Am. J. Hum. Genet. 66:1821–1832, 2000

Gaucher Disease: The Origins of the Ashkenazi Jewish N370S and 84GG Acid β-Glucosidase Mutations

George A. Diaz,[1,2,*] Bruce D. Gelb,[1,2,*] Neil Risch,[4] Torbjoern G. Nygaard,[5] Amos Frisch,[6,8] Ian J. Cohen,[7,8] Clara Sa Miranda,[9] Olga Amaral,[9] Irene Maire,[10] Livia Poenaru,[11] Catherine Caillaud,[10] Moishe Weizberg,[1] Pram Mistry,[1,3] and Robert J. Desnick[1,2]

Departments of [1]Human Genetics, [2]Pediatrics, and [3]Medicine, Mount Sinai School of Medicine, New York; [4]Department of Genetics, Stanford University, Stanford; [5]Department of Neuroscience, UMDNJ-New Jersey Medical School, Newark; [6]Felsenstein Medical Research Center, Rabin Medical Center, and [7]Schneider Children's Medical Center of Israel, Petah Tikva, Israel; [8]Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv; [9]Instituto de Genetica Medica, Porto, Portugal; [10]Hôpital Debrousse, Lyon; and [11]Laboratoire de Génétique, Hôpital Cochin, Paris

Type 1 Gaucher disease (GD), a non-neuronopathic lysosomal storage disorder, results from the deficient activity of acid β-glucosidase (GBA). Type 1 disease is panethnic but is more prevalent in individuals of Ashkenazi Jewish (AJ) descent. Of the causative GBA mutations, N370S is particularly frequent in the AJ population, ($q \sim .03$), whereas the 84GG insertion ($q \sim .003$) occurs exclusively in the Ashkenazim. To investigate the genetic history of these mutations in the AJ population, short tandem repeat (STR) markers were used to map a 9.3-cM region containing the GBA locus and to genotype 261 AJ N370S chromosomes, 60 European non-Jewish N370S chromosomes, and 62 AJ 84GG chromosomes. A highly conserved haplotype at four markers flanking GBA (PKLR, D1S1595, D1S2721, and D1S2777) was observed on both the AJ chromosomes and the non-Jewish N370S chromosomes, suggesting the occurrence of a founder common to both populations. Of note, the presence of different divergent haplotypes suggested the occurrence of de novo, recurrent N370S mutations. In contrast, a different conserved haplotype at these markers was identified on the 84GG chromosomes, which was unique to the AJ population. On the basis of the linkage disequilibrium (LD) δ values, the non-Jewish European N370S chromosomes had greater haplotype diversity and less LD at the markers flanking the conserved haplotype than did the AJ N370S chromosomes. This finding is consistent with the presence of the N370S mutation in the non-Jewish European population prior to the founding of the AJ population. Coalescence analyses for the N370S and 84GG mutations estimated similar coalescence times, of 48 and 55.5 generations ago, respectively. The results of these studies are consistent with a significant bottleneck occurring in the AJ population during the first millennium, when the population became established in Europe.

Introduction

The unique demographic history of the Ashkenazi Jewish (AJ) population and the occurrence of several lysosomal storage disorders at high frequency in this group have attracted considerable interest and debate concerning the mechanism(s) underlying the prevalence of these diseases among the Ashkenazim (Chase and McCusick 1972; Myrianthopolous et al. 1972; Goodman 1979; Zlotogora et al. 1988; Bonne-Tamir and Adam 1992; Motulsky 1995; Zoosmann-Diskin 1995). The recent availability of high-density genetic maps has facilitated the use of linkage disequilibrium (LD) and coalescence analyses to provide estimated dates for disease-causing mutations. Earlier applications of these methods to AJ genetic diseases demonstrated that >90% of the mutant alleles causing idiopathic torsion dystonia (ITD) arose from a single founder mutation (Risch et al. 1995). The relatively recent coalescence point of the ITD mutation (~16 generations, corresponding to ~400 years), coupled with its relatively high incidence (~1/300) in the present-day AJ, underscored the role that nonuniform population expansion has in establishing the current high frequency of the ITD disease allele in the AJ. Similar analyses have been reported for other mutations in the AJ population, including the coagulation factor XI type II lesion (Peretz et al. 1997; Goldstein et al. 1999), the BRCA1 185delAG deletion (Bar-Sade et al. 1998), and the APC I1307K mutation (Patael et al. 1999). Preliminary studies have suggested that the AJ acid β-glucosidase (GBA) mutation, N370S, the most frequent cause of Gaucher disease (GD [MIM 230800]) in this popu-

Received November 17, 1999; accepted for publication March 14, 2000; electronically published April 21, 2000.

Address for correspondence and reprints: Dr. Robert J. Desnick, Department of Human Genetics, Mount Sinai School of Medicine, New York, NY 10029. E-mail: rjdesnick@vaxa.crc.mssm.edu

* These authors contributed equally to this work.

0002-9297/2000/6606-0011$02.00

lation, shares a common origin with the allele present in the non-Jewish European population, but these studies have not provided a dating estimate for this mutation (Diaz et al. 1998).

In this study, linkage and haplotype analyses were used to investigate the origins of the most prevalent GBA mutations in the AJ population: N370S ($q \sim .03$) and 84GG ($q \sim .003$) (Eng et al. 1997). For these studies, a linkage map of 10 short tandem repeats (STRs) flanking the GBA gene was constructed and then were used to genotype AJ and non-Jewish N370S and AJ 84GG chromosomes. Haplotype and LD analyses indicated that the N370S mutation existed in the general European population prior to the founding, circa the 1st millennium, of the AJ population in Europe. Similar analyses revealed that GD 84GG chromosomes had a unique, conserved haplotype, indicating the occurrence of an ancestral AJ founder. Coalescence analyses suggested that the N370S and 84GG mutations coalesced in the AJ population 48 and 55.5 generations ago (~1,200–1,400 years before present [YBP]), around the time that the AJ population became established in Europe.

Subjects, Material, and Methods

Subjects and Sample Collection

Blood was obtained, with informed consent, from patients with GD type 1 and from relevant family members with GBA mutations N370S and 84GG who were enrolled in the study; these individuals included N370S homozygotes, N370S heterozygotes, and N370S/84GG compound heterozygotes. By means of the Puregene DNA isolation kit (Gentra), genomic DNA was isolated from peripheral leukocytes. Phasing of the STR markers on mutation-bearing chromosomes was initially performed with a subset of 25 AJ and 29 non-Jewish European N370S chromosomes and 4 AJ 84GG chromosomes, from families in which parents and/or siblings were also available. In addition, 60 AJ and 7 Portuguese homozygotes were studied to determine the allele frequencies of the STRs. To obtain allele-frequency data on the 84GG chromosomes, 32 N370S/84GG compound heterozygotes were analyzed. To assess other ethnic groups for the presence of the common N370S founder haplotype, 18 N370S chromosomes from patients of various ethnic origins were haplotyped, including patients with type 1 GD who were of British ($n = 8$), Greek ($n = 4$), Macedonian ($n = 2$), Sephardic Jewish ($n = 2$), Algerian ($n = 1$), and West Indian ($n = 1$) ancestry.

On the basis of preliminary findings, a second set of AJ subjects, including 58 N370S homozygotes and 36 N370S/84GG compound heterozygotes, were genotyped for coalescence analyses. Analysis of the first set of subjects revealed that the most informative markers for LD analysis of N370S chromosomes were D1S2624 and D1S1600; therefore, only these markers were typed for the additional N370S homozygotes. Similarly, initial results indicated that markers D1S305 and D1S2715 were most informative for 84GG chromosomes, so only these markers were typed for 84GG/N370S compound heterozygotes.

Linkage Mapping

Searching the public databases of the Whitehead Institute for Biomedical Research/MIT Center for Genome Research and The Cooperative Human Linkage Center identified maximally informative polymorphic STR markers mapping to 1q21; these included D1S305, D1S1595, D1S1600, D1S1653, D1S2128, D1S2624, D1S2715, D1S2721, D1S2777, and the trinucleotide-repeat marker PKLR. For linkage analysis, genotype data for these markers were obtained from 61 of the CEPH reference families, representing 1,036 meioses. STRs were PCR-amplified under standard conditions (Hudson et al. 1995)—except for markers D1S1595 and D1S1600, which required annealing temperatures of 61°C and 50°C, respectively. The map order and distance between the markers were determined by use of the MultiMap software package (Matise et al. 1994).

Analysis

To estimate allele frequencies on population-specific control chromosomes, as required for LD analysis, 83 AJ chromosomes (39 from family members of patients with GD plus 44 from unrelated AJ families), 63 French chromosomes (15 from unaffected family members plus 48 from French CEPH families) and 61 Portuguese chromosomes (25 from unaffected family members plus 36 from random unrelated individuals) were surveyed. Allele frequencies were determined by direct allele counting.

Estimates of marker-allele frequencies for N370S chromosomes from French and Portuguese subjects also could be obtained directly by allele counting, and variances of these frequencies were obtained with the formula $p(1 - p)/N$, where p is the observed frequency and N is the number of alleles. The situation was more complicated for the AJ population, because of the inclusion of the N370S/84GG compound heterozygotes. In this case, the EM algorithm was used to obtain maximum-likelihood estimates (MLEs) of the allele frequencies (Ceppellini et al. 1995; also see the Appendix). Significance of LD was then assessed by comparison of allele frequencies of mutation-bearing chromosomes versus those of control chromosomes. Specifically, let p_I and σ_p be the allele frequency and its standard error for allele i on mutation-bearing chromosomes, and let r_i and σ_r be

the corresponding frequency and standard error for control chromosomes. Then, statistical significance can be derived by the statistic $Z = (p_i - r_i)/\sqrt{\sigma_p^2 + \sigma_r^2}$. Because multiple alleles were tested at a locus, a statistical correlation for multiple tests was employed. A Bonferroni correction was applied by multiplication of the obtained significance level, *P,* for a one-sided test (since each allele was tested for evidence of *increased* frequency on mutation chromosomes) by the number of alleles tested. Every allele with a frequency >5% on the mutation chromosomes was tested, and, therefore, nominal *P* values were multiplied by the number of such alleles, to obtain a final *P* value. This correction may be conservative, especially for loci with a large number of alleles. In these cases, we also considered the anticonservative correction, $1 - (1 - P)^n$, where n is the number of alleles tested, to determine a range for the actual *P* values

The degree of LD at loci flanking the conserved core markers was assessed by calculation of the parameter δ (Bengtsson and Thomson 1981; Risch et al. 1995). Chromosomes bearing the conserved haplotype (defined as markers varying, from the consensus sequence, at no more than one position) were identified in each population, and associated allele frequencies were calculated for the flanking markers as described above. The δ value was obtained by use of the formula $\delta = (p - r)/(1 - r)$, where p is the frequency of the associated allele on mutation-bearing chromosomes and r is the frequency of the same allele on control chromosomes. A confidence interval for δ was calculated under an assumption of independence for the sampled chromosomes. Since $\delta = 1 - (1 - p)/(1 - r)$ is a function of the ratio of two independent random variables, the variance of δ can be approximated by

$$\mathrm{Var}(\delta) = \frac{1}{(1 - \mathrm{r})^4}\sigma_\mathrm{r}^2(\sigma_\mathrm{p}^2 + (1 - \mathrm{p})^2) + \frac{\sigma_\mathrm{p}^2}{(1 - \mathrm{r})^2} \ .$$

An approximate confidence interval of δ can then be obtained as $\delta + 1.96\sigma_\delta$, where $\sigma_\delta^2 = \mathrm{Var}(\delta)$.

To estimate the coalescence time for the founder mutations, or the number of generations to the most recent common ancestor (MRCA), the formula used was

$$G = \frac{\ln\delta}{\ln(1 - \theta)} \ , \qquad (1)$$

where θ is the recombination fraction with regard to the marker analyzed (Risch et al. 1995). For calculation of *G,* the GBA locus was assumed to colocalize with D1S2777, since the degree of LD observed on N370S and 84GG chromosomes was generally highest at this marker. A minimum confidence interval for *G* was calculated by use of the formula given above, with the minimum confidence interval for δ.

It is difficult to calculate a confidence interval for the estimate of *G,* because, under the assumption of LD, the mutation-bearing chromosomes are not independent (Risch et al. 1995; Rannala and Slatkin 1998). The degree of nonindependence—and, hence, the confidence interval—is strongly influenced by assumptions regarding the demographic history of the population. Rapid expansion after initial introduction of the mutation leads to greater statistical independence and smaller confidence intervals, whereas constant population size leads to less statistical independence and broader confidence intervals (Rannala and Slatkin 1998; Goldstein et al. 1999). Here we provide a minimum confidence interval for *G,* using the calculated 95% confidence interval for δ and equation (1). We note that this interval is appropriate only for a rapidly expanding mutation (which is not unlikely for the mutations under consideration here) and that the true confidence interval may be substantially larger. A more refined estimate of *G,* by examination of multiple markers simultaneously, might be possible in theory; however, in this case, because most of the chromosomes are unphased, it is not clear that such would be easily implemented or of great benefit.

Results

Genotyping Reveals a Founder N370S Haplotype

To facilitate haplotype and LD analyses of the GD N370S and 84GG mutations, a genetic map of polymorphic STR markers flanking the GBA gene was constructed (fig. 1). The genetic map spanned 9.3 cM and contained 10 STRs, several of which were at no measurable recombination distance from each other. The GBA gene was positioned relative to the mapped STRs, on the basis of available physical-mapping and sequence data for the GBA region and for the relevant STRs. The 5′ end of GBA is 71 kb upstream of the 5′ end of the liver-specific pyruvate kinase gene (Demina et al. 1998). An intragenic trinucleotide repeat in intron 11 of this gene, PKLR (Lenzner et al. 1994), colocalized with markers D1S1595 and D1S2721 on the genetic linkage map. Physical-mapping data placed GBA on YACs containing marker D1S2777. Thus, the GBA locus lies very close to the four central markers PKLR, D1S1595, D1S2721, and D1S2777.

For the N370S mutation, 54 homoallelic or heteroallelic AJ individuals in 19 families were genotyped for the 10 mapped STRs. Haplotype analysis of these families with GD revealed a highly conserved 6-5-4-6 haplotype for the markers PKLR-D1S1595-D1S2721-D1S2777, respectively (table 1). Additional genotyping of 61 AJ patients who were homoallelic for the N370S mutation demonstrated conservation of the ancestral haplotype at three or all four of these core STRs, in 133

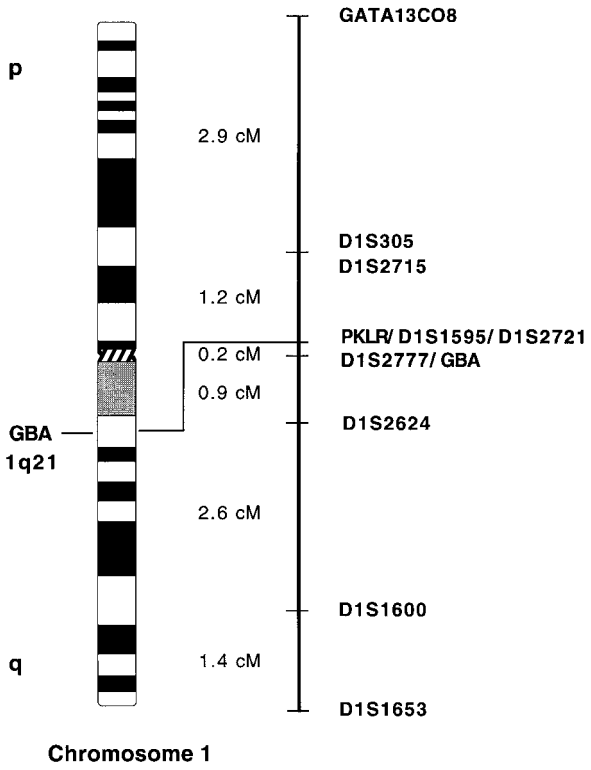

Figure 1 Linkage map of the GBA region. The order of 10 polymorphic STRs flanking the GBA gene on 1q21 was determined by linkage mapping using families from the CEPH panel. Markers are arrayed from centromere (*top*) to telomere (*bottom*), with the obtained recombination distances shown. The position of the GBA gene as derived from physical mapping, sequence, and LD data is indicated.

(95%) of the total 140 informative AJ N370S alleles, with only seven chromosomes varying at two or more STRs (table 2). Of note, the conserved haplotype was uncommon, found in only 3 (3.6%) of 83 of the control AJ chromosomes (data not shown). These could also be N370S carriers, since the population frequency of the mutation is ~6% (Eng et al. 1997).

Genotyping of non-Jewish European patients with GD type 1 who were homoallelic or heteroallelic for the N370S mutation revealed the same founder haplotype in the majority of individuals (table 1). This group included 13 Portuguese and 12 French families with GD, from which 14 Portuguese and 15 French N370S haplotypes were phased. An additional 14 Portuguese and 2 French chromosomes from patients homoallelic for N370S also were genotyped. The ancestral 6-5-4-6 haplotype was not detected in 44 French control haplotypes (29 derived from French families in the CEPH Genotype Database and 15 derived from unaffected French individuals in this study) or in 25 Portuguese control haplotypes, thereby confirming its rarity in these populations. Thus, a majority of the N370S alleles in AJ and non-Jewish European populations shared the common 6-5-4-6 haplotype.

Among the French and the Portuguese N370S chromosomes, the degree of allelic diversity at the conserved core markers was strikingly greater than that in the AJ chromosomes. Approximately 18% (3 of 17) and 33% (10 of 27) of French and Portuguese N370S chromosomes, respectively, varied from the ancestral 6-5-4-6 haplotype at two or more markers, as compared with 5% of the AJ N370S chromosomes (table 2). Indeed, the frequency of the associated founder allele for each

Table 1

Allele-Frequency Estimates for N370S, 84GG, and CEPH Control Chromosomes, for the Central Four Markers

Marker and Allele[a]	Frequency of Chromosomes				
	Controls	N370S		84GG	
	CEPH)	Ashkenazim	French	Portuguese	Ashkenazim
	N = 400	*N* = 117	*N* = 17	*N* = 23	*N* = 37
PKLR:					
1	.165	.000	.000	.043	.865
2	.092	.009	.000	.000	.135
3	.362	.009	.118	.000	.000
4	.215	.026	.059	.000	.000
5	.080	.026	.059	.000	.026
6	.078	.932	.706	.783	.015
7	.008	.000	.059	.174	.000
	N = 388	*N* = 142		*N* = 26	*N* = 39
D1S1595:					
2	.131	.007	.000	.000	.000
3	.062	.000	.000	.077	.000
4	.000	.000	.000	.077	.000
5	.330	.958	.824	.731	.078
7	.294	.028	.118	.115	.948
8	.103	.007	.059	.000	.000
	N = 374	*N* = 137	*N* = 16	*N* = 26	*N* = 38
D1S2721:					
1	.024	.007	.000	.000	.000
2	.064	.007	.000	.077	.000
3	.219	.029	.038	.077	.000
4	.235	.905	.563	.577	.000
5	.045	.007	.063	.038	1.000
6	.094	.015	.063	.038	.000
7	.035	.007	.000	.000	.000
8	.275	.022	.313	.192	.000
	N = 420	*N* = 147	*N* = 17	*N* = 27	*N* = 39
D1S2777:					
2	.000	.000	.000	.000	.000
3	.021	.000	.000	.000	.000
4	.093	.007	.059	.148	.000
5	.198	.014	.059	.111	1.000
6	.543	.973	.882	.741	.000
7	.064	.007	.000	.000	.000

[a] For clarity of presentation, uncommon alleles that are not present in disease chromosomes have been omitted.

Table 2

Conservation of N370S Chromosomes from Various Origins

Population	No. of N370S Chromosomes		% of N370S Chromosomes Conserved
	Total	Ancestral[a]	
Ashkenazi	140	133	95
Portuguese	28	18	64
French	17	14	82
Other	18	15	83

[a] At least three of the four core haplotype markers are conserved.

of the four loci was significantly lower ($P < .05$) in both the French and the Portuguese N370S chromosomes than in the AJ N370S chromosomes (table 1). This observation could be explained either by greater decay of the conserved haplotype in the non-Jewish populations or by recurrent N370S mutations in different haplotypes. The former possibility would imply that the N370S founder mutation had existed for a longer time in the non-Jewish populations than in the Ashkenazim and suggests the direction of gene flow for this particular allele.

Genotyping also was performed on GD N370S chromosomes of British ($n = 8$), Greek ($n = 4$), Macedonian ($n = 2$), Sephardic Jewish ($n = 2$), Algerian ($n = 1$), and West Indian ($n = 1$) ancestry. Haplotypes were inferred for 14 of these 18 chromosomes. The remaining four chromosomes were from two patients—one Greek and one Macedonian—who were homoallelic for N370S. The Greek patient's genotype was divergent at PKLR, for both alleles (i.e., 3/5) and at D1S1595 for one allele (i.e., 5/2); thus, one chromosome had the ancestral haplotype at three of the four conserved markers whereas the other had only two of the conserved alleles. The Macedonian patient was heterozygous at all of the markers, consistent with one completely conserved and one completely divergent chromosome. Of these 18 N370S chromosomes from diverse ethnic/demographic groups, 15 (83%) had the ancestral 6-5-4-6 haplotype, a frequency comparable to the 82% of ancestral chromosomes among the French N370S alleles (table 2).

LD at Markers Flanking the Conserved N370S Haplotype

Table 3 shows the allele frequencies for the STRs flanking the conserved 6-5-4-6 markers for the AJ N370S chromosomes bearing the conserved 6-5-4-6 haplotype, for all the 84GG chromosomes, and for the control AJ chromosomes. The allele frequencies for mutation-bearing chromosomes were compared with those for control chromosomes, to assess LD. For N370S, four markers showed significant LD. For three of the markers, a single allele was at significantly increased frequency (D1S2715, allele 9, $Z = 4.11$, corrected $P < .001$; D1S2624, allele 5, $Z = 8$, corrected $P < .001$; and D1S1600, allele 2, $Z = 7.11$, corrected $P < .001$). For D1S305, the situation was more complicated; three alleles (alleles 4, 5, and 6) were at increased frequency, although only the increase in frequency of alleles 5 and 6 was statistically significant (allele 5, $Z = 2.94$, corrected $P < .01$; and allele 6, $Z = 3.16$, corrected $P < .01$); however, it is likely that either the 4 allele or 5 allele is the ancestral allele, as detailed below.

Coalescence Estimates for AJ N370S Chromosomes

The presence of detectable LD at markers a defined map distance from the GBA locus permitted the estimation of the coalescence times for the founder N370S and 84GG mutations in the AJ population. The δ values given in table 3 (for all markers except D1S305 for N370S) and the recombination fractions from table 4 were used to calculate the values of G, the number of generations to the MRCA, for each mutation. As noted above, several alleles at marker D1S305 appeared to be associated with the N370S mutation. Therefore, all N370S chromosomes with the disease-associated allele 9 at D1S2715 (by phasing or homozygosity) were identified, and the frequencies of the various D1S305 alleles were calculated for this subset of 69 chromosomes. The D1S305 allele frequencies were as follows: allele 4, 63.8% ($n = 44$); allele 5, 23.2% ($n = 16$); allele 6, 3.0% ($n = 2$); and allele 8, 10.1% ($n = 7$). Two alleles had significantly increased frequencies: allele 4 ($Z = 2.05$, $P < .05$) and allele 5 ($Z = 2.59$, $P < .01$). In this subset, allele 6 did not have increased frequency. These results suggested either that allele 4 or allele 5 was the progenitor or that an early recombination or mutation event had led to an increased frequency of both of these alleles on N370S chromosomes. Therefore, a δ value was calculated, by combining of both of these alleles, leading to a δ estimate of $[.516 + .195 - (.459 + .071)]/[1 - (.459 + .071)] = .385$ (table 4). As noted in table 3, allele 4 at D1S305 and allele 9 at D1S2715 were very common in the AJ control chromosomes. Thus, only markers D1S2624 and D1S1600 were typed in the follow-up set of N370S chromosomes.

To account for the uncertainty inherent in the mapping of tightly linked loci, the θ values from the present study were compared with those from the genetic and integrated maps publicly available from the Center for Medical Genetics, Marshfield Medical Research Foundation and The Genetic Location Database of the University of Southampton (table 4). The fact that the median θ values at three of the four markers used for coalescence analysis were identical to the θ values ob-

Table 3

Allele-Frequency Estimates for N370S, 84GG, and Normal Control Chromosomes, with Associated δ Values

Locus and Allele	Frequency (δ) in						
	AJ Population			French Population		Portuguese Population	
	Control	N370S	84GG	Controls	N370S	Controls	N370S
	N = 85	*N* = 193	*N* = 58	*N* = 66	*N* = 16	*N* = 24	*N* = 27
D1S305:							
3	.094	.017	.100	.015	.000	.083	.000
4	.459	.516 (.105)	.154	.439	.438 (0)	.375	.185 (<0)
5	.071	.195* (.133)	.059	.015	.125 (.112)	.042	.000 (<0)
6	.024	.121* (.099)	.027	.152	.313 (.190)	.125	.259 (.153)
8	.235	.106	.588 (.461)	.227	.125	.167	.056
9	.094	.031	.052	.121	.000	.167	.333
10	.012	.010	.020	.015	.000	.041	.000
	N = 83	*N* = 182	*N* = 53	*N* = 58	*N* = 16	*N* = 51	*N* = 26
D1S2715:							
2	.000	.006	.452** (.452)	.103	.063	.118	.077
3	.108	.015	.165	.052	.063	.059	.038
4	.301	.133	.130	.293	.313	.353	.308
5	.012	.017	.000	.034	.000	.039	.000
6	.060	.060	.000	.000	.063	.000	.000
8	.024	.006	.019	.069	.063	.078	.000
9	.422	.688** (.460)	.099	.293	.375 (.120)	.333	.500 (.250)
10	.024	.076	.098	.317	.125	.020	.000
	N = 83	*N* = 268	*N* = 32	*N* = 60	*N* = 16	*N* = 50	*N* = 24
D1S2624:							
1	.024	.000	.000	.050	.125	.000	.000
2	.313	.134	.088	.250	.438	.300	.208
3	.036	.008	.000	.067	.000	.100	.042
4	.373	.198	.912** (.860)	.267	.313	.260	.458
5	.205	.631** (.536)	.000	.317	.125 (<0)	.320	.291 (<0)
6	.048	.030	.000	.000	.000	.000	.000
	N = 75	*N* = 258	*N* = 32	*N* = 61	*N* = 17	*N* = 24	*N* = 24
D1S1600:							
2	.067	.326** (.316)	.000	.033	.000 (<0)	.208	.042 (<0)
3	.427	.207	.394	.344	.294	.333	.458
4	.333	.346	.368	.410	.412	.167	.333
5	.173	.075	.238 (.079)	.180	.294	.125	.167

* $P < .01$.
** $P < .001$.

tained in the present study supports the overall accuracy of the current linkage map. The age estimates for N370S, for D1S305, D1S2715, D1S2624, and D1S1600, varied with the genetic map used but ranged from a minimum of 17 generations (marker D1S2624) to a maximum of 136 generations (marker D1S305), both estimates being based on the Southampton map. When the medians of the θ values from the three maps were used, the range was 32–68 generations. The minimum confidence intervals when median θ values were used were 37–123 generations for D1S305, 34–84 generations for D1S2715, 31–53 generations for D1S2624, and 26–40 generations for D1S1600. The latter two were based on substantially larger sample sizes. The median value for *G* was 48 generations. If it is assumed that a generation comprises 25 years, the coalescence time for this mutation was 1,200 YBP, or approximately year 800 of the common era (C.E.). It is noteworthy that the founding of the AJ population at ~900 C.E., when several thousand Jews migrated into the Rhineland and subsequently expanded into a much larger population (Weinryb 1972), falls within the confidence intervals for these estimates of the coalescence. We note that our analysis did not include mutation at the markers (e.g., see Goldstein et al. 1999). Presumably, the marker mutation rate is substantially less than the recombination rate considered here (i.e.,

Table 4

Estimates of Coalescence Generations for N370S and 84GG Chromosomes in the AJ Population.

	N370S					84GG				
		G (θ)[a]					G[a]			
MARKER	δ	Marshfield	Southampton	Present Study	Median	δ	Marshfield	Southampton	Present Study	Median
D1S305	.385	56 (.017)	136 (.007)	68 (.007)	68 (.014)	.461	45	110	55	55
D1S2715	.460	45 (.017)	111 (.007)	55 (.007)	55 (.014)	.452	46	113	56	56
D1S2624	.536	41 (.015)	17 (.036)	69 (.009)	41 (.015)	.860	10	4	17	10
D1S1600	.316	38 (.030)	29 (.039)	32 (.035)	32 (.035)	.079	83	64	71	71

[a] Recombination distances Marshfield = Center for Medical Genetics, Marshfield Medical Research Foundation; Southampton = The Genetic Location Database of the University of Southampton. θ Values are from the various maps, under the assumption that GBA colocalizes with D1S2777.

1.5%–3.5%). Inclusion of mutation within the coalescence analysis would lead to coalescence times slightly more recent than those given in table 4.

LD Analysis of the Non-Jewish N370S Chromosomes

The degree of LD with associated alleles on N370S founder chromosomes in the AJ population can be contrasted with that in the French and Portuguese populations (table 3). At D1S2715, the associated allele 9, which has a δ value of .46 in the AJ population, was also associated in the French and Portuguese populations, but to a lesser extent (δ = .25 for the Portuguese population, and δ = .12 for the French population), and the difference was statistically significant for each ($P <$ .05). At D1S2624, the associated allele 5 allele was not increased in the Portuguese or the French samples ($\delta <$ 0, $P <$ .05). Finally, at D1S1600, the associated allele 2 allele was not increased in either the Portuguese or the French populations (again, $P <$ .05). Indeed, no other consistent allelic associations at any of these markers were present in the Portuguese and the French populations. This could not be due to the smaller sample size of the French and the Portuguese populations, since a comparable level of LD (i.e., δ = .46 or .536) would have also been statistically significant in those populations. Thus, there was minimal LD on N370S progenitor chromosomes at 0.9%–1.5% recombination distance in the non-Jewish European populations. These results strongly suggest that the origin of the founder N370S mutation was more widespread and more ancient than it would be if it had occurred with the founding of the AJ population.

Although the absence of significant LD in the non-Jewish European population precluded direct coalescence analysis, it was possible to estimate the relative age of the N370S mutation in this population. The age of a mutation in a population is proportional to logδ, so that, even in the absence of known θ values, the ratio of ages can be computed by the ratio of the logδ values. Table 5 show, for the four conserved markers, the δ values and the ratios of logδ values in the AJ population versus those in the French or the Portuguese populations. For the French population, the median R value is 5.0; for the Portuguese population, the median R value is 7.0; and the median R value for both populations is 5.6. Assuming an age of 1,200 years for the AJ population gives an estimate age of ~6,700 years for the non-Jewish European N370S mutation. This value may be a slight overestimation, because of the inclusion of de novo mutations, as discussed below. Nonetheless, it suggests that little LD existed at the flanking loci in the extant population at the time of the founding of the Ashkenazim and that survival of a single AJ N370S chromosome from the founding has led to LD at the adjacent markers.

Evidence of De Novo N370S Mutations

The allelic diversity at the conserved N370S 6-5-4-6 haplotype, especially in the non-Jewish populations, raised the possibility of distinct, recurrent mutations. Thus, N370S chromosomes carrying atypical haplotypes were screened for the Pv1.1 intragenic polymorphism (Horowitz et al. 1989). This marker is one of 12 intragenic polymorphisms in the GBA gene that form two haplotypes (i.e., + and −) that are in complete LD with each other. The (−) haplotype occurs at a frequency of ~.7 in the general population and has been invariably associated with the N370S mutation, a finding that is suggestive of a common founder mutation for the N370S allele (Zimran et al. 1990; Beutler et al. 1992). Of the 23 atypical chromosomes, defined as having two or more variant alleles at the conserved 6-5-4-6 haplotype, 3 had a Pv1.1(+) allele (table 6), which was confirmed by direct sequencing (not shown), providing evidence of recurrent N370S mutations. If a proportion, q, of these 23 atypical chromosomes were nonancestral, then .3q would be the frequency of the Pv1.1(+) alleles among these chromosomes. Thus, .3q = 3/23 = .13, so that we can estimate q as .43; that is, 43% of the atypical hap-

Table 5

δ Values and Ratio, *R*, of logδ, for French and Portuguese N370S Chromosomes, Compared with Those of AJ Chromosomes

Marker	δ (*R*) for AJ Population	French Population	Portuguese Population
PKLR	.926	.681 (5.0)	.765 (3.5)
D1S1595	.937	.737 (4.7)	.598 (7.9)
D1S2721	.876	.429 (6.4)	.447 (6.1)
D1S2777	.941	.742 (4.9)	.433 (13.8)

lotypes carry de novo mutations, while the remainder represent recombination or marker mutation events.

Haplotype and LD Analyses for the 84GG Mutation

The presence of a conserved haplotype at the markers PKLR-D1S1595-D1S2721-D1S2777 was apparent from examination of the N370S/84GG genotype data. At each marker, two alleles were much more frequent than they were in control chromosomes, with the first allele corresponding to the conserved N370S allele and with the second allele defining the conserved 84GG allele. The allele frequencies given in table 1 were calculated by direct counting of the non-N370S allele at each marker, except in two individuals, one of whom was heterozygous for a nonconserved N370S allele and the other of whom was heterozygous for a conserved 84GG allele at single markers. These alleles defined a 1-7-5-5 haplotype at the markers PKLR-D1S1595-D1S2721-D1S2777 for the 84GG chromosomes. Phasing of four 84GG chromosomes demonstrated that three of these chromosomes had the completely conserved haplotype and that the fourth had a single variation (i.e., allele 2 at PKLR). Thus, the haplotypes from the phased chromosomes concurred with the haplotype predicted by the allele-frequency estimates.

The allele frequencies at nonconserved markers in the 84GG chromosomes were estimated by maximum likelihood (see Subjects, Material, and Methods). As seen in table 3, three markers flanking the conserved haplotype showed significant LD, which, in each case, was attributable to a single allele (D1S305, allele 8, $Z = 4.16$, corrected $P < .001$; D1S2715, allele 2, $Z = 6.52$, corrected $P < .001$; and D1S2624, allele 4, $Z = 7.17$, corrected $P < .001$). At D1S1600, one allele (allele 5) was modestly increased but did not reach statistical significance.

Coalescence Analysis of the 84GG Mutation

Coalescence analysis of the 84GG mutation gave similar age estimates for the two proximal markers, D1S305 and D1S2715 (table 4), but gave somewhat disparate results for the distal markers, D1S2624 and D1S1600. When median θ values were used, the minimum confidence intervals for D1S305, D1S2715, and D1S1600 were overlapping (30–93, 38–81, and 33–∞ generations, respectively). The interval for D1S2624 (0–25 generations) did not overlap the minimum intervals for the other three markers, perhaps reflecting the fact that these are minimal intervals and do not account for potential errors in θ. A median value of G for the 84GG chromosomes was ~55.5 generations (or ~1,400 YBP), similar to that observed for N370S. This is probably a reasonable estimate, because marker D1S2715 ($G = 56$) provided the greatest accuracy, since allele 2 on 84GG chromosomes was absent on control chromosomes and was nearly absent on N370S chromosomes. In contrast, the associated D1S2624 allele 4 ($G = 10$) was very common (frequency .373) on control chromosomes.

Discussion

In the present article, we have reported detailed haplotype and LD analyses that build on our earlier work (Diaz et al. 1998), in an attempt to determine the genetic history of the GBA N370S and 84GG mutations causing GD. Previous studies of the intragenic GBA polymorphism, Pv1.1 in intron 6, found that all N370S chromosomes examined occurred on the Pv1.1(−) haplotype (Zimran et al. 1990; Beutler et al. 1992; Lacerda et al. 1994*a*). The results of those studies were suggestive of a common founder but were not conclusive, since the (−) haplotype was found to have a population frequency of ~.7 (Beutler et al. 1992; Lacerda et al. 1994*a*). In

Table 6

Atypical N370S Chromosomes

Population	Allele(s) for Marker PKLR	D1S1595	D1S2721	D1S2777	Pv1.1 Allele(s)
Ashkenazim	6	7	7/4	4	−
	4	8	8/3	6	−
	5	5	3	6	−
	3/4[a]	5/7	5/8	6/6	−/−
	4/2[a]	7/5	8/1	6/6	+/+
Portuguese	6	5	8	4	−
	7	4	4	6	−
	7	5	8, 5	4	−
	No data	5/5	8/8	5/5	−/−
	6/6[a]	7/7	2/2	4/6	−/−
	6/1[a]	3/5	8/3	5/4	−/−
French	5	5	8	6	−
	6	7	8	6	−
	7	5	8/5	4	−
Macedonian	5/6[a]	8/5	3/4	5/6	+/−
Greek	3/5[a]	2/5	4/4	6/6	−/−
English	6	7	8	7	−−

[a] Unphased genotype data from N370S homozygotes.

contrast, the results described here clearly demonstrate that most N370S chromosomes derive from an ancestral founder. In addition, our study has documented the existence of de novo N370S mutations occurring on a Pv1.1(+) haplotype. Such alleles are uncommon, since only three Pv1.1(+) alleles were identified in the 303 N370S chromosomes haplotyped. The mechanism of recurrence is unclear, since the mutation does not occur at a CpG dinucleotide; nor are degenerate repeat motifs present that could contribute to replication errors.

Because of its high prevalence in the AJ population, N370S has been considered to be a Jewish mutation. It was speculated that its presence in non-Jewish European populations might have resulted from AJ admixture (Lacerda et al. 1994*b*). However, the results of LD analyses reported here, of N370S chromosomes with the ancestral haplotype from AJ and from non-Jewish French, Portuguese, and other ethnic groups, are not consistent with gene flow from the AJ into the non-AJ population. Despite the fact that the conserved four haplotype markers tightly linked to the GBA locus show very strong disequilibrium for the N370S chromosomes, only the AJ chromosomes show detectable LD at the more distant markers on the 9.3-cM map. The increased allelic diversity in non-Jewish N370S chromosomes compared with that in the AJ N370S chromosomes is consistent with a more ancient coalescence point of N370S in the non-Jewish European populations. The age estimate (~6,700 YBP) of the non-Jewish European N370S mutation, calculated by means of the ratio of the logδ values, suggests that little LD existed at the flanking markers in the extant population at the time of the Ashkenazi founding and that survival of a single AJ N370S chromosome from the founding has led to LD at these adjacent markers.

Although confidence intervals are given for the estimates of the mutation coalescence in this study, there are several additional sources of uncertainty that are worth noting. First, map distances and even some marker orders varied among the available genetic maps. This is not surprising, since the short distances (and correspondingly small number of observed recombinant events between markers) can be readily altered by small-sample bias or by single genotyping errors. Second, the position of the GBA gene relative to the four conserved markers is not precisely known. The *G*- estimate calculations reported here assume that GBA colocalizes with D1S2777, for two reasons: (1) the δ values for the AJ and the French N370S chromosomes and for the AJ 84GG chromosomes were maximal at this marker, and (2) instances of alternative alleles for D1S2777 on N370S chromosomes were generally consistent with mutation at the marker (i.e., the flanking markers were conserved), whereas this was less true for the other three central markers. Third, coalescence analysis does not distinguish recombination between markers from mutations of a marker. High mutation frequencies, which lead to the allelic variability that make these markers polymorphic, tend to reduce δ values and to result in a more ancient estimate for the coalescence time. Even if we account for these sources of uncertainty, it is striking that the confidence intervals for the coalescence times of the N370S and 84GG mutations include the approximate time of the Ashkenazi founding, supporting a founder effect to explain the prevalence of these mutations in the population.

Our mutation-dating results are widely discrepant with those in a recent letter comparing haplotypes observed in AJ and Spanish patients with GD and dating the N370S mutation in the Jewish population to an interval ~4,200–9,500 YBP (Diaz et al. 1999). Although some variation can be attributed to the use of different map distances, this large discrepancy suggests the existence of a calculation error. When the data presented in that report and the median recombination distance described above were used to estimate *G*, a value of $G = 50$ generations was obtained, in close agreement with the results of the present study.

Several disease mutations present in the AJ can be traced back to Middle Eastern progenitor Jewish populations. These include the factor XI type II mutation (Peretz et al. 1997; Goldstein et al. 1999), the BRCA1 185delAG mutation (Bar-Sade et al. 1998), and the APC I1307K mutation (Patael et al. 1999). These examples appear to be ancient mutations that are unique to the Jewish people. In contrast, the results reported here indicate that the N370S founder haplotype exists widely in non-Jewish European chromosomes. The distribution of the N370S mutation is more like that of the factor II polymorphism, G20210A, which is associated with an increased risk of thromboembolism. The G20210A allele is found at high prevalence among the AJ but is also found at lower levels in various non-Jewish populations (Zivelin et al. 1998). The presence of this factor II mutation in multiple European populations, in conjunction with its absence in non-European populations, suggests that it originated in an ancient progenitor population postdating the divergence of European and Asian populations. The same inference can be made for the N370S mutation, which has not been found in any Asian populations studied to date (Kim et al. 1996; Choy et al. 1997; Eto and Ida 1999).

The underlying genetic mechanism(s) accounting for the high prevalence of the N370S and 84GG mutations in the AJ population, as well as for other common AJ mutations, has been the subject of considerable debate. The relatively high prevalence of several lysosomal disorders in the AJ population, each resulting from several independent mutations, suggests the hypothesis that heterozygote advantage was the operative genetic mech-

anism. Although this suggestion is intuitively appealing, a clear biological benefit of heterozygosity for these lysosomal mutations has never been documented. The finding that a common N370S allele is concurrently present in the AJ population and in the neighboring non-Jewish European populations allows some inferences to be drawn. To explain the current N370S allele prevalence among the AJ by heterozygote advantage would require that selection operate only in this population and not on proximate non-AJ populations. In instances in which heterozygosity has conferred demonstrable selective advantage to human populations, by affecting susceptibility to common infectious diseases, such uneven selection has not been observed. Selection for β-globin mutations, which offer protection against malarial infection (Williams et al. 1996), has been observed in multiple populations inhabiting malaria-endemic regions. Similarly, the prevalence of the HIV-coreceptor mutation, CCR5Δ32, is approximately the same in the AJ population as in other, northern-European populations. This mutation, which has no disease phenotype, is proposed to have recently risen to its current frequency of .10–.13 in northern-European populations by conferring immunity to bacterial pathogens that enter cells by means of the CCR5 receptor (Libert et al. 1998; Stephens et al. 1998). If heterozygosity for N370S confers resistance to a common pathogen, it is unlikely that increased allele prevalence would be limited to the AJ population.

In contrast to N370S, the 84GG mutation is unique to the AJ population and appears to coalesce at approximately the same historical period as does the N370S lesion. If the ancestral alleles for both of these mutations were extant at the founding of the AJ population, ~1,200 YBP, as suggested by the coalescence analyses, then the striking differences in mutation prevalence in the current AJ population highlight the importance of demographic (or epigenetic) factors in the shaping of allele frequencies, at least within this frequency range. Although 84GG is a null allele that is lethal in homozygosity (Tayebi et al. 1997; R. J. Desnick and C. M. Eng, unpublished results), the negative-selection coefficient is minimal, because of the low frequency of other genetic-lethal GD alleles in the AJ population. Previous work by Risch et al. (1995) has suggested that the present-day AJ population is not the product of uniform population growth but, rather, is derived primarily from expansion of a subpopulation. The recent finding that the prevalence of Tay-Sachs carriers among carriers of GD is lower than would be expected on the basis of population-frequency data suggests that these mutations have not yet reached genetic equilibrium, a finding that is consistent with the presence of population subdivisions within the AJ population (Peleg et al. 1998). Thus, a likely explanation for the currently observed GD allele frequencies appears to be nonuniform expansion of the AJ population after the founding bottleneck. In support of this conclusion, computer simulations modeling the introduction of disease alleles into a rapidly growing population have produced results consistent with skewing of disease-allele frequencies, as is seen in the AJ population, in the absence of selection and in the time frame under discussion (N. Risch, personal communication). Taken as a whole, the findings in this and other studies indicate the importance of genetic drift in shaping the disease-allele frequencies observed in the contemporary AJ population.

Acknowledgments

The authors would like to thank Drs. Christine Eng, Marie Grace, and Lea Peleg, for additional clinical materials used in these studies, and Dr. Rina Zaizov, for longstanding active involvement in the clinical and research aspects of this project. These studies were supported in part by a Lucy Moses Foundation Fellowship (to G.A.D.) and by NIH grant 5 P30 HD 28822 (to the Mount Sinai Child Health Research Center).

Appendix

Estimation of Allele Frequencies in Disease Chromosomes, by EM Algorithm

Suppose that there are m marker alleles denoted 1–m. Let p_i represent the frequency of allele i on N370S chromosomes, and let q_i represent the frequency of allele i on 84GG chromosomes. Let a_i represent the number of i alleles observed on N370S chromosomes (either from homozygotes or from phased heterozygotes), and let b_i represent the corresponding number on 84GG chromosomes, and let c_{ij} represent the number of ij genotypes found on the N370S/84GG compound heterozygotes. Let $A = \Sigma a_i$, $B = \Sigma b_i$, and $C = \Sigma c_{ij}$ (note that $c_{ij} = c_{ji}$). The following recursion formulas can then be used to obtain the MLEs of p_i and q_i:

$$p_i' = \frac{a_i + c_{ii} \sum_{j \neq i} c_{ij} * p_i q_j / (p_j q_i + p_i q_j)}{A + 1/2C}$$

and

$$q_i' = \frac{b_i + c_{ii} \sum_{j \neq i} c_{ij} * p_i q_j / (p_j q_i + p_i q_j)}{B + 1/2C} .$$

Standard errors for the estimates of p_i and q_i can be obtained by standard likelihood theory, by inverting

the expected information matrix; for example, for frequencies p_i and q_i, the respective variances, σ_p and σ_q, are given by $\sigma_p^2 = \omega_{22}/(\omega_{11}\omega_{22} - \omega_{12}^2)$ and $\sigma_q^2 = \omega_{11}/(\omega_{11}\omega_{22} - \omega_{12}^2)$, where

$$\omega_{11} = \frac{A}{p_i(1-p_i)} + \frac{C(p_i + q_i - 2p_iq_i)}{p_i(1-p_i)} + \frac{C(1-2q_i)^2}{(p_i + q_i - 2p_iq_i)} ,$$

$$\omega_{22} = \frac{B}{q_i(1-q_i)} + \frac{C(p_i + q_i - 2p_iq_i)}{q_i(1-q_i)} + \frac{C(1-2p_i)^2}{(p_i + q_i - 2p_iq_i)} ,$$

and

$$\omega_{12} = \frac{C}{(p_i + q_i - 2p_iq_i)} .$$

Electronic-Database Information

Accession numbers and URLs for data in this article are as follows:

CEPH Genotype Database, http://www.cephb.fr/cephdb
Cooperative Human Linkage Center, The, http://lpg.nci.nih.gov/CHLC
Genetic Location Database, The, http://cedar.genetics.soton.ac.uk/public_html/ldb.html
Center for Medical Genetics, Marshfield Medical Research Foundation, http://www.marshmed.org/genetics
Online Mendelian Inheritance in Man (OMIM), http://www.ncbi.nim.nih.gov/omim (for GD [MIM 230800])
Whitehead Institute for Biomedical Research/MIT Center for Genome Research, http://www-genome.wi.mit.edu

References

Bar-Sade RB, Kruglikova A, Modan B, Gak E, Hirsh-Yechezkel G, Theodor L, Novikov I, et al (1998) The 185delAG BRCA1 mutation originated before the dispersion of Jews in the Diaspora and is not limited to Ashkenazim. Hum Mol Genet 7:801–805

Bengtsson BO, Thomson G (1981) Measuring the strength of associations between HLA antigens and diseases. Tissue Antigens 18:356–363

Beutler E, West C, Gelbart T (1992) Polymorphisms in the human glucocerebrosidase gene. Genomics 12:795–800

Bonne-Tamir B, Adam A (1992) Genetic diversity among Jews. Oxford University Press, New York

Ceppellini R, Siniscalco M, Smith CAB (1955) The estimation of gene frequencies in a random-mating population. Ann Hum Genet 20:97–115

Chase GA, McCusick VA (1972) Controversy in human genetics: founder effect in Tay-Sachs disease. Am J Hum Genet 24:339–340

Choy FY, Humphries ML, Shi H (1997) Identification of two novel and four uncommon missense mutations among Chinese Gaucher disease patients. Am J Med Genet 71:172–178

Demina A, Boas E, Beutler E (1998) Structure and linkage relationships of the region containing the human L-type pyruvate kinase (PKLR) and glucocerebrosidase (GBA) genes. Hematopathol Mol Hematol 11:63–71

Diaz A, Montfort M, Cormand B, Zeng B, Pastores GM, Chabas A, Vilageliu L, et al (1999) Gaucher disease: the N370S mutation in Ashkenazi Jewish and Spanish patients has a common origin and arose several thousand years ago. Am J Hum Genet 64:1233–1238

Diaz GA, Risch N, Nygaard T, Maire I, Poenaru L, Caillaud C, Sa Miranda C, et al (1998) Gaucher disease: the Ashkenazi Jewish N370S mutation occurred on an ancient European haplotype. Am J Hum Genet Suppl 63:A211

Eng CM, Schechter C, Robinowitz J, Fulop G, Burgert T, Levy B, Zinberg R, et al (1997) Prenatal genetic carrier testing using triple disease screening. JAMA 278:1268–1272

Eto Y, Ida H (1999) Clinical and molecular characteristics of Japanese Gaucher disease. Neurochem Res 24:207–211

Goldstein DB, Reich DE, Bradman N, Usher S, Seligsohn U, Peretz H (1999) Age estimates of two common mutations causing factor XI deficiency: recent genetic drift is not necessary for elevated disease incidence among Ashkenazi Jews. Am J Hum Genet 64:1071–1075

Goodman RM (1979) Genetic disorders among the Jewish people. Johns Hopkins University Press, Baltimore

Horowitz M, Wilder S, Horowitz Z, Reiner O, Gelbart T, Beutler E (1989) The human glucocerebrosidase gene and pseudogene: structure and evolution. Genomics 4:87–96

Hudson TJ, Stein LD, Gerety SS, Ma J, Castle AB, Silva J, Slonim DK, et al (1995) An STS-based map of the human genome . Science 270:1945–1954

Kim JW, Liou BB, Lai MY, Ponce E, Grabowski GA (1996) Gaucher disease: identification of three new mutations in the Korean and Chinese (Taiwanese) populations. Hum Mutat 7:214–218

Lacerda L, Amaral O, Pinto R, Aerts J, Sa MMC (1994*a*) The N370S mutation in the glucocerebrosidase gene of Portuguese type 1 Gaucher patients: linkage to the *Pvu*II polymorphism. J Inherit Metab Dis 17:85–88

Lacerda L, Amaral O, Pinto R, Oliveira P, Aerts J, Sa MMC (1994*b*) Gaucher disease: N370S glucocerebrosidase gene frequency in the Portuguese population. Clin Genet 45: 298–300

Lenzner C, Jacobasch G, Reis A, Thiele B, Nurnberg P (1994) Trinucleotide repeat polymorphism at the PKLR locus. Hum Mol Genet 3:523

Libert F, Cochaux P, Beckman G, Samson M, Aksenova M, Cao A, Czeizel A, et al (1998) The δccr5 mutation conferring protection against HIV-1 in Caucasian populations has a single and recent origin in Northeastern Europe. Hum Mol Genet 7:399–406

Matise TC, Perlin M, Chakravarti A (1994) Automated con-

struction of genetic linkage maps using an expert system (MultiMap): a human genome linkage map. Nat Genet 6: 384–390

Motulsky AG (1995) Jewish diseases and origins. Nat Genet 9:99–101

Myrianthopoulos NC, Naylor AF, Aronson SM (1972) Founder effect in Tay-Sachs disease unlikely. Am J Hum Genet 24:341–342

Patael Y, Figer A, Gershoni-Baruch R, Papa MZ, Risel S, Shtoyerman-Chen R, Karasik A, et al (1999) Common origin of the I1307K APC polymorphism in Ashkenazi and non-Ashkenazi Jews. Eur J Hum Genet 7:555–559

Peleg L, Frisch A, Goldman B, Karpaty M, Narinsky R, Bronstein S, Frydman M (1998) Lower frequency of Gaucher disease carriers among Tay-Sachs disease carriers. Eur J Hum Genet 6:185–186

Peretz H, Mulai A, Usher S, Zivelin A, Segal A, Weisman Z, Mittelman M, et al (1997) The two common mutations causing factor XI deficiency in Jews stem from distinct founders: one of ancient Middle Eastern origin and another of more recent European origin. Blood 90:2654–2659

Rannala B, Slatkin M (1998) Likelihood analysis of disequilibrium mapping, and related problems. Am J Hum Genet 62:459–473

Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, et al (1995) Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. Nat Genet 9:152–159

Stephens JC, Reich DE, Goldstein DB, Shin HD, Smith MW, Carrington M, Winkler C, et al (1998) Dating the origin of the CCR5-Δ32 AIDS-resistance allele by the coalescence of haplotypes. Am J Hum Genet 62:1507–1515

Tayebi N, Cushner SR, Kleijer W, Lau EK, Damschroder-Williams PJ, Stubblefield BK, Den Hollander J, et al (1997) Prenatal lethality of a homozygous null mutation in the human glucocerebrosidase gene. Am J Med Genet 73:41–47

Weinryb BD (1972) The Jews of Poland: a social economic history of the Jewish community in Poland from 1100–1800. Jewish Publication Society of America, Philadelphia

Williams TN, Maitland K, Bennett S, Ganczakowski M, Peto TE, Newbold CI, Bowden DK, et al (1996) High incidence of malaria in alpha-thalassaemic children. Nature 383: 522–525

Zimran A, Gelbart T, Beutler E (1990) Linkage of the *Pvu*II polymorphism with the common Jewish mutation for Gaucher disease. Am J Hum Genet 46:902–905

Zivelin A, Rosenberg N, Faier S, Kornbrot N, Peretz H, Mannhalter C, Horellou MH, et al (1998) A single genetic origin for the common prothrombotic G20210A polymorphism in the prothrombin gene. Blood 92:1119–1124

Zlotogora J, Zeigler M, Bach G (1988) Selection in favor of lysosomal storage disorders? Am J Hum Genet 42:271–273

Zoossmann-Diskin A (1995) ITD in Ashkenazi Jews—genetic drift or selection? Nat Genet 11:13–15