Mitochondrial DNA variant associated with Leber hereditary optic neuropathy and high-altitude Tibetans

Fuyun Ji^a, Mark S. Sharpley^{b,c}, Olga Derbeneva^{b,c}, Leonardo Scherer Alves^{b,c}, Pin Qian^d, Yaoli Wang^a, Dimitra Chalkia^{b,c}, Maria Lvova^{b,c}, Jiancheng Xu^a, Wei Yao^a, Mariella Simon^c, Julia Platt^c, Shiqin Xu^c, Alessia Angelin^{b,c}, Antonio Davila^{b,c}, Taosheng Huang^c, Ping H. Wang^e, Lee-Ming Chuang^f, Lorna G. Moore^g, Guisheng Qian^a, and Douglas C. Wallace^{b,c,1}

^aInstitute of Human Respiratory Disease, ^dInstitute of Field Internal Medicine, Xinqiao Hospital, Third Military Medical University, Chongqing 400037, China; ^bCenter for Mitochondrial and Epigenomic Medicine, Children's Hospital of Philadelphia, University of Pennsylvania, Philadelphia, PA 19104; ^cCenter for Molecular and Mitochondrial Medicine and Genetics, and ^eDepartment of Medicine, Center for Diabetes Research and Treatment, University of California, Irvine, CA 92697; ^fDepartment of Internal Medicine, National Taiwan University Hospital, Taipei 100, Taiwan; and ^gGraduate School of Arts and Sciences, Wake Forest University, Winston-Salem, NC 27157

Contributed by Douglas C. Wallace, March 15, 2012 (sent for review October 19, 2011)

The distinction between mild pathogenic mtDNA mutations and population polymorphisms can be ambiguous because both are homoplasmic, alter conserved functions, and correlate with disease. One possible explanation for this ambiguity is that the same variant may have different consequences in different contexts. The NADH dehydrogenase subunit 1 (ND1) nucleotide 3394 T > C (Y30H) variant is such a case. This variant has been associated with Leber hereditary optic neuropathy and it reduces complex I activity and cellular respiration between 7% and 28% on the Asian B4c and F1 haplogroup backgrounds. However, complex I activity between B4c and F1 mtDNAs, which harbor the common 3394T allele, can also differ by 30%. In Asia, the 3394C variant is most commonly associated with the M9 haplogroup, which is rare at low elevations but increases in frequency with elevation to an average of 25% of the Tibetan mtDNAs (odds ratio = 23.7). In high-altitude Tibetan and Indian populations, the 3394C variant occurs on five different macrohaplogroup M haplogroup backgrounds and is enriched on the M9 background in Tibet and the C4a4 background on the Indian Deccan Plateau (odds ratio = 21.9). When present on the M9 background, the 3394C variant is associated with a complex I activity that is equal to or higher than that of the 3394T variant on the B4c and F1 backgrounds. Hence, the 3394C variant can either be deleterious or beneficial depending on its haplogroup and environmental context. Thus, this mtDNA variant fulfills the criteria for a common variant that predisposes to a "complex" disease.

adaptation | hypoxia | OXPHOS | complex disease

The line demarcating pathogenic mtDNA polypeptide mutations from benign population polymorphisms has become increasingly blurred (1, 2). Classic pathogenic mtDNA mutations are recent, heteroplasmic, change highly conserved amino acids, and cause biochemical changes in mitochondrial functions in trans-mitochondrial cybrids (3, 4). However, population-specific groups of related mtDNA haplotypes (haplogroups) are commonly founded by amino acid substitution variants that change amino acids that can be conserved throughout much of the animal kingdom. This anomaly has led to the hypothesis that certain mtDNA variants are adaptive in specific environments and become enriched by natural selection. However, these same mtDNA haplogroups can be associated with various degenerative diseases in other contexts (5–7).

The first inherited mtDNA disease mutation was a missense mutation in the mtDNA ND4 gene at nucleotide 11778 G > A arginine 340 to histidine (R340H) that causes the midlife blindness Leber hereditary optic neuropathy (LHON) (8). Subsequently, several mtDNA mutations were found to cause LHON in Europeans: NADH dehydrogenase subunit 1 (ND1) G3460A (A52T) (9), ND6 T14484C (M64V) (10), and ND6 G14459A (A72V) (11). The ND6 14459A mutation generates the most severe complex I

defect (12) and causes LHON when heteroplasmic but dystonia when homoplasmic (11, 13). The remaining mutations cause milder complex I defects (14–18) and cause LHON when near homoplasmic (8–10). For the milder LHON mutations (11778A, 3460A, 14484C) the penetrance of blindness is increased when the mutations arise on mtDNA haplogroup J (19–22).

In addition to the common European LHON variants (11778A, 3460A, 14484C, and 14459A), a large number of rarer mtDNA variants have also been associated with LHON, either together with the 3460, 11778, and 14484 mutations or in isolation. One such variant is a T > C mutation found at nucleotide 3394 in the ND1 gene, which changes an evolutionarily conserved tyrosine (Y) 30 to histidine (H) (Y30H). This variant was first reported associated with LHON in a survey of 47 European LHON pedigrees, where four patients harbored the 3394C variant, 8.5% of LHON pedigrees vs. 1.1% of controls. However, one of the 3394C patients also harbored the 11778A mutation and the other three the 14484C mutation, all on haplogroup J (21). The 3394C variant was found in another European pedigree in association with the 14484C mutation and haplogroup J (23).

In Chinese studies, the 3394C variant was reported in four pedigrees in association with the 11778A mutation and haplogroup M9a (24). However, in four other Chinese LHON patients the 3394C variant was the only potentially pathogenic mutation, occurring on haplogroups D4b and M9a (25). Therefore, the 3394C variant lies at the interface between a pathogenic and a polymorphic variant and permits addressing the question: What level of mitochondrial complex I defect is necessary to produce LHON?

The mtDNA has a high mutation rate, but the most deleterious mtDNA mutations are removed by intraovarian selection (26, 27). As a result, mtDNA variants, which mildly perturb mitochondrial bioenergetics, are continuously being introduced into the population. Although some of these variants cause significant mitochondrial defects and disease (1), milder variants can increase in frequency in certain populations, suggesting that they may be advantageous in particular local environments. Adaptive mtDNA missense mutations should change a functionally important amino

Author contributions: F.J., J.X., L.G.M., G.Q., and D.C.W. designed research; F.J., M.S.S., O.D., L.S.A., P.Q., D.C., M.L., and A.A. performed research; F.J., Y.W., W.Y., M.S., J.P., S.X., A.D., T.H., P.H.W., L.-M.C., L.G.M., and D.C.W. contributed new reagents/analytic tools; F.J., M.S.S., O.D., L.S.A., P.Q., D.C., M.L., and A.A. analyzed data; and D.C.W. wrote the paper.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. [FJ748704](https://www.ncbi.nlm.nih.gov/nuclot/FJ748704)–[FJ748759](https://www.ncbi.nlm.nih.gov/nuclot/FJ748759)).

¹To whom correspondence should be addressed. E-mail: wallaced1@email.chop.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1202484109/-DCSupplemental.

acid, be enriched in a particular population, appear on several different mtDNA backgrounds, initiate mtDNA haplogroup lineages in particular environmental contexts, correlate in frequency with the intensity of the selective environmental condition, and significantly affect mitochondrial function. If the mutation is also deleterious in another context, then the variant would have been enriched in opposition to purifying selection, which would be unlikely to occur by genetic drift. Rather, the environment must have nullified the variant's adverse effects (5–7).

In Asia, the 3394C variant appears to be such a context-specific deleterious or adaptive mtDNA variant. Although it has been reported to cause LHON in low-altitude populations, it is enriched in high-altitude Tibetan and Indian populations, particularly in association with macrohaplogroup M haplogroups M9 and C4a4, respectively. Furthermore, the 3394C variant is associated with reduced complex I activity on B4c and F1 haplogroups but high complex I activity on haplogroup M9. Therefore, the 3394C variant appears to be deleterious in one mtDNA and environmental context but beneficial in another.

Results

Biochemical Alterations Associated with the T3394C Variant. The ND1 nucleotide T3394C Y30H variant changes an amino acid that is conserved in 99% (939 of 946) of mammalian mtDNAs and thus should be functionally important. To determine if there is a biochemical defect associated with the 3394C variant, we screened 390 Southern California Han Chinese to identify individuals whose mtDNAs harbored the 3394T and the 3394C variants on the same haplogroup background and found both alleles associated with the B4c and F1 haplogroups (Table 1 and Table S1). We then transferred the 3394T and 3394C mtDNAs on to a uniform nuclear DNA (nDNA) background by fusion of platelets from the subjects to 143B(TK⁻) ρ^0 cells (4). Cybrid fusions were performed for two subjects with the 3394C B4c mtDNA, two

subjects with the 3394T B4c mtDNA, one with the 3394C F1 mtDNA, and two with the 3394T F1 mtDNA. The mtDNAs from these subjects and a number of their transmittochondrial cybrid clones were completely sequenced (Tables S2–S4).

To detect the relatively subtle complex I defect expected for the ND1 T3394C mutation, we developed a highly precise complex I assay using alamethicin permeabilized mitochondria. To test the precision of this system, we analyzed the complex I activity of one of the B4c 3394T cybrid clones, which we found had acquired a de novo missense mutation in the ND1 gene at nucleotide T3515C (L70P), which disrupts a transmembrane helix (Table S3). Because the 3515T vs. 3515C clones were derived from the same B4c 3394T subject, their mtDNA sequences were otherwise identical. The nucleotide 3515C (L70P) missense mutation reduced the complex I specific activity by 65% ($P < 0.0001$). The extremely low P value indicates the ultra-low variance around the mean of the multiple determinations made for the two cell lines. This finding confirms the reliability of both our cybrid technique and complex I assay (Fig. 1A).

Applying this analysis to the B4c subjects revealed that the 3394C variant was associated with a 28% reduction in complex I activity relative to the 3394T variant ($P = 0.0006$). On the F1 background, the 3394C variant was associated with a 15% reduction in complex I activity ($P = 0.047$) (Fig. 1A). When measuring cellular O₂ consumption using the Oroboros respirometer and site I substrates (Fig. 1B), the basal respiration rate associated with the 3394C B4c mtDNAs was 17.3% less than the 3394T B4c rate ($P = 0.04$) and the uncoupled rate was reduced 7.3% ($P = 0.016$). Similarly, the basal rate of the 3394C F1 mtDNAs was 12.3% less than the 3394T F1 mtDNAs ($P = 0.092$), but the uncoupled rate was reduced 16.5% ($P = 0.006$) (Fig. 1B). Therefore, 3394C reduces both complex I activity and NADH-linked respiration, thus confirming that it can contribute to the pathophysiology of LHON.

Table 1. Haplogroup distribution of Tibetan and Han Chinese samples collected for this study

Haplogroup	Tibet sample 1 (n = 216)	Tibet sample 2 (n = 73)	Southwest China (n = 137)	Taiwan (n = 1,117)	Southern California Chinese (n = 390)
A	11.57	15.28	7.30	3.58	6.92
B	4.17	4.17	18.98	17.91	19.23
C	5.56	2.78	2.92		3.08
D	16.67	26.39	29.20	24.26	25.38
E				1.07	
F	9.26	9.72	21.90	17.55	18.72
G	11.57	9.72	2.92	2.15	2.82
H					0.26
M		8.33		3.13	0.51
M2	0.46		0.73		
M21				0.09	
M28				0.09	
M7			5.83	11.81	7.18
M8			2.19	6.45	3.59
M9	27.31	19.44	2.92	1.43	2.05
M10			0.73	1.52	0.51
M11	1.85				
M13	4.63	2.78			
M14	3.24				
M15	1.39				
M33					0.51
N		1.39			
N9a			1.46	6.18	5.13
R			0.73	2.69	1.03
U1	0.93				
Y				0.09	0.26
Z	1.39		2.19		2.82

Fig. 1. Alternations in complex I-specific activity and the cybrid physiology associated with the 3394C variant. (A) NADH: coenzyme q_1 oxidoreductase activity of mitochondria isolated from B4c, F1, and M9 haplogroup cybrids. The number of independent determinations for each sample is: B4c-T 3515C ($n = 6$); B4c-C ($n = 19$); B4c-T ($n = 21$); F1-C ($n = 5$); F1-T ($n = 5$); M9-C ($n = 11$). (B) The cellular respiration rates (basal and uncoupled) of the B4c, F1, and M9 cybrid clones. The number of independent determinations for each sample is: B4c-C: ($n = 4$); B4c-T ($n = 6$); F1-C ($n = 5$); F1-T ($n = 5$); and M9-C ($n = 9$). On the x axis, the letters next to each haplogroup represent the 3394 variants, C: 3394C and T: 3394T. The error bars represent the SEM.

Association of 3394C Variant with Haplogroup M9 and Their Enrichment in Tibetans. Although the above studies showed that the 3394C variant can inhibit complex I activity in Southern California Han Chinese, haplogroup M9, which contains the 3394C variant, was found in 2.2% of the population (8 of 390) (Table 1). This result contrasts with a frequency of the 3394C allele associated with F1 and B4 haplogroups of 0.1–0.3% (Table S1). This finding raised the question: Why should the 3394C variant be maintained in the Chinese population on haplogroup M9 at polymorphic frequencies? The answer to this question unexpectedly came from our Chinese population studies, which revealed that M9 with the 3394C variant is strongly enriched in Tibetans.

Analysis of Tibetan Haplogroup Distribution. In two independent Tibetan expeditions, we collected samples from six Tibetan villages located above 3,000-m elevation. Expedition sample 1 recruited 216 individuals and expedition sample 2, the earlier of the two, recruited 73 individuals (28) (Table 1). These high-altitude samples were compared with three low-altitude Chinese samples: the 390 Han Chinese living in Southern California, 137 subjects from southwestern (SW) China, and 1,117 from Taiwan (Table 1). The mtDNA haplogroups of the Tibetan and low-altitude Chinese samples were determined by analyzing restriction fragment-length polymorphisms (RFLP), screening for additional informative variants, sequencing the hypervariable section I (HVS I) of the mtDNA control region, and sequencing the entire mtDNAs of representative mtDNAs (see *SI Materials and Methods*).

Tibetan sample 1 mtDNAs. Tibetan sample 1 was found to encompass 14 haplogroups (Fig. 2A, Table 1, and Tables S5 and S6). All of these haplogroups have previously been reported in Asians, with the exception of mtDNA haplogroups M14, M15, and U1.

All Asian haplogroups can be grouped into two Asian macrohaplogroups, M and N. Tibetan sample 1 macrohaplogroup N

Fig. 2. Distribution of mtDNA haplogroups in Tibetan and low-land Chinese populations. (A) Tibetan samples 1 and 2, $n = 289$. (B) SW Chinese, Taiwanese, and southern California Chinese samples, $n = 1,644$. Haplogroups H, M2, M21, M28, M33, N, U1, and Y with percentages <0.7 in both populations were excluded for visual clarity.

mtDNAs include haplogroups A, B, F, and U (26% of subjects); macrohaplogroup M mtDNAs included C, D, G, M8, M9, M13, and the rare M haplogroups M2, M11, M14, and M15 (74% of subjects) (see *SI Results*).

Tibetan sample 2 mtDNAs. Tibetan sample 2 (28) was found to encompass 10 major haplogroups (Fig. 2A, Table 1, and Tables S7 and S8). Of these haplogroups, about 30% were N and 70% were M.

Analysis of Low-Altitude Asian mtDNA Haplogroups. The Southern California Han Chinese sample encompassed 17 major haplogroups (Fig. 2B and Table 1) and 25 lineages (Table S1). The SW China Han Chinese sample encompassed 14 mtDNA haplogroups (Fig. 2B, Table 1, and Tables S9 and S10) and the Taiwan Han Chinese sample encompassed 13 mtDNA haplogroups and sub-haplogroups (Fig. 2B and Table 1), all of which have been previously described (29–33). The average macrohaplogroup distribution of the low-altitude populations was N = 49% and M = 51% (Table 1).

Altitude Association of the M9 Haplogroup and the T3394C Variant. Comparison of the haplogroup frequency distribution of the two high-altitude-adapted Tibetan samples with the three low-altitude-adapted Asian samples revealed a striking increase in frequency the Tibetan haplogroup M9. The average M9 frequency of the low-altitude Asian populations was 1.7% (range 1.4–2.9%). In contrast, the M9 frequency of our two Tibetan populations was 25.3% (Fig. 2 and Table 1). Thus, there is a highly significant association between M9 and Tibetans, with the odds ratio (OR) of observing the M9 haplogroup in Tibetan sample 1 being 12.5-times higher than that for SW China sample, and OR being 19.5-times higher between the two Tibetan samples and three low-altitude Asian populations (Pearson χ^2 test, $P < 0.001$) (Table 2 and Fig. S14).

An independent survey of 680 Tibetan mtDNAs from 1,500 m and up (34) also revealed an elevated frequency of M9 of 16.4%. Comparison of the combined three Tibetan samples [samples 1 + 2 + Zhao et al. (34)] vs. all non-Tibetan samples both examined here and in published reports (mean M9 frequency of 0.6%) gave an OR = 23.7 (Table 2 and Fig. S14). Hence, the M9 mtDNA haplogroup is significantly enriched in high-altitude Tibetan populations.

In addition to the defining M9 variant in ND2 nucleotide G4491A (V8I), sequence analysis of six Tibetan M9 mtDNAs revealed that all contained the ND1 3394C (Y30H) variant. Further screening of Tibetan mtDNAs for the 3394C variant

Table 2. Statistics for 2 × 2 contingency tables of the variables: Mitochondrial haplogroup (M9, non-M9) or variant 3394 (C, T) and altitude (high, low)

Variable tested	Dataset*	Sample size	Pearson χ^2 -value	P value		OR	95% Confidence intervals	
				Asymptotic	Exact			
Haplogroup	1	353	34.0	<0.0001	7.39 E-09	12.5	4.42	35.3
	2	1,929	274.7	<0.0001	2.71 E-41	19.5	12.3	30.8
	3	11,302	1173.3	<0.0001	4.93 E-129	23.7	18.4	30.5
Variant 3394	1	353	35.77	<0.0001	1.43 E-09	13.1	4.6	36.9
	2	1,929	268.48	<0.0001	1.27 E-40	17.9	11.5	27.8
	3	11,302	896.7	<0.0001	5.66 E-106	14.7	11.7	18.4
	4	11,302	1004.8	<0.0001	4.83 E-117	16.54	13.2	20.7
	5	11,302	1272.8	<0.0001	2.76 E-144	21.88	17.4	27.6

*Dataset 1: Tibetan sample 1 vs. SW China Han; Dataset 2: Tibetan samples 1 + 2 vs. SW China Han, Taiwan Han, and southern California Han; Dataset 3: Tibetan samples 1 + 2 + Zhao et al. (34) Tibetans vs. SW China Han, Taiwan Han, southern California Han, and NCBI complete mtDNA sequences (last updated 05/17/2010), excluding five M9 sequences from Tibet (34); Dataset 4: Tibetan samples 1 + 2 + Zhao et al. (34) Tibetans + 10 C4a samples from Daccan plateau in southern India vs. SW China Han, Taiwan Han, southern California Han, and NCBI complete mtDNA sequences (last updated 05/17/2010), excluding five M9 sequences from Tibet (34); Dataset 5: Tibetan samples 1 + 2 + Zhao et al. (34) Tibetans + 10 C4a samples from Daccan plateau in southern India + two C4a samples from northern India, + 21 M9 samples from northern India vs. SW China Han, Taiwan Han, southern California Han, and NCBI complete mtDNA sequences (last updated 05/17/2010), excluding five M9 sequences from Tibet (34).

using the associated HaeIII restriction-site variant revealed that all of the M9 mtDNAs harbored the 3394C, as did one of the 27 D4 and one of the seven G3b mtDNAs from Tibetan sample 1 (Fig. 3A and Tables S6 and S11). The OR of finding the 3394C variant among Tibetan samples 1 and 2 vs. our three low-altitude populations was 17.9-times higher and that for all three Tibetan samples [1, 2 and Zhao et al. (34)] vs. the three Han Chinese samples was 14.7 (Table 2 and Fig. S1B).

An independent study of Indian Tibeto-Burman linguist populations revealed the 3394C variant in 21 M9, 1 M11a, and 12 C4a4 mtDNAs. All of the M9, the one M11a, and two of the C4a4 3394C mtDNAs were observed in several tribes living in Northeastern India adjacent to Tibet. The remaining 10 C4a4 3394C mtDNAs were found in the Jenu Kuruba tribe living on the Deccan plateau at 1500–1750 m altitude (35) (Table S11). Combining these results with ours revealed that the 3394C variant was observed in populations at 1,500 m or above on five different mtDNA haplogroups, and that all of the M9 haplogroup mtDNAs in the Himalaya and 83% of the C4a4 haplogroup mtDNAs in the Deccan plateau harbor the 3394C variant. Thus, the 3394C variant has arisen multiple times and been enriched on two different mtDNA backgrounds in two independent, higher-altitude populations. Combining all available data revealed that above 1,500 m the 3394C variant is 21.9-times more likely at high altitude than low (Table 2 and Fig. S1B).

To further verify that the M9 haplogroup and the 3394C variant were enriched by altitude, we calculated the frequencies of M9 and 3394C for each Tibetan village observed in our samples 1 and 2 and by Zhao et al. (34). This finding revealed that the frequency of both the M9 haplogroup ($R^2 = 0.40$) and the 3394C variant ($R^2 = 0.31$) increased as the altitude of the village sampled increased (Fig. 3B and C and Table S12).

Although the ND1 Y30 amino acid is highly conserved among mammals, it is variant among *Macaca mulatta* and *Macaca sylvanus* (Y30H), consistent with *M. mulatta* having formerly been found in southern China and Tibet at elevations up to 3,000 m (36). Similarly, *Theropithecus gelada* is Y30C and this primate lives in the Ethiopia and Eritrea highlands between 2,000 and 5,000 m. Hence, variation in ND1 Y30 can be associated with high altitude in other primates as well as humans (Table S13).

Biochemical Consequences of the 3394C Variant on the M9 Haplogroup.

To determine why the 3394C variant might be enriched in Tibetans on the M9 haplogroup, we prepared transmittochondrial cybrids

from southern Californian Han Chinese harboring M9-3394C mtDNAs (Table S2). Surprisingly, the complex I-specific activity and basal NADH-linked respiration of the M9 cybrids with the 3394C allele were much higher than those activities of either the B4c or the F1 cybrids with the 3394C allele and equal to or higher than the complex I activity of the B4c and F1 cybrids with the wild-type 3394T allele (Fig. 1). Moreover, we could not find any M9 mtDNAs with the 3394T allele, suggesting that the T allele may be incompatible with the M9 mtDNA background. Finally, every time that the 3394C allele was found at high altitude, it occurred on a macrohaplogroup M mtDNA (M9, C4a4, D4, and G3b).

Discussion

The 3394C variant provides evidence that a mtDNA allele can be deleterious in one context but advantageous in another. At low altitude the 3394C variant increases the risk of LHON (14–18), which is consistent with its 15–28% reduction in complex I-specific activity and a 7–17% reduction in maximum respiration rate.

However, mtDNA background is a major contributing factor to the complex I activity associated with the 3394C variant, with greater differences being found between different haplogroups with the same ND1 allele (3394T) than were found within the same mtDNA haplogroup with different 3394 alleles (T vs. C). Therefore, both individual sequence variants and haplotypes influence mitochondrial function. That these differences can be phenotypically relevant is confirmed by the report that haplogroup F mtDNAs are associated with increased risk of developing diabetes (37), and these cybrids had the lowest complex I activities.

The 3394C variant fulfills all of the criteria expected for an adaptive variant enriched by high altitude. The 3394C variant is enriched in both Tibetan and the Indian subcontinent populations living above 1,500 m, OR = 21.9; it has been found on five different haplogroups above 1,500 m, with M9 predominating in the Himalayas and C4a4 in the Deccan plateau. Variants at amino acid ND1 30Y are found in other high-altitude primates. Finally, the variant is associated with LHON in low-altitude populations, and if this were true at high altitudes blindness would be common in Tibetans and the mutant allele would have had to be enriched by drift in the face of natural selection.

The 3394C variant is only found on macrohaplogroup M mtDNAs at high altitude and macrohaplogroup M mtDNAs themselves increase in frequency with altitude, the ratio of N to M mtDNAs for all three low elevation Asian populations being 49:51; for Tibetan samples 1 and 2 it was 26:74 and 30:70,

Fig. 3. Multiple independent origins of the 3394C variant in Tibet. (A) A simplified phylogenetic tree of mtDNA haplogroups and subhaplogroups observed among native Tibetans in sample 1. Steps in which the T3394C variant must have occurred are shown, revealing three independent mutational events. Note that all 3394C variants are associated with macrohaplogroup M, which is also true for the C4a4 mtDNAs of the Deccan plateau, Karnataka State. (B) A positive linear relationship between the frequency of the haplogroup M9 and altitude in Tibetan villages ($R^2 = 0.4018$). (C) A positive linear relationship between the frequency of the 3394C variant and altitude in Tibetan villages ($R^2 = 0.3748$). The names of the villages sampled are presented next to each symbol and the M9 and 3394C frequencies are provided in Table S12.

respectively. Therefore, both the 3394C variant and haplogroup M mtDNAs are enriched at higher elevations.

Although the 3394C variant on macrohaplogroup N B4 and F1 mtDNAs was associated with reduced complex I activity, the 3394C variant on the macrohaplogroup M mtDNA M9 had the highest complex I activity. Hence, the functional benefit of the 3394C allele is only manifested in the context of macrohaplogroup M. Still, haplogroup background is insufficient to explain all of the phenotypic consequences of the 3394C variant, because the 3394C variant on the M9 haplogroup has been associated with LHON in low-altitude Asian populations (24, 25). Thus, environment must also modulate the phenotypic consequence of the 3394C variant.

The deleterious consequences of the 3394C variant when occurring on macrohaplogroup N may be the result of most macrohaplogroup N mtDNAs harboring two additional mtDNA

amino acid variants, ND3 (nucleotide A10398C, A114T) and ATP6 (nucleotide G8701A, A59T), which macrohaplogroup M lacks (1, 38). The nucleotide 10398 and nucleotide 8701 variants have been correlated with alterations in mitochondrial matrix calcium and pH (39). European haplogroup J, a derivative of macrohaplogroup N, lacks the nucleotide 10398 variant but has the additional amino acid-altering variants, nucleotide 13708 in ND5 and nucleotides 14798 or 15257 in cytochrome b.

The environmental factors that render the 3394C variant on M9 deleterious at low altitude remain unknown (25), although reduced oxygen tension is one likely factor. Physiological alterations have been observed in high-altitude Tibetan, Peruvian Andean Native Americans, and India subcontinent populations (40, 41), which may be the result of changes in mitochondrial physiology (42, 43). For example, an alteration in complex I might increase mitochondrial reactive oxygen species production, which increases to toxic levels under high- but not low-oxygen tension. Alternatively, increased NADH/NAD⁺ ratios could impede the tricarboxylic acid cycle, increasing succinate and fumarate levels which could stabilize hypoxia-inducible factors 1 and 2 (HIF-1, HIF-2) and increase adaptation to hypoxia (44). The genes of the HIF-1 and -2 pathways have been found to be altered in Tibetan and Andean populations (45–48).

Whatever the interactions between the 3394C allele, macrohaplogroup M, and high altitude, it seems clear that the 3394C variant can be adaptive in some contexts but deleterious in others. Thus, genetic variants that alter bioenergetics, which most commonly occur in the mtDNA, may be the long-sought common genetic alleles that predispose to “complex” diseases.

Materials and Methods

Biochemical Analysis of the mtDNA ND1Y30H Variant. Preparation of 3394T vs. 3394C mtDNA cybrids. Cybrids were prepared by fusing subject blood platelets to human mtDNA-deficient (p^0) osteosarcoma 143B (TK⁻) cells. Cybrids were selected in DMEM with 10% (vol/vol) FCS containing 30 μ g/mL BrdUrd in the absence of uridine (49).

Mitochondrial complex I (NADH: coenzyme Q₁ oxidoreductase) activity. Mitochondria were isolated by differential centrifugation and the pellets were resuspended in iso-osmotic solution and frozen. The mitochondria were permeabilized by incubation in assay buffer with alamethicin (40 μ g/mL). Complex I was assayed using 25 μ g of mitochondria, coenzyme q₁, and NADH. NADH oxidation was monitored at 340 nm using a dual-beam spectrophotometer, with or without rotenone, and rotenone-sensitivity calculated (SI Materials and Methods).

The complex I assays of the B4c cybrids involved two independent subjects for each allele, encompassing eight and nine separate mitochondrial isolations from B4c 3394T and B4c 3394C clones, respectively. The B4c 3394T 3515C cybrid clone complex I was assayed six times.

The complex I activity of the F1a cybrid clones involved three separate mitochondrial isolations from a cybrid clone of the F1a 3394C subject, two mitochondrial isolations from a cybrid clone of one of the F1a 3394T subjects and one mitochondrial isolation from the cybrid clone of the other F1a 3394C subject. The results of 10 complex I assays were averaged for the three cybrid clones (see SI Results).

The mtDNA/ndDNA ratios of all clones assayed for complex I activity were within 20% of the mean of the mtDNA/ndDNA ratio of all clones and did not correlate with enzyme activities (SI Materials and Methods and Fig. S2).

Cybid oxygen consumption. O₂ consumption rates of 5×10^5 cells/mL were monitored in the Oroboros Oxygraph-2K in medium without serum at 37 °C. Basal and uncoupled rates [0.4 μ M carbonyl cyanide p-(trifluoromethoxy) phenylhydrazone (FCCP) dissolved in DMSO] were monitored.

Population Sampling and Analysis. All sampled individuals could trace their ancestry through the maternal lineage back at least three generations. Samples were collected by informed consent consistent with National Institutes of Health guidelines and the provisions of the respective countries.

mtDNA analysis. All nucleotide positions are reported relative to the Revised Cambridge Reference Sequence (50). The mtDNA was amplified from whole blood in 22 overlapping PCR fragments, and each fragment subjected to digestion with 18 informative restriction endonucleases (28, 51) (SI Materials and Methods). The entire mtDNA was sequenced for 39 Tibetan mtDNAs

and 18 SW China Han mtDNAs (*SI Materials and Methods*). The GenBank accession nos. are FJ748704–FJ748759.

Sequence collections and analyses. The Tibetan mtDNA data from the RFPL/HVS I analysis and the complete mtDNA sequence data were classified into haplogroups using the Mitomap Phylogeny (52). In addition, the complete sequences of 7,088 human mtDNAs were downloaded from National Center for Biotechnology Information (NCBI) (May 2010). Statistical analysis used SAS (see *SI Materials and Methods*).

ACKNOWLEDGMENTS. The authors thank all the Tibetans and Han Chinese who participated in this study; Goar Melkonian, for coordinating material exchange and laboratory management at the Center for Molecular and

Mitochondrial Medicine and Genetics (MAMMAG), University of California at Irvine; Abigail Bigham, Jin Li, and Zhenghua Wei for blood sample transportation and storage; and Sen Dan for assistance in recruiting the southern California Han Cohort. This work was supported by Grant 30772456 from the National Science Foundation of China and a grant from the Xinqiao Hospital for International Collaborative Research between Xinqiao Hospital of the Third Military Medical University and MAMMAG, University of California at Irvine (to F.J.); National Science Foundation Grant BNS-8919645 and National Institutes of Health (NIH) Grants HLBI 079647 and TW001188 (to L.G.M.); NIH Clinical and Translational Science Awards Grant TR000153 (to UC Irvine); NIH Grants AG24373, NS21328, DK73691, and AG13154 (to D.C.W.); Doris Duke Clinical Interfaces Grant 2005057 (to D.C.W.); and California Institute of Regenerative Medicine Comprehensive Grant RC1-00353-1 (to D.C.W.).

- Wallace DC (2005) A mitochondrial paradigm of metabolic and degenerative diseases, aging, and cancer: A dawn for evolutionary medicine. *Annu Rev Genet* 39:359–407.
- Zaragoza MV, Brandon MC, Diegoli M, Arbustini E, Wallace DC (2011) Mitochondrial cardiomyopathies: how to identify candidate pathogenic mutations by mitochondrial DNA sequencing, MITOMASTER and phylogeny. *Eur J Hum Genet* 19:200–207.
- Wallace DC, Lott MT, Procaccio V (2007) Mitochondrial genes in degenerative diseases, cancer and aging. *Emery and Rimoin's Principles and Practice of Medical Genetics*, Chapter 13, eds Rimoin DL, Connor JM, Pyeritz RE, Korf BR (Churchill Livingstone Elsevier, Philadelphia), 5th Ed, Vol 1, 194–298.
- Trounce IA, Kim YL, Jun AS, Wallace DC (1996) Assessment of mitochondrial oxidative phosphorylation in patient muscle biopsies, lymphoblasts, and transmittochondrial cell lines. *Methods Enzymol* 264:484–509.
- Wallace DC, Ruiz-Pesini E, Mishmar D (2003) mtDNA variation, climatic adaptation, degenerative diseases, and longevity. *Cold Spring Harb Symp Quant Biol* 68:479–486.
- Mishmar D, et al. (2003) Natural selection shaped regional mtDNA variation in humans. *Proc Natl Acad Sci USA* 100:171–176.
- 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.
- Wallace DC, et al. (1988) Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. *Science* 242:1427–1430.
- Huoponen K, Vilkkilä J, Aula P, Nikoskelainen EK, Savontaus ML (1991) A new mtDNA mutation associated with Leber hereditary optic neuroretinopathy. *Am J Hum Genet* 48:1147–1153.
- Johns DR, Neufeld MJ, Park RD (1992) An ND-6 mitochondrial DNA mutation associated with Leber hereditary optic neuropathy. *Biochem Biophys Res Commun* 187:1551–1557.
- Jun AS, Brown MD, Wallace DC (1994) A mitochondrial DNA mutation at np 14459 of the ND6 gene associated with maternally inherited Leber's hereditary optic neuropathy and dystonia. *Proc Natl Acad Sci USA* 91:6206–6210.
- Jun AS, Trounce IA, Brown MD, Shoffner JM, Wallace DC (1996) Use of trans-mitochondrial cybrids to assign a complex I defect to the mitochondrial DNA-encoded NADH dehydrogenase subunit 6 gene mutation at nucleotide pair 14459 that causes Leber hereditary optic neuropathy and dystonia. *Mol Cell Biol* 16(3):771–777.
- Novak BJ, et al. (2007) Exhaled methyl nitrate as a noninvasive marker of hyperglycemia in type 1 diabetes. *Proc Natl Acad Sci USA* 104:15613–15618.
- Brown MD, Allen JC, Van Stavern GP, Newman NJ, Wallace DC (2001) Clinical, genetic, and biochemical characterization of a Leber hereditary optic neuropathy family containing both the 11778 and 14484 primary mutations. *Am J Med Genet* 104:331–338.
- Brown MD, Trounce IA, Jun AS, Allen JC, Wallace DC (2000) Functional analysis of lymphoblast and cybrid mitochondria containing the 3460, 11778, or 14484 Leber's hereditary optic neuropathy mitochondrial DNA mutation. *J Biol Chem* 275:39831–39836.
- Carelli V, et al. (1999) Biochemical features of mtDNA 14484 (ND6/M64V) point mutation associated with Leber's hereditary optic neuropathy. *Ann Neurol* 45:320–328.
- Carelli V, et al. (1997) Leber's hereditary optic neuropathy: biochemical effect of 11778/ND4 and 3460/ND1 mutations and correlation with the mitochondrial genotype. *Neurology* 48:1623–1632.
- Baracca A, et al. (2005) Severe impairment of complex I-driven adenosine triphosphate synthesis in Leber hereditary optic neuropathy cybrids. *Arch Neurol* 62:730–736.
- Brown MD, et al. (2002) The role of mtDNA background in disease expression: A new primary LHON mutation associated with Western Eurasian haplogroup J. *Hum Genet* 110:130–138.
- Brown MD, Sun F, Wallace DC (1997) Clustering of Caucasian Leber hereditary optic neuropathy patients containing the 11778 or 14484 mutations on an mtDNA lineage. *Am J Hum Genet* 60:381–387.
- Brown MD, Torroni A, Reckord CL, Wallace DC (1995) Phylogenetic analysis of Leber's hereditary optic neuropathy mitochondrial DNA's indicates multiple independent occurrences of the common mutations. *Hum Mutat* 6:311–325.
- Torroni A, et al. (1997) Haplotype and phylogenetic analyses suggest that one European-specific mtDNA background plays a role in the expression of Leber hereditary optic neuropathy by increasing the penetrance of the primary mutations 11778 and 14484. *Am J Hum Genet* 60:1107–1121.
- Puomila A, et al. (2007) Epidemiology and penetrance of Leber hereditary optic neuropathy in Finland. *Eur J Hum Genet* 15:1079–1089.
- Zhang M, et al. (2010) Mitochondrial haplogroup M9a specific variant ND1 T3394C may have a modifying role in the phenotypic expression of the LHON-associated ND4 G11778A mutation. *Mol Genet Metab* 101:192–199.
- Liang M, et al. (2009) Leber's hereditary optic neuropathy is associated with mitochondrial ND1 T3394C mutation. *Biochem Biophys Res Commun* 383:286–292.
- Fan W, et al. (2008) A mouse model of mitochondrial disease reveals germline selection against severe mtDNA mutations. *Science* 319:958–962.
- Stewart JB, et al. (2008) Strong purifying selection in transmission of mammalian mitochondrial DNA. *PLoS Biol* 6:e10.
- Torroni A, et al. (1994) Mitochondrial DNA analysis in Tibet: Implications for the origin of the Tibetan population and its adaptation to high altitude. *Am J Phys Anthropol* 93:189–199.
- Yao YG, Kong QP, Bandelt HJ, Kivisild T, Zhang YP (2002) Phylogeographic differentiation of mitochondrial DNA in Han Chinese. *Am J Hum Genet* 70:635–651.
- Kong QP, et al. (2003) Mitochondrial DNA sequence polymorphisms of five ethnic populations from northern China. *Hum Genet* 113:391–405.
- Kong QP, et al. (2003) Phylogeny of east Asian mitochondrial DNA lineages inferred from complete sequences. *Am J Hum Genet* 73:671–676.
- Yao YG, Kong QP, Wang CY, Zhu CL, Zhang YP (2004) Different matrilineal contributions to genetic structure of ethnic groups in the silk road region in China. *Mol Biol Evol* 21:2265–2280.
- Saillard J, Forster P, Lynnerup N, Bandelt HJ, Nørby S (2000) mtDNA variation among Greenland Eskimos: The edge of the Beringian expansion. *Am J Hum Genet* 67:718–726.
- Zhao M, et al. (2009) Mitochondrial genome evidence reveals successful Late Paleolithic settlement on the Tibetan Plateau. *Proc Natl Acad Sci USA* 106:21230–21235.
- Chandrasekar A, et al. (2009) Updating phylogeny of mitochondrial DNA macro-haplogroup m in India: dispersal of modern human in South Asian corridor. *PLoS ONE* 4:e7447.
- Stammach E (1987) Desert, forest, and montane baboons: Multilevel societies. *Primate Societies*, eds Smuts BB, Cheney DL, Seyfarth RM, Wrangham RW, Struhsaker TT (Univ of Chicago Press, Chicago), pp 112–120.
- Fuku N, et al. (2007) Mitochondrial haplogroup N9a confers resistance against type 2 diabetes in Asians. *Am J Hum Genet* 80:407–415.
- Wallace DC (2007) Why do we still have a maternally inherited mitochondrial DNA? Insights from evolutionary medicine. *Annu Rev Biochem* 76:781–821.
- Kazuno AA, et al. (2006) Identification of mitochondrial DNA polymorphisms that alter mitochondrial matrix pH and intracellular calcium dynamics. *PLoS Genet* 2:e128.
- Moore LG (2001) Human genetic adaptation to high altitude. *High Alt Med Biol* 2:257–279.
- Beall CM, Song K, Elston RC, Goldstein MC (2004) Higher offspring survival among Tibetan women with high oxygen saturation genotypes residing at 4,000 m. *Proc Natl Acad Sci USA* 101:14300–14304.
- Hochachka PW, et al. (1996) ³¹P magnetic resonance spectroscopy of the Sherpa heart: A phosphocreatine/adenosine triphosphate signature of metabolic defense against hypobaric hypoxia. *Proc Natl Acad Sci USA* 93:1215–1220.
- Zhao XX, Gao WX, Gao YQ, Suo L, Chen J (2007) Placental mitochondrial respiratory function of native Tibetan at high altitude (in Chinese). *Zhonghua Yi Xue Za Zhi* 87:894–897.
- Wallace DC, Fan W, Procaccio V (2010) Mitochondrial energetics and therapeutics. *Annu Rev Pathol* 5:297–348.
- Suzuki K, et al. (2003) Genetic variation in hypoxia-inducible factor 1α and its possible association with high altitude adaptation in Sherpas. *Med Hypotheses* 61:385–389.
- Yi X, et al. (2010) Sequencing of 50 human exomes reveals adaptation to high altitude. *Science* 329:75–78.
- Simonson TS, et al. (2010) Genetic evidence for high-altitude adaptation in Tibet. *Science* 329:72–75.
- Bigham A, et al. (2010) Identifying signatures of natural selection in Tibetan and Andean populations using dense genome scan data. *PLoS Genet* 6:1–14.
- Chomyn A (1996) Platelet-mediated transformation of human mitochondrial DNA-less cells. *Methods Enzymol* 264:334–339.
- Andrews RM, et al. (1999) Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 23:147.
- Starikovskaya EB, et al. (2005) Mitochondrial DNA diversity in indigenous populations of the southern extent of Siberia, and the origins of Native American haplogroups. *Ann Hum Genet* 69:67–89.
- Ruiz-Pesini E, et al. (2007) An enhanced MITOMAP with a global mtDNA mutational phylogeny. *Nucleic Acids Res* 35(Database issue):D823–D828.