{ij} = \frac{P_{ij} - P_i P_j}{\sqrt{P_i(1-P_i)P_j(1-P_j)}}$ where p_i and q_j are the observed frequencies for alleles i and j , respectively, at the two loci. We computed the pairwise G^2 as $\sum_i (P_i - Q_i)^2 / (P_i + Q_i)$ where P_i is the estimated two-locus frequency for haplotype i and Q_i is the expected haplotype frequency based on independence of individual allele frequencies at the same two loci. Both statistics were calculated using haplotypes phased from family data. Haplotypes were manually constructed for the four parental chromosomes by inferring phase from their genotypes and those of the child. Before genotyping the 7 four-member families, one child was selected at random. Final haplotype frequencies were estimated using an EM algorithm.

ACKNOWLEDGMENTS

We thank all the study participants and Sister Ria van Wyk for assistance with the recruitment of the families. We also thank Gonçalo Abecasis for useful comments. Support for this work was provided by grants from the EJLB Foundation, the Essel Foundation, and NIH R01 MH61399-01 (to MK).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

- Balakrishnan, V. and Sanghvi, L.D. 1968. Distance between populations on the basis of attribute data. *Biometrics* **24**: 859–865.
- Brink, P.A., Steyn, L.T., Coetzee, G.A., and van der Westhuyzen, D.R. 1987. Familial hypercholesterolemia in South African Afrikaners. PvuII and StuI DNA polymorphisms in the LDL-receptor gene consistent with a predominating founder gene effect. *Hum. Genet.* **77**: 32–35.
- Brink, P.A., Ferreira, A., Moolman, J.C., Weymar, H.W., van der Merwe, P.L., and Corfield, V.A. 1995. Gene for progressive familial heart block type I maps to chromosome 19q13. *Circulation* **91**: 1633–1640.
- Chapman, N.H. and Wijsman, E.M. 1998. Genome screens using linkage disequilibrium tests: Optimal marker characteristic and feasibility. *Am. J. Hum. Genet.* **63**: 1871–1885.
- Eaves, I.A., Merriman, T.R., Barber, R.A., Nutland, S., Tuomilehto-Wolf, E., Tuomilehto, J., Cucca, F., and Todd, J.A. 2000. The genetically isolated populations of Finland and Sardinia may not be a panacea for linkage disequilibrium mapping of common disease genes. *Nat. Genet.* **25**: 320–323.
- Freimer, N.B., Service, S.K., and Slatkin, M. 1997. Expanding on population studies. *Nat. Genet.* **17**: 371–373.
- Goddard, K.A.B. 1999. 'Design issues in the analysis of complex genetic traits.' PhD diss., University of Washington, Seattle.
- Goldman, A., Krause, A., Ramsay, M., and Jenkins, T. 1996. Founder effect and the prevalence of myotonic dystrophy in South Africans: Molecular studies. *Am. J. Hum. Genet.* **59**: 445–452.
- Gordon, D., Simonic, I., and Ott, J. 2000. Significant evidence for linkage disequilibrium over a 5-cM region among Afrikaners. *Genomics* **66**: 87–92.
- Groenewald, J.Z., Liebenberg, J., Groenewald, I.M., and Warnich, L. 1998. Linkage disequilibrium analysis in a recently founded

- population: Evaluation of the variegated porphyria founder in South African Afrikaners. *Am. J. Hum. Genet.* **62**: 1254–1258.
- Guo, S.W. and Thompson, E.A. 1992. Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361–372.
- Hayden, M.R., MacGregor, J.M., and Beighton, P.H. 1980. The origin of Huntington's chorea in Afrikaner population. *S. Afr. Med. J.* **58**: 197–200.
- Hedrick, P.W. 1987. Gametic disequilibrium measures: Proceed with caution. *Genetics* **117**: 331–341.
- Jorde, L.B. 1995. Linkage disequilibrium as a gene-mapping tool. *Am. J. Hum. Genet.* **56**: 11–14.
- Leitersdorf, E., van der Westhuyzen, D.R., Coetzee, G.A., and Hobbs, H.H. 1989. Two common low density lipoprotein receptor gene mutations cause familial hypercholesterolemia in Afrikaners. *J. Clin. Invest.* **84**: 954–961.
- Lewontin, R.C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics* **49**: 49–67.
- McPeck, M.S. and Strahs, A. 1999. Assessment of linkage disequilibrium by the decay of haplotype sharing, with application to fine-scale genetic mapping. *Am. J. Hum. Genet.* **65**: 858–875.
- Mohlke, K.L., Lange, E.M., Valle, T.T., Ghosh, S., Magnuson, V.L., Silander, K., Watanabe, R.M., Chines, P.S., Bergman, R.N., Tuomilehto, J., et al. 2001. Linkage disequilibrium between microsatellite markers extends beyond 1 cM on chromosome 20 in Finns. *Genome Res.* **11**: 1221–1226.
- Pritchard, J.K. and Rosenberg, N.A. 1999. Use of unlinked genetic markers to detect population stratification in association studies. *Am. J. Hum. Genet.* **65**: 220–228.
- Pronk, J.C., Gibson, R.A., Savoia, A., Wijker, M., Morgan, N.V., Melchionda, S., Ford, D., Temtamy, S., Ortega, J.J., Jansen, S., et al. 1995. Localisation of the Fanconi anaemia complementation group A gene to chromosome 16q24.3. *Nat. Genet.* **11**: 338–340.
- Risch, N. and Merikangas, K. 1996. The future of genetic studies of complex human disease. *Science* **273**: 1516–1517.
- Roby, P., Eyre, S., Worthington, J., Ramesar, R., Cilliers, H., Beighton, P., Grant, M. and Wallis, G. 1999. Autosomal dominant (Beukes) premature degenerative osteoarthritis of the hip joint maps to an 11-cM region on chromosome 4q35. *Am. J. Hum. Genet.* **64**: 904–908.
- Rosendorff, J., Bernestein, R., Macdougall, L. and Jenkins, T. 1987. Fanconi anaemia: Another disease of high prevalence in the Afrikaans population of South Africa. *Am. J. Med. Genet.* **27**: 793–797.
- Schneider, S., Roessli, D., and Excoffier, L. 2000. Arlequin: A software for population genetic data analysis. User manual ver 2.000. Ver. 2.000. Geneva: Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva, Geneva.
- Shifman, S. and Darvasi, A. 2001. The value of isolated populations. *Nat. Genet.* **28**: 309–310.
- Torrington, M. and Viljoen, D.L. 1991. Founder effect in Afrikaners kindreds with pseudoxanthoma elasticum. *S. Afr. Med. J.* **79**: 7–11.
- Warnich, L., Kotze, M.J., Groenewald, I.M., Groenewald, J.Z., van Brakel, M.G., van Heerden, C.J., de Villiers, J.N., van de Ven, W.J., Schoenmakers, E.F., Taketani, S. et al. 1996. Identification of three mutations and associated haplotypes in the protoporphyrinogen oxidase gene in South Africa. *Hum. Mol. Genet.* **5**: 981–984.
- Zavattari, P., Deidda, E., Whalen, M., Lampis, R., Mulaglia, A., Loddo, M., Eaves, I., Mastio, G., Todd, J.A., and Cucca, F. 2000. Major factors influencing linkage disequilibrium by analysis of different chromosome regions in distinct populations: Demography, chromosome recombination frequency and selection. *Hum. Mol. Genet.* **9**: 2947–2957.

WEB SITE REFERENCES

- <http://anthropologie.unige.ch/arlequin/>; Arlequin program.
- <http://research.marshfieldclinic.org/genetics/>; Haplotypes from Eaves et al. 2000.
- <http://genome.ucsc.edu/>; Human Genome Project Working Draft, Santa Cruz.
- <http://www.gdb.org/>; Genome Database (GDB).
- http://www-gene.cimr.cam.ac.uk/todd/public_data/chr18/LDwebpage.html; Marshfield genetic maps.

Received January 29, 2002 ; accepted in revised form March 29, 2002.