Am. J. Hum. Genet. 59:377–384, 1996

A Locus for Fanconi Anemia on 16q Determined by Homozygosity Mapping

Michele Gschwend,[1] Orna Levran,[2] Leonid Kruglyak,[3] Koustubh Ranade,[1] Peter C. Verlander,[2] Susan Shen,[1] Sabine Faure,[5] Jean Weissenbach,[5] Cigdem Altay,[6] Eric S. Lander,[3,4] Arleen D. Auerbach,[2] and David Botstein[1]

[1]Department of Genetics, Stanford University, Stanford; [2]Laboratory of Human Genetics and Hematology, Rockefeller University, New York; [3]Whitehead Institute for Biomedical Research and [4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA; [5]Généthon, Evry; and [6]Faculty of Medicine, Hacettepe University, Ankara

Summary

We report the results of a genomewide scan using homozygosity mapping to identify genes causing Fanconi anemia, a genetically heterogeneous recessive disorder. By studying 23 inbred families, we detected linkage to a locus causing Fanconi anemia near marker D16S520 (16q24.3). Although ~65% of our families displayed clear linkage to D16S520, we found strong evidence (P = .0013) of genetic heterogeneity. This result independently confirms the recent mapping of the *FAA* gene to chromosome 16 by Pronk et al. Family ascertainment was biased against a previously identified *FAC* gene on chromosome 9, and no linkage was observed to this locus. Simultaneous search analysis suggested several additional chromosomal regions that could account for a small fraction of Fanconi anemia in our families, but the sample size is insufficient to provide statistical significance. We also demonstrate the strong effect of marker allele frequencies on LOD scores obtained in homozygosity mapping and discuss ways to avoid false positives arising from this effect.

Introduction

Homozygosity mapping is a powerful method that uses affected, inbred individuals to map a rare, recessive disease (Smith 1953; Lander and Botstein 1987). In the past several years, researchers (Hamida et al. 1993; Pollak et al. 1993; German et al. 1994) have successfully applied this method to locate genes that cause monogenic Mendelian diseases. The applicability of homozygosity mapping has been enhanced by the recent progress in the number of genetic markers available (Weissenbach et al. 1992; Gyapay et al. 1994) and the development of computer programs, such as MAPMAKER/HOMOZ (Kruglyak et al. 1995), that efficiently calculate multipoint LOD scores in inbred pedigrees.

In theory, homozygosity mapping can also be applied to genetically heterogeneous diseases, but the sample size must be increased (Lander and Botstein 1987) and different analytical tools must be applied. However, this approach has not yet been applied in practice. To allow for genetic heterogeneity, heterogeneity LOD scores (hLODs) can be calculated by incorporating in the genetic model a parameter α, which represents the fraction of linked families. The heterogeneity parameter α does not need to be specified in advance but can be determined by maximum-likelihood estimation (Terwilliger and Ott 1994). In fact, it is possible to extend this approach to allow for linkage to several specified loci, by estimating the fractions, α and β, of families in which the disease gene is linked to each of two different loci. This method is called "simultaneous search" (Lander and Botstein 1986) and, in principle, may facilitate the detection of linkage in a genetically heterogeneous disease.

The recessive disease Fanconi anemia (FA) is a good candidate for the application of homozygosity mapping and simultaneous search to a genetically heterogeneous disease. The disease is rare and known to be genetically heterogeneous. FA is characterized by progressive bone marrow failure and other, more variable clinical manifestations (Schroeder et al. 1976; Gordon-Smith and Rutherford 1989; Alter 1992). FA cells are uniquely hypersensitive to bifunctional DNA crosslinking agents such as diepoxybutane and mitomycin C (Cervenka et al. 1981; Auerbach 1993; Giampietro et al. 1993); complementation of this cellular phenotype has been used to demonstrate genetic heterogeneity. Complementation tests between cells from different FA patients have shown the presence of at least five complementation groups, most simply interpreted as five FA-causing genes (Joenje et al. 1995; Strathdee et al. 1992*b*). FA complementation group A has been shown to be the most com-

Received December 18, 1995; accepted for publication May 9, 1996.

Address for correspondence and reprints: Dr. David Botstein, Department of Genetics, Stanford University School of Medicine, Stanford, CA 94305-5120. E-mail: botstein@genome.stanford.edu

0002-9297/96/5902-0013$02.00

mon gene, representing 67% of FA cells (Buchwald 1995). One of the genes causing this disease, *FAC* (for FA complementation group C), was cloned by functional complementation, proved to be causative by mutation analysis, and mapped to chromosome 9q22.3 (Strathdee et al. 1992*a*, 1992*b*). Whitney et al. (1995) have mapped the *FAD* gene to chromosome 3p by functional complementation using microcell-mediated chromosome transfer. Independent of this study, another FA gene was recently localized to chromosome 16q24.3 by linkage analysis in several families assigned to complementation group A (Pronk et al. 1995).

In this paper, we present the result of a genomewide screen for genes causing FA in a panel of consanguineous pedigrees. We independently detected the presence of a gene causing FA on chromosome 16q24.3 in the majority of families analyzed. A simultaneous search analysis for additional FA-causing loci was also performed and indicated several other potential loci. However, the relatively small number of families linked to these other loci precludes a conclusion of significant evidence for or against linkage.

Subjects and Methods

Consanguineous Families and Markers

DNA from a total of 26 pedigrees segregating FA was obtained from the International Fanconi Anemia Registry at Rockefeller University (table 1). Not all families were available for genotyping at the outset of the project. Approximately half of the pedigrees were ascertained in Turkey by one of the authors (C.A.), even though this country does not have an increased incidence of FA. Individuals were considered affected by the presence of congenital malformations or hematological abnormalities typical of FA (Auerbach et al. 1989) and by the hypersensitivity of their cells to diepoxybutane (Auerbach 1993). Thirty-three of 39 affected individuals used in this study were not found to have mutations in *FAC*, by SSCP analysis (Verlander et al. 1994). Seven of 39 affected individuals were not screened for mutations in *FAC*.

A two-stage genomewide screen with different marker resolutions was performed (table 2). An initial low-resolution (15-cM) screen consisted of affected individuals and their parents from 14 inbred pedigrees (table 1). An additional family of two affected siblings with parents of unknown degree of relatedness were also included. These 48 DNAs were genotyped with 347 $(CA)_n$ repeat markers (Weissenbach et al. 1992), 278 (80%) of which produced genotype information. For the high-resolution (5 cM) screen, additional pedigrees had become available, and we were able to expand the number of inbred families to 23 (table 1). Pedigrees 3 and 19 were not included in the high-resolution screen, because of limited DNA resources. An additional 432 $(CA)_n$ repeat markers (Gyapay et al. 1994), of which 388 (90%) produced genotype information, were added to the previous set of markers to produce a genomewide screen with an average spacing of 5 cM between markers. For the combined genomewide screen it was not possible to score 113 (15%) of 779 of the markers, because of the presence of nonspecific products or poor amplification. All primer pairs were purchased from Research Genetics.

Determination of Genotypes

Genotypes were determined as described elsewhere (Clark and Gschwend 1995). In brief, single markers were amplified by PCR in 10-μl reactions composed of 25 ng DNA, 5 pmol each primer, 0.2 mM dNTPs, 1 × *Taq* buffer (Perkin Elmer), and 0.25 U of *Taq* polymerase (Perkin Elmer). PCR was performed with an initial denaturation for 5 min at 94°C; followed by 30 cycles at 94°C for 10 s, 55°C for 30 s, and 72°C for 30 s; and a final extension of 72°C for 5 min. After pooling and concentrating the reactions, 24 multiplexed PCR products plus an internal size standard were separated in a 5% polyacrylamide sequencing gel. The separated DNA was electroblotted onto a membrane, hybridized with a digoxigenin-labeled primer, and detected by chemiluminescence. Marker films were translated into digital image files using Whole Band Analyzer software (Bio Image). Genotypes were scored independently by two individuals; discordant calls were not used in the analysis. For comparison of allele sizes between gels, sizes were normalized to a CEPH individual (NA10859, Coriell Cell Repositories) with known genotype (Gyapay et al. 1994).

Analysis of Genotypes

Multipoint LOD scores under a hypothesis of genetic homogeneity (LOD) or genetic heterogeneity (hLOD) were computed by MAPMAKER/HOMOZ (version 0.9) (Kruglyak et al. 1995). For the 15-cM screen, allele frequencies for each marker were obtained online from the Genome Data Base (GDB; http://gdbwww.gdb.org/www). If an observed allele was not present in the database, an allele frequency of {1/number of chromosomes assayed} was used. For the 5-cM genomic screen, allele frequencies were calculated internally by examining the genotypes of the affected children of each inbred pedigree. Each allele was counted as it occurred; that is, for a homozygous individual the allele would be counted twice. Calculating allele frequencies in this manner is conservative, because it leads to inflation of allele frequencies, which reduces the LOD score in inbred progeny. Distances between markers were obtained from published maps (Weissenbach et al. 1992; Gyapay et al. 1994). The disease allele frequency was set at .004 with 100% penetrance (Swift 1971; Rogatko and Auerbach

Table 1

Description of Family Sample

Pedigree	Degree of Relationship	Country of Ancestry	No. of Affected[a]	No. of Parents[a]	No. of Siblings[b]	Genome Screen[c]
1	Second cousin	Palestine	1	2	4	Low, high
2	First cousin	Portugal	1	2	0	Low, high
3	Unknown	Dominican Republic	2	2	2	Low
4	First cousin	Italy	1	2	2	Low, high
5	First cousin	Sicily	3	1	1	Low, high
6	First cousin	Yemen	1	2	5	Low, high
7	First cousin	Mexico	1	2	0	Low, high
8	First cousin	Pakistan	2	2	1	Low, high
9	First cousin	Turkey	1	2	2	Low, high
10	First cousin	Turkey	1	2	1	Low, high
11	First cousin, once removed	Turkey	1	2	1	Low, high
12	Second cousin	Turkey	1	2	1	Low, high
13	First cousin	Turkey	1	2	1	Low, high
14	First cousin	Turkey	1	2	3	Low, high
15	First cousin	Turkey	1	2	0	Low, high
16	Second cousin	Ireland	1	2	0	High
17	Third cousin	Brazil	4	2	1	High
18	First cousin	Brazil	3	2	2	High
19a[d]	Second cousin	Saudi Arabia	3	2	1	Chromosome 11
19b[d]	First cousin	Saudi Arabia	2	2	5	Chromosome 11
20	First cousin, once removed	England	2	2	0	High
21	First cousin	Turkey	1	1	0	High
22	First cousin	Turkey	1	2	0	High
23	Second cousin	Turkey	1	1	0	High
24	First cousin, once removed	Turkey	1	2	0	High
25	First cousin	Turkey	1	2	0	High

[a] No. of individuals available for genotyping.

[b] No. of unaffected siblings available for genotyping from each pedigree. As described in Results, these individuals were genotyped only for chromosome 11 markers.

[c] Describes the participation of each pedigree in each of two genomewide screens. Several pedigrees were not available for the low-resolution screen. Because of limited resources, individuals from pedigree 19 were genotyped only for chromosome 11 markers.

[d] Pedigree 19 consisted of two inbreeding loops.

1988). Simultaneous search analysis was performed using a program written by L.K. and E.S.L. (unpublished data).

Radiation Hybrid (RH) Mapping

Sixteen $(CA)_n$ repeat markers linked to D20S109 were selected for physical mapping using RHs (Cox et al. 1990). Four markers could not be mapped, because of amplification of similarly sized rodent bands. Two other markers (D20S857 and D20S183) showed a discrepancy in order between the genetic and physical maps. Genotypes from these two markers were not used in the linkage analysis. The Stanford G3 panel of 83 RH DNA samples was purchased from Research Genetics. One hundred nanograms of each RH, including positive (human DNA) and negative (rodent DNA) controls, were amplified in duplicate with 10 pmol of each forward and reverse primer, 2.5 mM dNTPs, 1 × Perkin-Elmer *Taq* buffer, and 0.5 U of *Taq* polymerase in a 20-μl reaction. The PCR was done as described above. PCR products were detected by agarose gel electrophoresis. The RH data was analyzed using RHMAP (Boehnke et al. 1991).

Results

Low-Resolution Genomic Screen

We began by performing a low-resolution (15-cM) genomewide screen in which 256 markers were screened on 19 affected individuals and 29 parents from 15 pedigrees. Under a model of genetic homogeneity, several chromosomes (3, 11, 14, and 20) produced LOD scores >1.00, but no chromosomal region yielded a significant LOD score (>3.00). To account for the known genetic

Table 2

Number of Markers per Chromosome in the Genomewide Screen

Chromosome	Low-Resolution Screen		High-Resolution Screen		Total No. of Markers Scored
	No. of Markers	Number scored[a]	No. of Markers Added[b]	Number Scored[a]	
1	32	28	37	34	62
2	20	13	45	43	56
3	17	14	34	33	47
4	20	17	27	25	42
5	14	13	27	22	35
6	12	10	27	24	34
7	16	13	26	25	38
8	18	16	17	16	32
9	11	10	19	18	28
10	12	10	20	17	27
11	67	56	4	4	60
12	8	6	24	20	26
13	7	5	18	13	18
14	10	6	17	12	18
15	6	4	12	10	14
16	9	6	15	14	20
17	10	9	15	15	24
18	7	5	14	13	18
19	6	2	14	11	13
20	37	31	9	9	40
21	4	3	4	3	6
22	4	1	7	7	8
Overall	347	278	432	388	666

[a] No. of markers scored per chromosome. Markers were not scored when nonspecific products or poor amplification was observed.

[b] No. of markers added to the low-resolution map to fill in gaps >5 cM.

heterogeneity of FA, LOD scores were then calculated under a model of heterogeneity (hLOD). In this sample set, which consisted of mostly first-cousin offspring, an hLOD of ~3.70 is required for statistical significance (Lander and Kruglyak 1995). Two regions of the genome yielded hLODs exceeding this threshold (table 3).

Table 3

hLOD Scores Obtained from the Low-Resolution (15 cM) Genomic Screen and the Effect of Changes in Allele Frequency

	hLOD at D11S1343	hLOD at D20S109
GDB frequencies[a]	4.01	4.59
.05 minimum[b]	3.92	2.72
.07 minimum[c]	3.81	2.15

[a] Allele frequencies were obtained online from GDB (http://gdbwww.gdb.org/www). A minimum allele frequency of .02 was used.

[b] All allele frequencies <.05 were increased to .05. Allele frequencies >.05 were not changed.

[c] All allele frequencies <.07 were increased to .07. Allele frequencies >.07 were not changed.

An hLOD of 4.01 was observed on chromosome 11 at D11S1343 ($\theta = 0$; $\alpha = 0.58$), and an hLOD of 4.59 was observed on chromosome 20 at D20S109 ($\theta = 0$; $\alpha = 0.52$). (The parameter α represents the fraction of pedigrees linked to a given locus.)

The results of a homozygosity mapping analyses may be sensitive to allele frequencies. Because we used marker allele frequencies taken from the GDB, we carried out sensitivity analysis (Kruglyak et al. 1995) by increasing the minimum allele frequency. The hLOD peak near D11S1343 on chromosome 11 was relatively insensitive to changes in allele frequencies (table 3). However, the hLOD peak near D20S109 fell from 4.59 to below significance levels when allele frequencies were somewhat altered (table 3).

Further Examination of Peaks on Chromosomes 20 and 11

The LOD score results on chromosome 20 were extremely sensitive to changes in allele frequency, indicating that this apparent evidence of linkage was likely to be a false positive. To test further the possibility of link-

age to chromosome 20, we increased the marker density from 9 cM to 2 cM by typing additional genetic markers linked to D20S109. Because the map position and relative order of these markers was not known with precision, we physically mapped the markers using RH mapping (data not shown). The resulting high-density genetic map was used to reevaluate the support of linkage on chromosome 20. The maximum hLOD dropped from 4.59 (D20S109; $\theta = 0$; $\alpha = 0.52$) to -0.015 (D20S109; $\theta = 0$; $\alpha < 0.03$), as the previous genetic markers showing homozygosity were now seen to be interspersed with markers showing heterozygosity. This result indicates that the initial linkage on chromosome 20 was a false positive—probably due to inaccurate allele frequencies for the initial markers. Specifically, the alleles showing homozygosity were likely to be more prevalent in our population than the frequencies estimated in GDB.

Although the LOD score result on chromosome 11 at D11S1343 was insensitive to changes in allele frequencies, we examined the area further by augmenting the sample with newly available pedigrees and by typing further markers. The addition of 11 consanguineous pedigrees (20 affected and 20 parental genotypes) slightly reduced the hLOD at D11S1343 from 4.01 (original 15 pedigrees, $\alpha = 0.58$) to 3.83 (26 pedigrees, $\alpha = 0.42$). Furthermore, when genotypes from the 33 unaffected siblings were added to the analysis of 26 pedigrees, the hLOD at D11S1343 dropped from 3.83 ($\alpha = 0.42$) to 2.74 ($\alpha = 0.31$). The loss of support for linkage was due to the sharing of an identical haplotype between the proband and an unaffected sibling in two pedigrees (data not shown).

The peak on chromosome 11 was further examined by the addition of eight markers surrounding D11S1343 (Vanagaite et al. 1995). These markers were genotyped on DNA of 12 of 26 families showing positive LOD scores at D11S1343. The resulting hLOD for all 26 pedigrees at D11S1343 dropped from 3.83 ($\alpha = 0.42$) to 2.99 ($\alpha = 0.34$). When both unaffected siblings and additional marker genotypes are analyzed together, the hLOD at D11S1343 falls to 2.273 ($\alpha = 0.31$). In summary, addition of further families and further markers caused the hLOD score on chromosome 11 to drop from 4.01 to 2.27—a level that was not statistically significant.

High-Resolution Genomic Screen

Because we did not detect statistically significant linkage to any locus, we increased the resolution of our genomewide scan. We genotyped 32 affected inbred individuals for an additional 432 markers, comprising a 5-cM genomic screen. As the use of allele frequencies from GDB led to a false positive on chromosome 20 in the low-resolution screen, allele frequencies for the high-resolution screen were determined from within the sample. We adopted a very conservative approach, calculating allele frequencies for all markers by counting the alleles present in the genotype of each inbred, affected individual.

A single significant peak with a maximum hLOD score of 6.08 ($\alpha = 0.66$) at marker D16S520 was observed. This hLOD score is well above the threshold for statistical significance and indicates an FA-causing gene near the telomere of chromosome 16. When genetic homogeneity is assumed, the LOD score at D16S520 drops to 3.823, indicating significant genetic heterogeneity ($P = .0013$) at this locus. D16S520 maps 14 cM beyond the most distal marker that had been used in our low-resolution screen (D16S422). In retrospect, it was not surprising that our low-resolution screen failed to identify this region, which was 14 cM beyond the bounds of the initial search. Simulations using the observed allele frequencies showed that a disease locus in two-thirds of pedigrees occurring in this location would usually result in a nonsignificant hLOD score. Indeed, there would be only a 50% chance of finding an hLOD >1.92.

Since heterogeneity at D16S520 was observed in our analysis, we examined whether there was detectable linkage to the *FAC* gene on chromosome 9. Of the 7 of 23 pedigrees that had not been screened for *FAC* mutations by SSCP, only one pedigree had a slightly positive LOD score (0.1058) at D9S176, a marker that is tightly linked to *FAC* (Gibson et al. 1994). All other pedigrees had negative multipoint LOD scores at D9S176 (range: $-\infty$ to -0.0686). Thus, the heterogeneity observed at marker D16S520 is not explained by linkage to the previously identified *FAC* gene on chromosome 9. Examining the rest of the genome, we noted nine other chromosomal regions (on chromosomes 2, 4, 7, 10, 11, 18, and 19) with hLODs between 1.00 and 2.00. The list includes a peak at D11S1343 (hLOD = 1.43; $\alpha = 0.22$), which was also seen in the low-resolution screen.

To determine whether linkage to loci on chromosomes other than chromosome 16 could be detected within our data set, we performed a conditional search analysis (Dupuis et al. 1995) which is an adaptation of simultaneous search. By fixing a disease locus at D16S520, we sought to identify additional loci in the genome that accounted for FA in families unlinked to chromosome 16. In a conditional search with initial locus L_1 (in the present case, D16S520), one scans the genome to identify a second locus L_2 such that the two loci *together* explain the trait much better than locus L_1 alone. The strength of the evidence is measured by comparing the hLODs for the hypothesis of disease genes at L_1 and L_2 accounting for proportions α and β of families and the hypothesis of a disease gene at L_1 accounting for proportion α' of families; the hLOD

score must increase by ≥3.70 to provide statistically significant evidence for the role of locus L_2. There were six chromosomal regions (on chromosomes 2, 4, 9, 11, 18, and 19) for which the hLOD increased by 1.0–2.0, but no regions yielded an increase of 3.70 (table 4). These regions are thus potential sites to study for an additional FA gene in a larger family collection, but there is not sufficient statistical evidence to implicate them.

As an additional method to identify a locus other than D16S520, we also performed a separate hLOD score analysis of eight pedigrees in which the affected individuals were not homozygous at D16S520. Several markers showed hLODs >1.00, but no statistically significant hLOD was observed (table 5). The highest hLOD observed in these eight pedigrees was 2.154 at D18S56 ($\alpha = 0.83$).

Finally we performed a simultaneous search analysis on this data subset. In such an analysis, one compares the hypothesis that two loci L_1 and L_2 account for proportions α and β of families to the hypothesis that *neither* locus harbors a disease gene. Because the risk of false positives is greater (inasmuch as one is searching over all *pairs* of loci rather than a single locus), the threshold for significance must be increased. A simulation of these pedigrees showed that a simultaneous search LOD score of 6.50 results in a genomewide false-positive rate of <5% (David Siegmund, personal communication). Thus, a simultaneous search LOD score >6.50 for a pair of loci makes it unlikely that both loci are false positives but does not indicate whether one of the two loci might be a false positive. The simultaneous search analysis did not produce a statistically significant results. The strongest combined LOD score observed in these data (eight pedigrees) was 4.11 (D9S162, $\alpha = 0.36$; D18S56, $\beta = 0.63$).

Table 4

Results of the Conditional Search Analysis with Locus 1 (D16S520) Fixed

LOD Score	α[a]	Locus 2[b]	β[a]	ΔLOD[c]
7.65	.58	D2S206	.32	1.57
7.53	.63	D4S413	.21	1.45
7.79	.67	D9S162	.18	1.71
7.50	.66	D11S938[d]	.16	1.42
7.83	.63	D18S56	.32	1.75
7.36	.65	D19S405	.24	1.28

[a] The fraction of pedigrees (α, β) linked to locus 1 (D16S520) or locus 2, respectively.

[b] The marker(s) occurring at the LOD score maximum is indicated.

[c] The increase in LOD score compared to D16S520 alone (hLOD = 6.08).

[d] D11S938 maps 6 cM distal to D11S1343.

Table 5

hLOD Scores >1.00 Observed in an Analysis of Eight Pedigrees That Were Not Homozygous at D16S520

Marker Name[a]	hLOD	Fraction of Pedigrees Linked
D2S206	1.135	.62
D4S413	1.700	.63
D7S517	1.034	.42
D7S521, D7S667	1.020	.46
D9S162, D9S171	1.605	.42
D18S56	2.154	.83

[a] The marker(s) occurring at the LOD score maximum is indicated.

Discussion

By applying the method of homozygosity mapping, we identified a gene causing FA in two-thirds of 23 consanguineous pedigrees analyzed. Our results are entirely consistent with the recent mapping of the *FAA* gene to chromosome 16q by Pronk et al (1995). At present, candidate genes are lacking in this area.

The data show significant genetic heterogeneity at D16S520. This is not the result of linkage to *FAC;* our family set was deliberately chosen among families that did not have mutations in the *FAC* gene by SSCP (Verlander et al. 1994). Several potential regions that might harbor FA loci were suggested by our analysis (table 4). However the LOD scores at these positions, although showing some evidence of linkage, do not reach statistical significance. Simulation studies using this data set confirmed that we would not be able to detect significant linkage when only 25% of our samples are linked to a given locus. To determine whether one or any of these suggested loci are involved in FA, additional markers and pedigrees will need to be analyzed. Furthermore, complementation analysis cell culture may facilitate FA gene mapping, especially for rare complementation groups.

In this study we have successfully applied homozygosity mapping to a rare recessive disease in the face of heterogeneity. Several practical features of this method deserve mention. First, use of externally derived allele frequencies for homozygosity mapping can lead to spurious positives, as seen in our low-resolution genomewide screen. Fortunately, it is possible to test borderline positive results by genotyping affected individuals with a dense genetic map. False-positive results due to inaccurate allele frequencies will typically be revealed by the presence of heterozygosity at these additional loci. We recommend such a test for all borderline results obtained in a homozygosity mapping screen. In addition, when a data set contains enough families, it is possible to employ a conservative approximation by estimating allele frequencies from within the family set. As seen here,

such a conservative approach was not detrimental to identify linkage.

In summary, homozygosity mapping can be successfully applied to a heterogeneous disease as long as care is taken to follow up potential positives with additional markers and to use a sufficiently dense genetic map. The application of a simultaneous search analysis was not decisive, since our data set was rather small after the majority of the pedigrees were linked to 16q. Nonetheless, this analysis suggested several regions not identified by simple examination of the hLOD scores. It remains to be seen whether any of these potential regions contain another FA gene.

Acknowledgments

We thank David Siegmund for helpful discussions concerning the threshold for simultaneous search analysis, David Cox for analysis of our RH data, and Carl Rosenberg and Mark Daly for software assistance. We gratefully acknowledge the contribution of the many physicians who referred patients to the International Fanconi Anemia Registry at The Rockefeller University. Research was supported by a grant to D.B. from The Charles A. Dana Foundation, to A. D. Auerbach from the National Institutes of Health (HL 32987) and the Fanconi Anemia Research Fund, Inc., to E.S.L. from the National Institutes of Health (HG00098), and to L.K. from the National Center for Human Genome Research (HG00017). M.G. was supported by training grant GM07790 from the National Institutes of Health.

References

Alter BP (1992) Fanconi's anemia: current concepts. Am J Pediatr Hematol Oncol 14:170–176

Auerbach AD (1993) Fanconi anemia diagnosis and the diepoxybutane (DEB) test. Exp Hematol 21:731–733

Auerbach AD, Rogatko A, Schroeder-Kurth T (1989) International Fanconi anemia registry: relation of clinical symptoms to diepoxybutane sensitivity. Blood 73:391–396

Boehnke M, Lange K, Cox DR (1991) Statistical methods for multipoint radiation hybrid mapping. Am J Hum Genet 49: 1174–1188

Buchwald M (1995) Complementation groups: one or more per gene? Nat Genet 11:228–230

Cervenka J, Arthur D, Yasis C (1981) Mitomycin C test for diagnostic differentiation of idiopathic aplastic anemia and Fanconi anemia. Pediatrics 67:119–127

Clark C, Gschwend M (1995) Nonradioactive multiplex analysis of SSLP's. In: Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, et al (eds) Current protocols in human genetics. Vol 1. John Wiley & Sons, New York, pp 2.5.7–2.5.20

Cox DR, Burmeister M, Price ER, Kim S, Myers RM (1990) Radiation hybrid mapping: a somatic genetic method for constructing high-resolution maps of mammalian chromosomes. Science 250:245–250

Dupuis J, Brown PO, Siegmund D (1995) Statistical methods for linkage analysis of complex traits from high-resolution maps of identity by descent. Genetics 140:843–856

German J, Roe AM, Leppert MF, Ellis NA (1994) Bloom syndrome: an analysis of consanguineous families assigns the locus mutated to chromosome band 15q26.1. Proc Natl Acad Sci USA 91:6669–6673

Giampietro PF, Adler-Brecher B, Verlander PC, Pavlakis SG, David JG, Auerbach AD (1993) The need for more accurate and timely diagnosis in Fanconi anemia: a report from the International Fanconi Anemia Registry. Pediatrics 91:1116–1120

Gibson RA, Ford D, Jansen S, Savoia A, Havenga C, Milner RD, de Ravel TJ, et al (1994) Genetic mapping of the FACC gene and linkage analysis in Fanconi anaemia families. J Med Genet 31:868–71

Gordon-Smith EC, Rutherford TR (1989) Fanconi anaemia-constitutional, familial aplastic anaemia. Baillieres Clin Haematol 2:139–152

Gyapay G, Morissette J, Vignal A, Dib C, Fizames C, Millasseau P, Marc S, et al (1994) The 1993–94 Généthon human genetic linkage map. Nat Genet 7:246–339

Hamida BC, Doerflinger N, Belal S, Linder C, Reutenauer L, Dib C, Gyapay G, et al (1993) Localization of Friedreich ataxia phenotype with selective vitamin E deficiency to chromosome 8q by homozygosity mapping. Nat Genet 5:195–200

Joenje H, Lo ten Foe JR, Oostra AB, van Berkel CG, Rooimans MA, Schroeder-Kurth T, Wegner RD, et al (1995) Classification of Fanoni anaemia patients by complementation analysis: evidence for a fifth genetic subtype. Blood 86: 2156–2160

Kruglyak L, Daly MJ, Lander ES (1995) Rapid multipoint linkage analysis of recessive traits in nuclear families, including homozygosity mapping. Am J Hum Genet 56:519–527

Lander ES, Botstein D (1986) Mapping complex denetic traits in humans: new methods using a complete RFLP linkage map. Cold Spring Harbor Symp Quant Biol 51:49–62

——— (1987) Homozygosity mapping: a way to map human recessive traits with the DNA of inbred children. Science 236:1567–1570

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

Pollak MR, Chou YH, Cerda JJ, Steinmann B, La DB, Seidman JG, Seidman CE (1993) Homozygosity mapping of the gene for alkaptonuria to chromosome 3q2. Nat Genet 5:201–204

Pronk J, Gibson R, Savoia A, Wijker M, Morgan N, Melchionda S, Ford D, et al (1995) Localisation of the Fanconi anaemia complementation group A gene to chromosome 16q24.3. Nat Genet 11:338–340

Rogatko A, Auerbach AD (1988) Segregation analysis with uncertain ascertainment: application to Fanconi anemia. Am J Hum Genet 42:889–897

Schroeder TM, Tilgen D, Krüger J, Vogel F (1976) Formal genetics of Fanconi's anemia. Hum Genet 32:257–288

Smith CAB (1953) The detection of linkage in human genetics. J R Stat Soc B 15:153–192

Strathdee CA, Duncan AMV, Buchwald M (1992*a*) Evidence

for at least four Fanconi anaemia genes including FACC on chromosome 9. Nat Genet 1:196–198

Strathdee CA, Gavish H, Shannon WR, Buchwald M (1992*b*) Cloning of cDNAs for Fanconi's anaemia by functional complementation. Nature 356:763–767

Swift M (1971) Fanconi's anaemia in the genetics of neoplasia. Nature 230:370–373

Terwilliger JD, Ott J (1994) Handbook of human linkage analysis. The Johns Hopkins University Press, Baltimore

Vanagaite L, James MA, Rotman G, Savitsky K, Bar-Shira A, Gilad S, Ziv Y, et al (1995) A high-density microsatellite map of the ataxia-telangiectasia locus. Hum Genet 95:451–454

Verlander PC, Lin J, Udono M, Zhang Q, Gibson R, Mathew C, Auerbach A (1994) Mutation analysis of the Fanconi anemia gene *FACC*. Am J Hum Genet 54:595–601

Weissenbach J, Gyapay G, Dib C, Vignal A, Morissette J, Millasseau P, Vaysseix G, et al (1992) A second-generation linkage map of the human genome. Nature 259:794–801

Whitney M, Thayer M, Reifsteck C, Olson S, Smith L, Jakobs PM, Leach R, et al (1995) Microcell mediated chromosome transfer maps the Fanconi anaemia group D gene to chromosome 3p. Nat Genet 11:341–343