Report

The Power to Detect Disease Associations with Mitochondrial DNA Haplogroups

David C. Samuels,[1] Andrew D. Carothers,[2] Robin Horton,[3] and Patrick F. Chinnery[3,4,5]

[1]Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA; [2]Medical Research Council Human Genetics Unit, Western General Hospital, Edinburgh; [3]School of Neurology, Neurobiology and Psychiatry, [4]Mitochondrial Research Group, and [5]Institute of Human Genetics, The University of Newcastle upon Tyne, Newcastle upon Tyne, United Kingdom

Genetic variation of mitochondrial DNA (mtDNA) has been linked to a number of multifactorial diseases, but there is currently no tool available to predict the optimal size for these investigations. We used a simulation-based (Monte Carlo) permutation test to generate power curves for European mtDNA haplogroup studies, to derive a universal equation to enable power calculations for prospective studies across the globe, and to show that very large cohorts are required to reliably detect an association with complex human diseases. In some populations, geographical variation in haplogroup frequencies will prevent the reliable detection of subtle haplogroup associations with uncommon disorders.

Human mtDNA codes for 13 essential polypeptide components of the mitochondrial respiratory chain that generates the principal source of intracellular energy, ATP (DiMauro and Schon 2003). mtDNA is inherited almost exclusively through the maternal line and is highly polymorphic. Human populations can be divided into several mtDNA haplogroups on the basis of specific SNPs scattered throughout the mitochondrial genome, reflecting mutations accumulated by a discrete maternal lineage. Each mtDNA haplogroup forms a mutually exclusive category (Torroni and Wallace 1994). Among Europeans, 95% of the population belongs to 1 of 10 haplogroups: H, I, J, K, M, T, U, V, W, and X (Torroni et al. 1996). Given the fundamental role of the mitochondrial genome in cellular metabolism, there have been a number of studies investigating the association between mtDNA lineages and multifactorial diseases and aging (e.g., De Benedictis et al. 1999; Ruiz-Pesini et al. 2000; Carrieri et al. 2001; Niemi et al. 2003; Mancuso et al. 2004; van der Walt et al. 2004). Despite initially promising results, often reaching high levels of statistical significance (α), it has rarely been possible to replicate the findings of the original reports. There are a number of explanations for the inconsistency, including difficulties in matching cases and controls and when there is a genuine variation in the size of a genetic effect in different populations. However, the consistent inability to repro-

duce the original result brings into question the power of individual studies.

The standard statistical method for analyzing haplogroup distributions is to use a $2 \times N_H$ contingency table and the χ^2 test, where N_H is the number of haplogroups. However, low counts in the less common haplogroups will tend to inflate the χ^2 value and lead to a false-positive result (Roff and Bentzen 1989). This can been tackled in one of two ways, either by performing Fisher's exact test for each haplogroup in turn, with an appropriate correction for multiple significance testing, or by grouping the less frequent haplogroups. Both approaches potentially mask a real difference between two haplotypes, particularly if they are uncommon in the general population. In addition, correcting for multiple significance testing is not straightforward, since each haplogroup frequency is dependent on the others. To address this problem, we developed a Monte Carlo permutation test to determine the exact type I error (false-positive rate) associated with a specific data set under study. (See the "Methods" section in appendix A.) This Monte Carlo permutation test has the advantage that it is not dependent on any prior assumptions about the distribution of the data or the number of haplogroups under study. We then compared the results of the Monte Carlo simulations with the results of a number of published studies.

In most cases, direct simulation generated a level of

Table 1

Comparison of Published mtDNA Haplogroup Association Studies with the Results of the Present Study

| | | DATA FROM PREVIOUSLY PUBLISHED STUDY | | | |
PHENOTYPE	SOURCE	No. of Controls	No. of Cases	α^a	EXACT α^b
Centenarian	De Benedictis et al. 1999	51	26	.017	.0226
Sperm motility (males only)[c]	Ruiz-Pesini et al. 2000	180	365	.0311	.1435
Alzheimer disease, with an *APOE ε4* allele[d]	Carrieri et al. 2001	119	94	.018	.0233
Dementia, with Lewy bodies	Chinnery et al. 2000	179	84	NR	.1347
Parkinson disease	van der Walt et al. 2004	340	609	NR	.1374
Longevity (>90 years):	Niemi et al. 2003				
General controls		400	225	.01	.011
Comparison with healthy infants		257	225	.00005	.0003
Amyotrophic lateral sclerosis	Mancuso et al. 2004	150	222	.016[e]	.0204
Parkinson disease	Pyle et al. 2005	447	455	<.001	<.0001

NOTE.—Our reanalysis of the original data in table 3 of Mancuso et al (2004) revealed a significant difference in the overall haplogroup distribution between the patients and controls ($\chi^2 = 18.54; \alpha = .016$). Five of the cells in the χ^2 table had expected counts of <5, thus potentially elevating the χ^2 value and leading to an inappropriately low α. This was confirmed by Monte Carlo simulation in which the exact α was much greater, .0204. Our subsequent power studies show that very large sample sizes are required to reliably detect a modest difference in haplogroup frequencies between two groups. This does not mean that it is impossible to detect a difference with a smaller study size, but it does mean that the chance of detecting such a difference is much lower.

[a] NR = not reported.

[b] Determined by Monte Carlo simulation; 10,000 iterations.

[c] Comparison of males with asthenozoospermia and males who were nonasthenozoospermic. Haplogroups K and U were merged in this study.

[d] Patients with Alzheimer disease who carry an *APOE ε4* allele compared with patients with Alzheimer disease who do not carry an *APOE ε4* allele,

[e] Mancuso et al (2004) report "the frequency of haplogroups H, J, K, T, U, V, W and X did not differ between the two groups" (p. 160).

significance (α) that was greater than the theoretical value based on an assumed null χ^2 distribution (table 1). For some studies, the simulated result was not significant by conventional criteria. As expected, these studies were thorough in determining the frequency of rare haplogroups (defined as <5% of controls). However, because of the limited study size, a number of haplotype groups contained fewer than five individuals. The simulations confirmed that, under these circumstances, the χ^2 value is artificially elevated, and tabulated values of χ^2 should

not be used to determine statistical significance. This problem can be avoided by applying Cochran's "rule of thumb" that no expected frequency be <1 and that ≤20% of the expected frequencies be <5 (Cochran 1954). If the results of a study do not comply with this rule, then direct simulation provides an unbiased and reliable alternative. This is often the case for mtDNA association studies, in which 6 of the 10 European haplogroups are each found in <10% of the population but account for 17% of the total (Torroni et al. 1996).
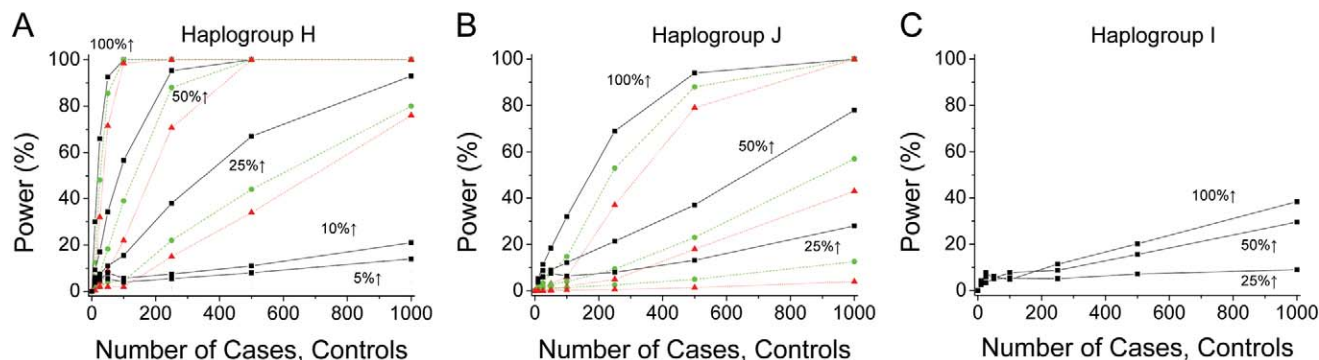

Figure 1 Examples of mtDNA haplogroup power curves determined by simulation—haplogroup H (*A*), haplogroup J (*B*), and haplogroup I (*C*)—on the basis of control data for the European population. The abscissa shows the number of cases and controls simulated (e.g., 400 = 400 cases and 400 controls). Black squares/solid line = .05 significance level; green circles/dashed line = .01 significance level; red triangles/dotted line = .001 significance level. The figures on the graph describe the percentage change in haplogroup frequency associated with a particular group of curves; for example, 100%↑ = haplogroup increases by 100%, with the remaining cases distributed evenly over the remaining nine haplogroups. For clarity, only the .05 level data is shown for haplogroup I.

We used a Monte Carlo simulation to determine the power to detect a difference between cases and controls at different levels of significance, on the basis of the known distribution of the 10 major European haplogroups (Torroni et al. 1996) (see fig. 1 for examples). We simulated both increases and decreases in the frequency of a each haplogroup, with the difference distributed proportionally between the remaining haplogroups, thus simulating a typical exploratory mtDNA haplogroup study for which there is no a priori assumption of an association with any one specific haplogroup. Figure 2A shows simulated data for five haplogroups, at $\alpha = .05$. As this plot shows, the power values do not have a simple relationship with

Table 2

Parameter Values for the Power, Used in Equation (A1)

Variable	$\alpha = .05$	$\alpha = .01$	$\alpha = .001$
A	-28 ± 5	-13 ± 2	-7 ± 1
x_0	2.6 ± 0.2	5.0 ± 0.1	7.4 ± 0.1
d	2.4 ± 0.1	2.5 ± 0.1	2.8 ± 0.1

NOTE.—These values can be used with equation (A2) to determine the minimum number of cases and controls required for a given power.
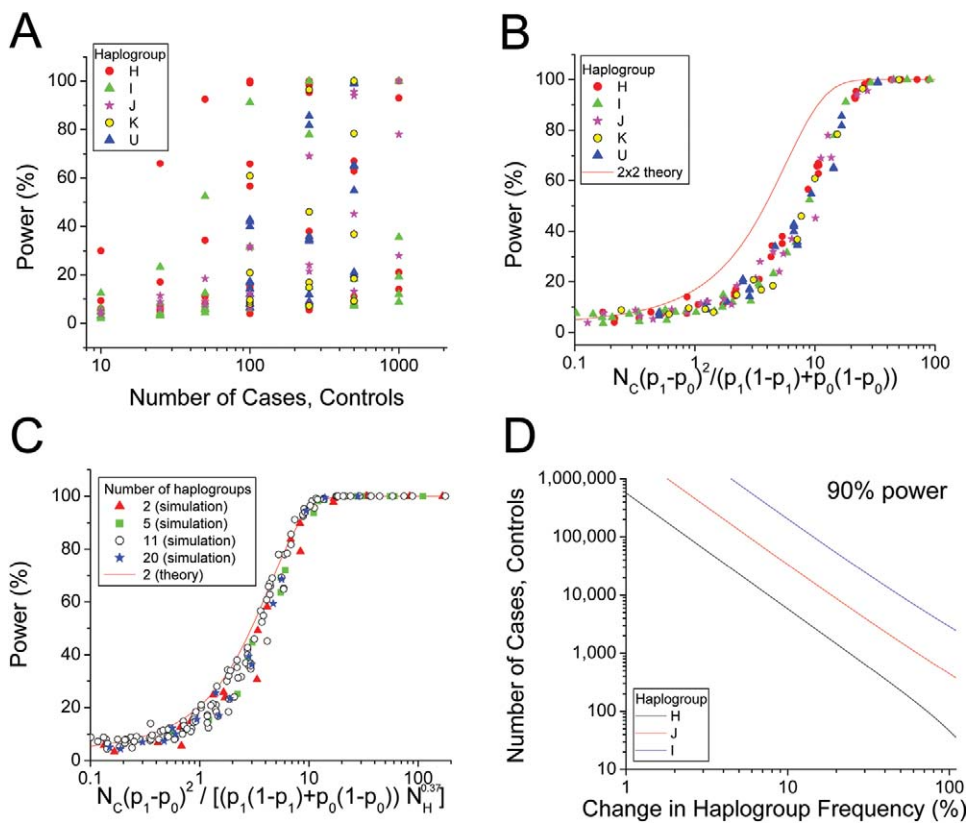

Figure 2   The power to detect an association between mtDNA haplogroups and disease. In all situations, the number of disease cases is assumed to be equal to the number of control subjects. *A,* Monte Carlo simulation data for the European haplogroups H, I, J, K, and U. Data from the individual simulations, including those shown in figure 1, are shown on the same graph for different changes in the percentage level of a particular haplogroup at the $\alpha = .05$ significance level. This demonstrates the scatter of the data points. *B,* These data collapse onto a simple sigmoid curve, with the X-axis scaling based on standard binomial theory. N_c is the number of disease cases (equal to the number of controls), p_0 is the frequency of the haplogroup in the control population, and p_1 is the frequency of the haplogroup in the cases. Data are shown for the significance level $\alpha = .05$ (for $\alpha = .01$ and $\alpha = .001$, see fig. A1). Both the data and the theoretical curve describe the same sigmoidal shape, with the European haplogroup simulation data (with the 10 major European haplogroups plus "others" being equivalent to a 2 × 11 table) shifted to the right. *C,* Simulations (symbols) and theoretical 2 × 2 curve (*red line*) for populations with different numbers of mutually exclusive haplogroups. The simulated subdivisions could correspond to superhaplogroups, haplogroup clusters, or any mutually exclusive sequence variants in any population. All of the data collapse onto a single curve when the X-axis is normalized by the number of haplogroups, N_H, raised to the power 0.37 (eq. [1]). Data are shown for the significance level $\alpha = .05$ (for $\alpha = .01$ and $\alpha = .001$, see figure A2). *D,* Example showing the number of cases and controls required to generate 90% power at the .05 significance level for a study of the 10 major European haplogroups ($N_H = 11$ to account for the <5% that do not fall into these 10 groups and are considered "others"). Haplogroup proportions in the control group are based on published values (haplogroup H = 0.41 in controls [*black line*]; I = .02 in controls [*blue line*]; J = .11 in controls [*red line*]) (Torroni et al. 1996).

Table 3

Scaled Number of Cases Required to Generate a Given Level of Power at a Given Level of Significance

POWER (%)	SCALED NO. OF CASES AND CONTROLS (N_{scaled})		
	$\alpha = .05$	$\alpha = .01$	$\alpha = .001$
50	3.6	5.5	7.8
75	6.0	8.1	10.7
90	8.5	10.8	13.8

NOTE.—The values are derived from a Bolzmann fit (eq. [A1]; $r^2 > 0.98$) of simulations shown in fig. 2C. The scaled number of cases is defined in equation (1). The scaled number of cases and controls required to generate a given power can be read from these curves or calculated from equation (A1) and equation (A2). For example, for a control population with N_H haplogroups, the number of cases and controls required to detect a change in the proportion of a given haplogroup from p_0 in controls to p_1 in the cases is given in equation (2). This approach was used to generate the example shown in figure 2D.

the raw number of cases and controls. In figure 2B, we normalize the X-axis, using a relationship derived from standard binomial theory (Armitage et al. 2002), and show that there is a simple relationship between the power and the normalized number of cases and controls. These results were compared with the theoretical curve for two haplogroups (fig. 2B for $\alpha = .05$; fig. A1 for $\alpha = .01$ and $\alpha = .001$). The curve described by the simulation data for the 10 European haplogroups was similar in shape to the theoretical 2×2 curve (corresponding to 2 haplogroups) but was displaced to the right. To explore the effect of different numbers of haplogroups on the position power curve, we then simulated other theoretical haplogroup distributions within the population (2, 5, or 20 discrete categories). These subdivisions could correspond to superhaplogroups, haplogroup clusters, or any mutually exclusive sequence variants in any population. The entire data set collapsed to a single universal curve when the X-axis was normalized by the number of haplogroups, N_H, raised to the power 0.37, by use of the expression

$$N_{scaled} = \frac{N_C(p_1 - p_0)^2}{[p_1(1 - p_1) + p_0(1 - p_0)]N_H^{0.37}}, \quad (1)$$

where N_c is the number of cases with a disease (assumed to be equal to the number of controls), p_0 is the frequency of the haplogroup in the control population, and p_1 is the frequency of the haplogroup in the disease group (fig. 2C for $\alpha = .05$; fig. A2 for $\alpha = .01$ and $\alpha = .001$). The power $\pm$ SD (0.37 ± 0.01) was determined by the standard approach of curve fitting. For the European haplogroup data studied here, $N_H = 11$, reflecting additional

category "other" for individuals not belonging to the 10 major haplogroups.

Equation (1) can be used to determine the minimum number of disease cases (N_{Cmin}) and control subjects required for a specific level of power,

$$N_{Cmin} = N_{scaled}N_H^{0.37}\left(\frac{p_1(1 - p_1) + p_0(1 - p_0)}{(p_1 - p_0)^2}\right), \quad (2)$$

by use of the scaling value, N_{scaled}, derived from the universal curves at each significance level α (table 2). Ninety percent power is achieved when N_{scaled} is 8.5 for $\alpha = .05$, 11 for $\alpha = .01$, and 14 for $\alpha = .001$, as derived from figure 2C (table 3). This approach can be used for any number of mutually exclusive genotypes in any population (fig. 3). Resolving equation (2) in terms of the odds ratio (OR) is complex, but equation (2) can be used by determining the proportion of cases that correspond to a specific OR for a given haplogroup, by use of the standard equation

$$OR = \frac{p_0/(1 - p_0)}{p_1/(1 - p_1)}. \quad (3)$$

The ORs, based on one published control data set (Torroni et al. 1996), corresponding to percentage changes in haplogroup frequency for the 10 European haplogroups are shown in tables A1 and A2. An example of
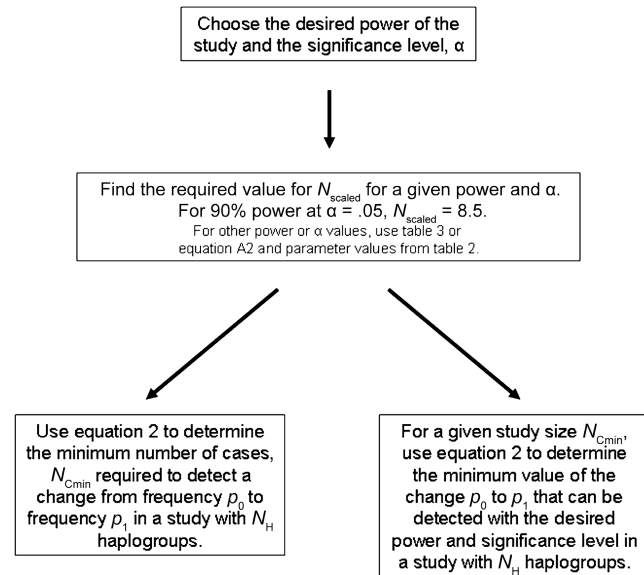

Figure 3   Algorithm to determine the power of an mtDNA haplogroup association study at a given level of statistical significance. The algorithm is based on the assumption that no specific haplogroup is suspected of being associated with a disease. α = statistical significance; N_H = the number of haplogroups (mutually exclusive genotypes) in a given study; p_0 = the frequency of the allele in control subjects; p_1 = the frequency of the allele in individuals with the disease; N_{Cmin} = the minimum number of cases to generate a given power.

this approach is given in figure 2*D*, with use of the European haplogroup distribution for 90% power at $\alpha = .05$. These results also show the limited power of moderate study sizes. With no prior reason to suspect an association with haplogroup H, ~6,000 cases and ~6,000 controls are required to form a study with 90% power to detect a 10% change in the frequency of haplogroup H in European populations (corresponding to an OR >1.18). Studies with only 500 cases and 500 controls would have only 90% power to detect >35% increase in the frequency of haplogroup H, corresponding to an OR >1.75. For less common haplogroups or more-subtle changes in haplogroup frequency, even greater sample sizes will be required (fig. 2*D*).

mtDNA is inherited almost exclusively down the maternal line and undergoes little, if any, intermolecular recombination. This makes the haplotype distribution in a given sample exquisitely sensitive to minor population genetic bottlenecks, which lead to differences in haplogroup frequency over short geographical distances. These factors increase the chance of detecting a spurious difference between cases and controls (false-positive or type I error). The standard approach to this problem is to consider the first study result as preliminary and to then generate a hypothesis that should be tested on an independent cohort. Here, we present a single equation that can be applied to mtDNA haplogroup studies around the globe, enabling a priori power calculations for exploratory or confirmatory studies. However, subtle changes in haplogroup frequency will require massive cohorts. In some populations, the frequency of mtDNA haplogroups in controls varies considerably over short geographic differences (Ghezzi et al. 2005), which makes it impossible to collect an adequately sized homogeneous study cohort, particularly for uncommon diseases. The same applies to population-genetics studies based on haplogroup comparisons between control populations across the globe, which have been used to deduce patterns of population migration (Mishmar et al. 2003; Ruiz-Pesini et al. 2004). These studies are equally likely to suffer from type I error, and they often show subtle differences in haplogroup frequency but require adequate sample sizes to demonstrate and confirm geographic differences in haplogroup distribution. An inability to reproduce the result with an adequately powered study would question the validity of the original result.

Acknowledgments

Appendix A

Methods

To determine whether there is significant heterogeneity in a given data set, the standard approach is to calculate χ^2, which is the sum of squares for the absolute difference between observed and expected values relative to the expected value, for each cell in a contingency table. For large sample sizes, statistical significance can be determined by comparing the calculated χ^2 value with tabulated values at specified degrees of freedom. However, when the number of observed counts is low, this approach is not reliable. Under these circumstances, the statistical significance can be determined by directly simulating the χ^2 distribution for a given data set. The Monte Carlo permutation test we present here is based on this principle and was adapted from the method of Roff and Bentzen (1989). For a given data set, the program generates a series of contingency tables by randomly assigning individual observations to each cell. The χ^2 value is then calculated for each simulated data set and is compared with the original χ^2 value. The probability of observing the original data set under the null hypothesis equates to the proportion of simulated χ^2 values greater than the χ^2 value for the original data set. Power simulations used the same algorithm, with simulation of different alterations in the haplogroup distribution for studies of different sizes. Subsequent analyses, curve fitting, and the generation of power curves were performed with Microcal Origin software (v. 6).

The power in figure 2*B* and 2*C* is also well described $\left(_{R^2 > 0.98} \right)$ by a Boltzmann equation,

$$\text{Power} = 100 + \frac{A - 100}{1 + e^{(N_{\text{scaled}} - x_0)/d}} , \qquad \text{(A1)}$$

which is easy to apply, where the power is expressed as a percentage and N_{scaled} is the scaled *X*-axis of figure 2, defined in equation (1). A, x_0, and d are parameters of the fit (values given in table 2) allowing the power to be derived for any haplogroup distribution, significance level, and study size. The Boltzmann equation

$$N_{\text{scaled}} = x_0 + d \ln \left(\frac{\text{Power} - A}{100 - \text{Power}} \right) \qquad \text{(A2)}$$

can be solved to give the value of N_{scaled} required to achieve a given power.

These equations are equally applicable to increases or decreases in the haplogroup frequency.
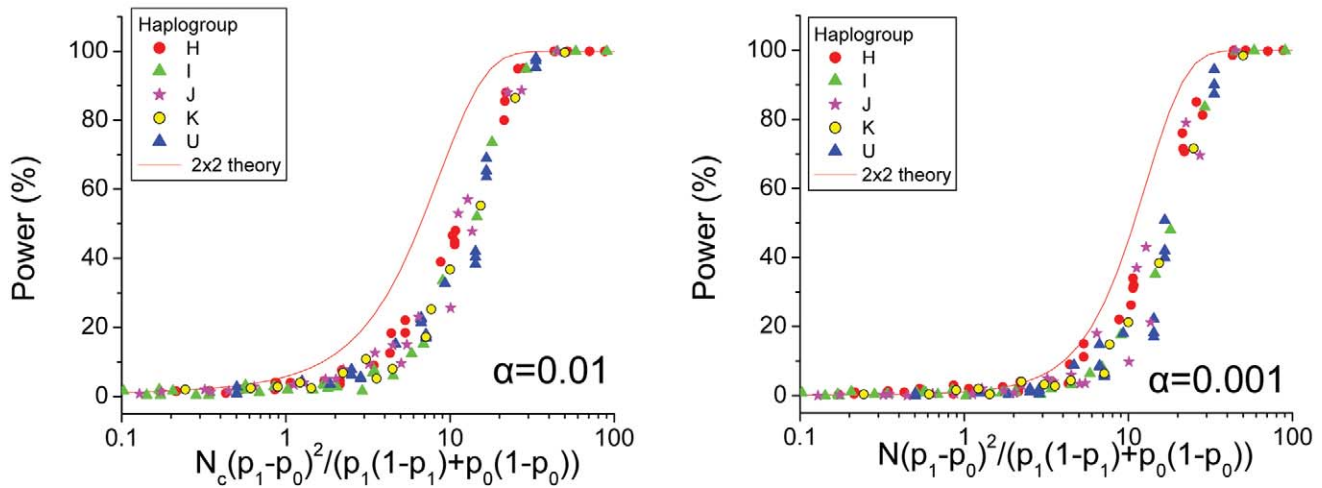
Figure A1 Monte Carlo simulation data for the European haplogroups H, I, J, K, and U. Data are from the individual simulations (fig. 1) and the theoretical curve for a 2 × 2 table based on binomial theory for the significance levels $\alpha = .01$ and $\alpha = .001$. These data describe a simple sigmoid curve, with the X-axis based on standard binomial theory. N_c is the number of disease cases (which is equal to the number of controls), p_0 is the frequency of the haplogroup in the control population, and p_1 is the frequency of the haplogroup in the cases. Both the data and the theoretical curve describe the same sigmoidal shape, with the European haplogroup simulation data (with the 10 major European haplogroups plus "others" being equivalent to a 2 × 11 table) shifted to the right.
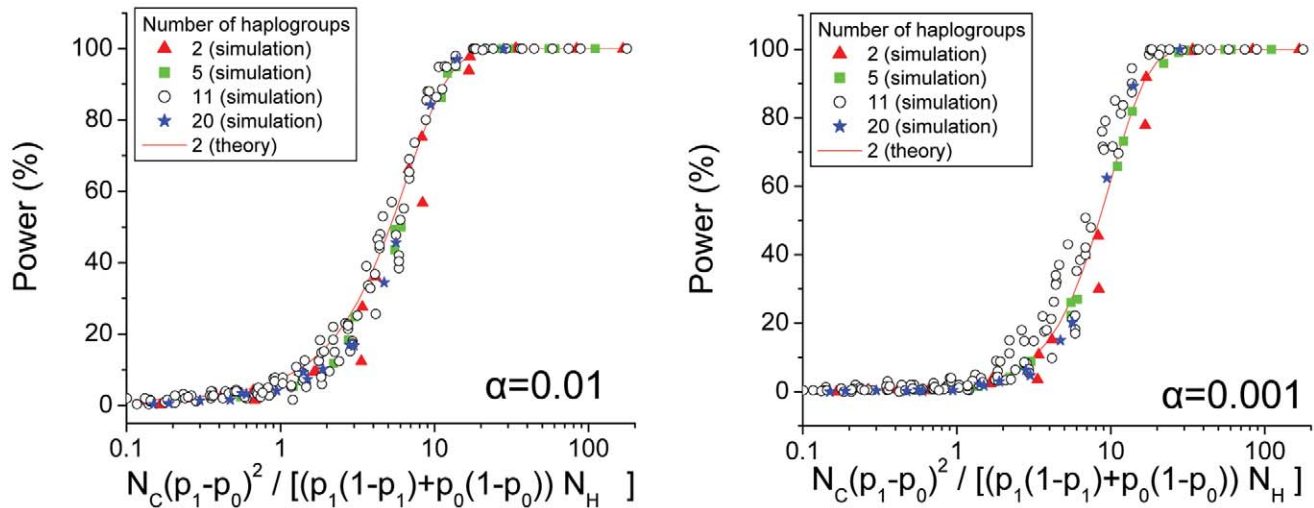
Figure A2 Simulations (*symbols*) and theoretical 2 × 2 curve (*red line*) for populations with different numbers of mutually exclusive haplogroups. The simulated subdivisions could correspond to superhaplogroups, haplogroup clusters, or any mutually exclusive sequence variants in any population. All of the data collapse onto a single curve when the X-axis is normalized by the number of haplogroups, N_H, raised to the power 0.37 (eq. [1]). Data are shown for significance level $\alpha = .01$ and $\alpha = .001$.

Table A1

ORs Corresponding to an Increase in Haplogroup Frequency for the 10 European Haplogroups

INCREASE IN HAPLOGROUP FREQUENCY (%)	OR BY HAPLOGROUP (FREQUENCY IN CONTROLS)									
	H (.41)	I (.02)	J (.11)	K (.08)	M (.01)	T (.13)	U (.15)	V (.03)	W (.02)	X (.02)
10	1.18	1.10	1.11	1.11	1.10	1.12	1.12	1.10	1.10	1.10
20	1.39	1.20	1.23	1.22	1.20	1.24	1.24	1.21	1.20	1.20
30	1.64	1.31	1.35	1.33	1.30	1.36	1.37	1.31	1.31	1.31
40	1.94	1.41	1.47	1.45	1.41	1.49	1.51	1.42	1.41	1.41
50	2.30	1.52	1.60	1.57	1.51	1.62	1.65	1.52	1.52	1.52
60	2.74	1.62	1.73	1.69	1.61	1.76	1.79	1.63	1.62	1.62
70	3.31	1.72	1.86	1.81	1.71	1.90	1.94	1.74	1.72	1.72
80	4.05	1.83	2.00	1.93	1.81	2.04	2.10	1.85	1.83	1.83
90	5.07	1.94	2.14	2.06	1.92	2.20	2.26	1.95	1.94	1.94
100	6.56	2.04	2.28	2.19	2.02	2.35	2.43	2.06	2.04	2.04
110	8.91	2.15	2.43	2.32	2.12	2.51	2.61	2.17	2.15	2.15
120	13.24	2.26	2.58	2.46	2.23	2.68	2.79	2.28	2.26	2.26
130	23.81	2.36	2.74	2.59	2.33	2.85	2.98	2.40	2.36	2.36
140	88.50	2.47	2.90	2.73	2.43	3.03	3.19	2.51	2.47	2.47
150	...	2.58	3.07	2.88	2.54	3.22	3.40	2.62	2.58	2.58
160	...	2.69	3.24	3.02	2.64	3.42	3.62	2.74	2.69	2.69
170	...	2.80	3.42	3.17	2.75	3.62	3.86	2.85	2.80	2.80
180	...	2.91	3.60	3.32	2.85	3.83	4.10	2.97	2.91	2.91
190	...	3.02	3.79	3.47	2.96	4.05	4.36	3.08	3.02	3.02
200	...	3.13	3.99	3.63	3.06	4.28	4.64	3.20	3.13	3.13
210	...	3.24	4.19	3.79	3.17	4.52	4.93	3.32	3.24	3.24
220	...	3.35	4.40	3.96	3.27	4.77	5.23	3.43	3.35	3.35
230	...	3.46	4.61	4.13	3.38	5.03	5.55	3.55	3.46	3.46
240	...	3.58	4.83	4.30	3.48	5.30	5.90	3.67	3.58	3.58
250	...	3.69	5.07	4.47	3.59	5.59	6.26	3.79	3.69	3.69
260	...	3.80	5.30	4.65	3.70	5.89	6.65	3.91	3.80	3.80
270	...	3.92	5.55	4.84	3.80	6.20	7.07	4.04	3.92	3.92
280	...	4.03	5.81	5.02	3.91	6.53	7.51	4.16	4.03	4.03
290	...	4.15	6.08	5.22	4.02	6.88	7.99	4.28	4.15	4.15
300	...	4.26	6.36	5.41	4.13	7.25	8.50	4.41	4.26	4.26
310	...	4.38	6.65	5.61	4.23	7.64	9.05	4.53	4.38	4.38
320	...	4.49	6.95	5.82	4.34	8.05	9.65	4.66	4.49	4.49
330	...	4.61	7.26	6.03	4.45	8.48	10.30	4.79	4.61	4.61
340	...	4.73	7.59	6.25	4.56	8.94	11.00	4.92	4.73	4.73
350	...	4.85	7.93	6.47	4.66	9.43	11.77	5.05	4.85	4.85
360	...	4.96	8.29	6.70	4.77	9.96	12.61	5.18	4.96	4.96
370	...	5.08	8.66	6.93	4.88	10.51	13.54	5.31	5.08	5.08
380	...	5.20	9.05	7.17	4.99	11.11	14.57	5.44	5.20	5.20
390	...	5.32	9.46	7.41	5.10	11.74	15.72	5.57	5.32	5.32
400	...	5.44	9.89	7.67	5.21	12.43	17.00	5.71	5.44	5.44

NOTE.—Values are derived from equation (3) and a published data set (Torroni et al. 1996). These values will differ for control groups with a different background-haplogroup frequency.

Table A2

ORs Corresponding to a Decrease In Haplogroup Frequency for the 10 European Haplogroups

DECREASE IN HAPLOGROUP FREQUENCY (%)	OR BY HAPLOGROUP (FREQUENCY IN CONTROLS)									
	H (.41)	I (.02)	J (.11)	K (.08)	M (.01)	T (.13)	U (.15)	V (.03)	W (.02)	X (.02)
10	.84	.90	.89	.89	.90	.89	.88	.90	.90	.90
20	.70	.80	.78	.79	.80	.78	.77	.80	.80	.80
30	.58	.70	.67	.68	.70	.67	.66	.69	.70	.70
40	.47	.60	.57	.58	.60	.57	.56	.59	.60	.60
50	.37	.49	.47	.48	.50	.47	.46	.49	.49	.49
60	.28	.40	.37	.38	.40	.37	.36	.39	.40	.40
70	.20	.30	.28	.28	.30	.27	.27	.29	.30	.30
80	.13	.20	.18	.19	.20	.18	.18	.20	.20	.20
90	.06	.10	.09	.09	.10	.09	.09	.10	.10	.10

NOTE.—Values derived from equation (3) and a published data set (Torroni et al. 1996). These values will differ for control groups with a different background-haplogroup frequency.

References

Armitage P, Berry G, Matthews JNS (2002) Statistical methods in medical research. Blackwell, Oxford, United Kingdom

Carrieri G, Bonafe M, De Luca M, Rose G, Varcasia O, Bruni A, Maletta R, Nicmias B, Sorbi S, Corsonello F, Feraco E, Andreev K, Yashin AI, Francheschi C, De Benedictis G (2001) Mitochondrial DNA haplogroups and APOE4 allele are not independent variables in sporadic Alzheimer's disease. Hum Genet 108:194–198

Chinnery PF, Taylor G, Howell N, Andrews RM, Morris CM, McKeith IG, Perry RH, Edwardson JA, Turnbull DM (2000) Mitochondrial DNA haplogroups and susceptibility to AD and dementia with Lewy bodies. Neurology 55:302–304

Cochran WG (1954) Some methods for strengthening the common χ^2 tests. Biometrics 10:417–451

De Benedictis G, Rose G, Carrieri G, De Luca M, Falcone E, Passarino G, Bonafe M, Monti D, Baggio G, Bertolini S, Mari D, Mattace R, Franceschi C (1999) Mitochondrial DNA inherited variants are associated with successful aging and longevity in humans. FASEB J 13:1532–1536

DiMauro S, Schon EA (2003) Mitochondrial respiratory-chain diseases. N Engl J Med 348:2656–2668

Ghezzi D, Marelli C, Achilli A, Goldwurm S, Pezzoli G, Barone P, Pellecchia MT, Stanzione P, Brusa L, Bentivoglio AR, Bonuccelli U, Petrozzi L, Abbruzzese G, Marchese R, Cortelli P, Grimaldi D, Martinelli P, Ferrarese C, Garavaglia B, Sangiorgi S, Carelli V, Torroni A, Albanese A, Zeviani M (2005) Mitochondrial DNA haplogroup K is associated with a lower risk of Parkinson's disease in Italians. Eur J Hum Genet 13:748–752

Mancuso M, Conforti FL, Rocchi A, Tessitore A, Muglia M, Tedeschi G, Panza D, et al (2004) Could mitochondrial haplogroups play a role in sporadic amyotrophic lateral sclerosis? Neurosci Lett 371:158–162

Mishmar D, Ruiz-Pesini E, Golik P, Macaulay V, Clark AG, Hosseini S, Brandon M, Easley K, Chen E, Brown MD, Sukernik RI, Olckers A, Wallace DC (2003) Natural selection shaped regional mtDNA variation in humans. Proc Natl Acad Sci USA 100:171–176

Niemi AK, Hervonen A, Hurme M, Karhunen PJ, Jylha M, Majamaa K (2003) Mitochondrial DNA polymorphisms associated with longevity in a Finnish population. Hum Genet 112:29–33

Pyle A, Foltynie T, Tiangyou W, Lambert C, Keers SM, Allcock LM, Davison J, Lewis SJ, Perry RH, Barker R, Burn DJ, Chinnery PF (2005) Mitochondrial DNA haplogroup cluster UKJT reduces the risk of PD. Ann Neurol 57:564–567

Roff DA, Bentzen P (1989) The statistical analysis of mitochondrial DNA polymorphisms: chi 2 and the problem of small samples. Mol Biol Evol 6:539–545

Ruiz-Pesini E, Lapeña A-C, Díez-Sánchez C, Pérez-Martos A, Montoya J, Alvarez E, Díaz M, Urriés A, Montoro L, López-Pérez MJ, Enriquez JA (2000) Human mtDNA haplogroups associated with a high or reduced spermatozoa motility. Am J Hum Genet 67:682–696

Ruiz-Pesini E, Mishmar D, Brandon M, Procaccio V, Wallace DC (2004) Effects of purifying and adaptive selection on regional variation in human mtDNA. Science 303:223–226

Torroni A, Huoponen K, Francalacci P, Petrozzi M, Morelli L, Scozzari R, Obinu D, Savontaus ML, Wallace DC (1996) Classification of European mtDNAs from an analysis of three European populations. Genetics 144:1835–1850

Torroni A, Wallace DC (1994) Mitochondrial DNA variation in human populations and implications for detection of mitochondrial DNA mutations of pathological significance. J Bioenerg Biomembr 26:261–271

van der Walt JM, Dementieva YA, Martin ER, Scott WK, Nicodemus KK, Kroner CC, Welsh-Bohmer KA, Saunders AM, Roses AD, Small GW, Schmechel DE, Murali Doraiswamy P, Gilbert JR, Haines JL, Vance JM, Pericak-Vance MA (2004) Analysis of European mitochondrial haplogroups with Alzheimer disease risk. Neurosci Lett 365:28–32