No association between microsomal triglyceride transfer protein (*MTP*) haplotype and longevity in humans

Almut Nebel*, Peter J. P. Croucher†, Rieke Stiegeler‡, Susanna Nikolaus‡, Michael Krawczak†, and Stefan Schreiber*‡§

Institutes for *Clinical Molecular Biology and †Medical Informatics and Statistics and ‡Hospital for General Internal Medicine, University Hospital Schleswig–Holstein, 24105 Kiel, Germany

Edited by Louis M. Kunkel, Harvard Medical School, Boston, MA, and approved April 20, 2005 (received for review November 22, 2004)

Human longevity is a multifactorial condition with a significant genetic contribution. A recent association study in two independent samples of long-lived U.S. Caucasians [long-lived individuals (LLI)] identified a SNP haplotype of the microsomal triglyceride transfer protein (*MTP*, 4q25) that was underrepresented among LLI when compared with younger controls. This suggested that variation in the *MTP* gene might modify human longevity. Because of its function in lipid metabolism, the *MTP* gene product could plausibly play a pivotal role in the physiology of aging. However, the association observed in the U.S. samples could not be replicated by the same authors in a larger French LLI sample. We have therefore investigated the *MTP* "risk" haplotype in our own collection of 1,589 German nonagenarians, centenarians, and appropriately matched controls. No statistically significant differences were observed between LLI and controls at the allele, genotype, or haplotype level. This indicates that a noteworthy influence of the respective *MTP* haplotype on human longevity in the German population is unlikely. Furthermore, in comparison with all other U.S. and European samples analyzed, the *MTP* "risk" haplotype was found to be overrepresented only in the U.S. controls. This implies that the putative association is more likely to reflect recent changes in the genetic structure of the U.S. Caucasian population as a whole, rather than genetic effects upon survival to old age. In our view, the original study therefore highlights potential problems that arise when the case-control design is used as a means to map longevity genes in humans.

case-control study design | replication study

Longevity in humans can be thought of as a multifactorial condition to which both genetic and environmental factors are likely to contribute. Studies in twins have shown that longevity is moderately heritable, and that genetic variation accounts for ≈25% of the observed variation in life expectancy (1, 2). Similarly, family studies of centenarians from the U.S. and Iceland have indicated a significant genetic component of longevity (3, 4).

Recently, the gene encoding microsomal triglyceride transfer protein (*MTP*) has been reported as the first positional candidate for a heritable influence upon human life span (5). This link was drawn in an association-based fine-mapping experiment around microsatellite D4S1564, a genetic marker for longevity previously identified via genome-wide linkage analysis (6). A two-marker haplotype within the *MTP* gene was found to be significantly underrepresented in a sample of 190 long-lived Americans of European extraction, suggesting that it confers a higher mortality earlier in life (5). The "risk" haplotype was defined by the minor allele of rs2866164 (allele G) and the major allele of *MTP* Q/H 95 (allele Q95). Marker rs2866164 is a SNP in the *MTP* gene promoter known to be in perfect linkage disequilibrium with another promoter mutation, rs1800591 (also described as −493 G/T; ref. 7). *MTP* Q/H 95 is a nonsynonymous single-base-pair substitution in exon 3 of the *MTP* gene, leading to an exchange of glutamine with histidine at amino acid position 95 (8). Additional support for an influence of *MTP* gene variation upon longevity was obtained by the replication of the initial haplotype association in independent sets of 250 proactively matched U.S. Caucasian cases and controls (5). However, in the same study, no haplotype frequency differences were detected in a large French sample. Only a departure from the Hardy–Weinberg equilibrium was observed among French long-lived individuals (LLI) but not in the controls (5).

The *MTP* gene product has been proposed to be involved in the correct assembly of very low-density lipoprotein and chylomicron particles (9), and drugs that inhibit MTP activity improve lipoprotein profiles (10). This function appears to lend additional credibility to an association of *MTP* gene variation with human longevity, because coronary artery disease attributed to unfavorable lipid profiles is one of the major causes of death in the western world (5). However, in view of the observed discrepancies between the U.S. and French samples and the consequent uncertainty about the possible involvement of MTP in human longevity, we attempted to replicate the findings of Geesaman *et al.* (5) in an extensive collection of >1,000 LLI and 550 controls of German ancestry. In our view, the observed lack of notable haplotype frequency differences in the German samples highlights the particular problems arising from the inherent absence of age-matching in case-control longevity studies.

Subjects and Methods

Study Population. The "case" sample comprised 1,039 unrelated individuals of German ancestry who were between 95 and 109 years of age at the time of recruitment (mean age, 98.2 years). They represented the top percentiles of the respective birth cohort-specific age distributions (95th percentile for females and 98th percentile for males). Three hundred seventy-three (36%) of the long-lived participants were centenarians or supercentenarians (mean age, 101.4 years). The sex ratio in the entire sample was 74% females vs. 26% males.

Nonagenarians and centenarians were recruited from different geographic regions throughout Germany. Their names and addresses were obtained from local registry offices ("Einwohnermeldeamt") and from the Federal Administrative Office, using the Centenarian Graduation List recently issued by the German President of State. All potential participants were contacted by a letter in which the study was explained and permission was asked to send a questionnaire and a blood sampling kit. Probands were advised to consult their general practitioner for taking a blood sample and for help with filling out the questionnaire. A full socioeconomic, quality of life, and health status assessment was performed. Primary criteria for

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: MTP, microsomal triglyceride transfer protein; LLI, long-lived individuals.

§To whom correspondence should be addressed. E-mail: s.schreiber@mucosa.de.

© 2005 by The National Academy of Sciences of the USA

Table 1. Summary statistics for SNP rs2866164 in German, French, and U.S. Caucasian LLI and controls

Sample (no. of individuals)	Frequency of minor allele (G)	P_{allele}	$P_{genotype}$	Odds ratio (Wald 95% CI)
Germans				
LLI (1,033)	0.269	0.34	0.13	1.18 (0.96–1.46)
Centenarians (370)	0.276	0.28	0.30	1.21 (0.93–1.58)
Controls (540)	0.253			
French				
LLI (559)	0.253	0.54	NA	NA
Controls (552)	0.264			
U.S. Caucasians				
LLI, sample 1 (175)	0.260	0.039	NA	NA
Controls, sample 1 (183)	0.331			
LLI, sample 2 (244)	0.252	0.029	NA	NA
Controls, sample 2 (231)	0.316			

P_{allele}, *P* value from a χ^2 test (1 df) for allele frequency difference; $P_{genotype}$, *P* value from a χ^2 test (2 df) for genotype frequency difference; CI, confidence interval; NA, not available. Data of French and U.S. Caucasian samples are as reported by Geesaman *et al.* (5). The second sample of the U.S. Caucasians represents proactively matched LLI and controls. For the U.S. Caucasians, the minor allele frequency and *P* value were recalculated from the reported counts and differ slightly from those in the original publication.

inclusion in the study were good health, physical activity, and absence of major diseases. Furthermore, it was required that participants be mentally fit and able to answer a health history questionnaire routinely used in clinical studies. A total of 550 unrelated German control individuals aged between 60 and 75 years (mean age, 67.2 years) were recruited by the same means and in the same regions as the nonagenarians and centenarians. The controls matched the LLI by ancestry (all German), geographic origin within the country, and sex.

All participants gave written informed consent to participate in the study. Lifestyle and health-related and genetic data were completely anonymized before being logged into an electronic database. Approval for the project was received from the Ethics Committee of the University Hospital Schleswig–Holstein, Campus Kiel, and from local data protection authorities.

Genotyping. Markers rs1800591, rs2866164, and *MTP* Q/H 95 were genotyped by using the Taqman Allelic Discrimination method with fluorescence-labeled minor groove-binding probes on an Applied Biosystems 7700 Sequence Detector. Genotype data were managed by means of an integrated database system (11). Additionally, positive controls on each genotyping plate were checked for consistency. All markers were tested for deviation from the Hardy–Weinberg equilibrium, as described below. The call rate for each marker exceeded 99%.

Statistical Analysis. Markers were assessed for significant deviations from the Hardy–Weinberg equilibrium by using an exact test with 1,000 permutations, as implemented in ARLEQUIN, Ver. 2.001 (http://lgb.unige.ch/arlequin). The significance of allele and genotype frequency differences between samples was tested by using a χ^2 statistic with the appropriate degrees of freedom. Genotype-based odds ratios were computed, combining carriers of the minor allele of each marker in one genotype class, and 95% confidence intervals were calculated by using the Web-based Simple Interactive Statistical Analysis (SISA) tool, available at http://home.clara.net/sisa. Haplotype frequencies were estimated from nonphased genotype data by using an expectation maximization (EM) algorithm. To evaluate the possible association with longevity of individual haplotypes, two different approaches were followed. First, EM-inferred haplotype frequencies were subjected to a regression-based haplotype association test with WHAP (www.genome.wi.mit.edu/~shaun/whap). In the second approach (called HAPRAND), case and control haplotype frequency estimates were used to calculate a pseudo-χ^2 statistic (12). Significance was assessed in both instances by using 10,000 random permutations of cases and controls. Power calculations were also carried out by means of a Web-based tool, available at the University of California, Los Angeles, Department of Statistics (http://calculators.stat.ucla.edu/pc).

Results

Both the entire collection of 1,039 German LLI and a subsample of 373 centenarians were subjected to a gender-matched case-control analysis of markers rs1800591, rs2866164, and Q/H 95, located in the *MTP* gene. All three SNPs were found to be in Hardy–Weinberg equilibrium. No evidence for association was detected between any of the tested SNPs and the longevity phenotype in either sample at the allele or at the genotype level. The association statistics for the most informative SNP, rs2866164 (for which comparable data on the U.S. Caucasian and French samples were also available), are presented in Table 1.

In all four LLI samples from Germany, France, and the U.S., the frequency of the minor rs2866164 allele (G) fell into the range between 0.252 and 0.269 (heterogeneity $\chi^2 = 1.18$, 3 df, $P = 0.76$). The allele frequency distribution in the German and French control groups was also found to be very similar ($\chi^2 = 0.39$, 1 df, $P = 0.53$). Only the two U.S. control samples were characterized by a significantly increased frequency of the G allele of rs2866164 when compared with the two European controls ($\chi^2 = 8.62$, 2 df, $P = 0.013$ for U.S. sample 1, $\chi^2 = 6.76$, 2 df, $P = 0.034$ for U.S. sample 2).

As expected, markers rs1800591 and rs2866164 were found to be in perfect linkage disequilibrium when all three SNPs were jointly analyzed in the German samples. Haplotype reconstruction yielded only three of the eight possible haplotypes (Table 2). For both the German LLI and centenarians, the permutation tests did not indicate any significant haplotype frequency differences between cases and controls. ''Risk'' haplotype rs2866164-G, *MTP* Q95 was found to be overrepresented in the U.S. control samples (Table 3) instead of being underrepresented in any of the LLI, as suggested by Geesaman *et al.* (5).

The American and European samples were also characterized by important demographic differences. For instance, the American controls were on average much younger (38.6 years) than the German (67.2 years) and French (51.2 years) controls. To

GENETICS

Table 2. *MTP* haplotype frequencies in German LLI, centenarians, and controls

Haplotype*	Controls	LLI	P_{WHAP}	$P_{HAPRAND}$	Centenarians	P_{WHAP}	$P_{HAPRAND}$
G-C–Q	0.749	0.732	0.35	0.31	0.724	0.29	0.24
T-G–Q	0.195	0.217	0.16	0.15	0.224	0.20	0.15
T-G–H	0.055	0.051	0.58	0.64	0.052	0.81	0.76
Global test			0.36	0.25		0.41	0.41

P_{WHAP}, permutation *P* value from a regression-based test for a haplotype frequency difference between LLI or centenarians, respectively, and controls; $P_{HAPRAND}$, permutation *P* value for a pseudo-χ^2 statistic.
*Haplotypes are given in marker order rs1800591, rs2866164, *MTP* Q/H 95.

rule out an age effect masking the putative *MTP* association, we also typed 359 younger German controls (mean age, 39.8 years) for SNP rs2866164. The G allele of rs2866164 had a frequency of 0.231 in the younger German control group, which did not differ significantly from the 0.253 observed in the older controls ($\chi^2 = 1.09$, 1 df, $P = 0.30$). Consequently, no age-related trend toward a higher frequency of G in younger individuals, as suggested by the American samples, exists among the Germans. The American LLI also comprised a larger proportion of males than the European collections. However, stratification of the case-control comparison by gender did not reveal any gender-specific association in the German sample (data not shown).

Discussion

In a recent case-control study of U.S. Caucasians, a two-SNP-haplotype in the *MTP* gene on chromosome 4q was implicated as a modifier of human life span (5). In the same publication, however, the authors reportedly failed to confirm their findings in a much larger sample from France, raising concerns about the validity of the original association. To critically evaluate these results, we have performed a large-scale study on >1,000 German nonagenarians and centenarians and 550 matched controls. The validity of our study population for longevity research has been supported by an examination of the apolipoprotein E (*APOE*) ε2, ε3, and ε4 haplotypes, which revealed an increased frequency of ε2 in LLI when compared with controls (0.118 vs. 0.085; OR [CI] = 1.45 [1.12–1.86]) (OR, odds ratio; CI, confidence interval) and a concurrently decreased frequency of ε4 (0.071 vs. 0.145; OR [CI] = 0.45 [0.36–0.57]). These results agree with other longevity studies and corroborate both the validity and efficacy of our study sample. Assuming an *MTP* risk factor frequency of 0.2 and a relative risk of 1.4 (characteristic values of the U.S. samples), our study sample would have had a power >80% to replicate the reported association. Yet, no statistically significant frequency differences were observed between German LLI and controls at the allele, genotype, or haplotype level.

Table 3. Frequency of *MTP* "risk" haplotype rs2866164-G, *MTP* Q95 in German, French, and U.S. Caucasian LLI and controls

Sample (no. of haplotypes)	Frequency
Germans	
LLI (2,036)	0.217
Centenarians (728)	0.224
Controls (1,066)	0.195
French	
LLI (1,126)	0.202
Controls (1,121)	0.197
U.S. Caucasians	
LLI, sample 1 (346)	0.176*
Controls, sample 1 (364)	0.286
LLI, sample 2 (488)	0.193*
Controls, sample 2 (462)	0.253

Data on French and U.S. Caucasian samples are from Geesaman *et al.* (5).
*Haplotype frequency differences between American cases and controls were reported as statistically significant by the same authors. The second U.S. Caucasian sample represents proactively matched LLI and controls.

Geesaman *et al.* (5) reported that the minor allele (G) of rs2866164 and the rs2866164-G, *MTP* Q95 haplotype of the *MTP* gene were significantly underrepresented in long-lived U.S. Caucasians, suggesting this allele or haplotype conferred a higher mortality earlier in life. However, comparison with all other samples analyzed reveals that the *MTP* "risk" haplotype is in fact consistently overrepresented in U.S. controls. Thus, the question arises as to which factors could have contributed to this discrepancy.

First, the higher frequencies of the rs2866164-G allele and of the "risk" haplotype in the U.S. control samples could have been false-positive results due to population stratification. U.S. Caucasians represent a considerably admixed population to which immigrants from a number of different European countries have contributed over the last 400 years. Because their origin has changed with time, the ethnic and therefore genetic composition of the LLI sample (going back to the beginning of the 20th century) in the U.S. study (5) is likely to differ from that of the control population (comprising more recent generations of U.S. citizens). Geesaman *et al.* (5) went to great lengths to limit the likelihood of a stratification artifact by (*i*) replicating the association in independent U.S. Caucasian LLI and control samples, (*ii*) correcting for population structure using the genomic control approach based on 60 random SNPs, and (*iii*) undertaking proactive matching of their cases and controls. However, it has recently been shown that, in studies designed to identify modest genetic risk factors, the use of a few dozen SNPs may in general lack sufficient power to reveal even moderate levels of stratification (13–15). Thus, that proactive matching was deemed necessary in the original study at all implies that, from the outset, the U.S. LLI and controls may not have been well matched. The genomic control method that was subsequently applied, although state of the art at the time, was perhaps insufficient to detect the residual stratification. In this case, differential gene flow and admixture could have contributed to the *MTP* haplotype frequency difference between the two U.S. Caucasian control samples and all other samples analyzed.

Second, inappropriate matching of cases and controls under even modest levels of population structure not only can lead to false positives but may also cause false-negative findings (15). Although the German control individuals were selected to ensure they were as close as possible to the LLI in terms of their age and geographic origin, systematic genetic differences between these samples cannot be completely excluded. Thus, latent population structure may have been responsible for the failure of the present study to confirm the association observed in the Americans, but this would of course not explain the negative finding in the French samples. Another reason for the lack of replication in both European samples could have been the higher proportion of males in the U.S. LLI. Men reach 100 years of age less often than women. It would therefore seem plausible that particular risk alleles impact more upon male than female

mortality, so that these alleles are more strongly underrepresented in male than female centenarians. However, no gender-specific effect has been observed in the German samples.

Third, the observed discrepancy between the American control individuals and both European populations could reflect differences brought about by recent changes of the selective forces acting upon the *MTP* gene in the U.S. Caucasian population. During the first half of the 20th century, U.S. Caucasians became one of the most advanced societies in the world, with medical, technological, and sociological improvements that had not been experienced by earlier generations (represented by the U.S. LLI) or by contemporary Europeans. In large populations, even small changes of selective pressure can create considerable changes in allele frequencies over relatively short periods of time (16). Because it is not known whether MTP has functions (other than those in lipid metabolism, for instance, in reproductive or cognitive processes) that would render it a suitable target of natural selection, the type of selective pressure that might be involved remains undefined at present.

Fourth, as originally pointed out by Geesaman *et al.* (5), strong environmental factors, such as diet, that modify the influence of a risk allele in different populations may have contributed to the observed inconsistency between studies. However, because the main difference between the Americans and the Europeans applies to the *MTP* allele and haplotype frequency in the control individuals, this explanation seems unlikely.

More work will be required to discern which of the above explanations for the lack of replication of an original association finding is correct in the case of the *MTP* gene. Taken together, however, currently available evidence suggests that the putative association observed in the Americans (5) was due to recent changes in the genetic structure of the U.S. Caucasian population as a whole, rather than to genetic effects upon survival to old age. Although the known function of MTP in lipid metabolism renders it a plausible candidate for human longevity, the reported association appears likely to have been artifactual. The study by Geesaman *et al.* (5) thus highlights the potential pitfalls that arise when a case-control design is used as a means to map longevity genes in humans. Such studies would generally be hampered by the inherent difficulty of appropriately matching case and control samples simply because, if living individuals are compared, they were born generations apart. Our own study also emphasizes the importance of replicating initial association findings in very large samples from different populations, especially for complex phenotypes in which genes with only weak or moderate effects are likely to play a role.

We thank all probands for their participation. We are grateful to I. Urbach, T. Henke, B. Petersen, and T. Wesse for technical assistance; R. Vogler for database management; and J. Heinze for help with recruitment. This study was funded by the Federal Ministry of Science and Education through an Explorative Project of the National Genome Research Network (NGFN-2). We acknowledge the support of our collaborators in the Genetics of Healthy Aging Consortium (GEHA), 6th Framework Program of the European Commission.

1. McGue, M., Vaupel, J. W., Holm, N. & Harvald, B. (1993) *J. Gerontol.* **48,** B237–B244.
2. Herskind, A. M., McGue, M., Holm, N. V., Sorensen, T. I., Harvald, B. & Vaupel, J. W. (1996) *Hum. Genet.* **97,** 319–323.
3. Gudmundsson, H., Gudbjartsson, D. F., Frigge, M., Gulcher, J. R. & Stefansson, K. (2000) *Eur. J. Hum. Genet.* **8,** 743–749.
4. Perls, T. T., Wilmoth, J., Levenson, R., Drinkwater, M., Cohen, M., Bogan, H., Joyce, E., Brewster, S., Kunkel, L. & Puca, A. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 8442–8447.
5. Geesaman, B. J., Benson, E., Brewster, S. J., Kunkel, L. M., Blanche, H., Thomas, G., Perls, T. T., Daly, M. J. & Puca, A. A. (2003) *Proc. Natl. Acad. Sci. USA* **100,** 14115–14120.
6. Puca, A. A., Daly, M. J., Brewster, S. J., Matise, T. C., Barrett, J., Shea-Drinkwater, M., Kang, S., Joyce, E., Nicoli, J., Benson, E., *et al.* (2001) *Proc. Natl. Acad. Sci. USA* **98,** 10505–10508.
7. Karpe, F., Lundahl, B., Ehrenborg, E., Eriksson, P. & Hamsten, A. (1998) *Arterioscler. Thromb. Vasc. Biol.* **18,** 756–761.
8. Narcisi, T. M., Shoulders, C. C., Chester, S. A., Read, J., Brett, D. J., Harrison, G. B., Grantham, T. T., Fox, M. F., Povey, S., de Bruin, T. W., *et al.* (1995) *Am. J. Hum. Genet.* **57,** 1298–1310.
9. Jamil, H., Chu, C. H., Dickson, J. K., Jr., Chen, Y., Yan, M., Biller, S. A., Gregg, R. E., Wetterau, J. R. & Gordon, D. A. (1998) *J. Lipid Res.* **39,** 1448–1454.
10. Wetterau, J. R., Gregg, R. E., Harrity, T. W., Arbeeny, C., Cap, M., Connolly, F., Chu, C. H., George, R. J., Gordon, D. A., Jamil, H., *et al.* (1998) *Science* **282,** 751–754.
11. Hampe, J., Wollstein, A., Lu, T., Frevel, H. J., Will, M., Manaster, C. & Schreiber, S. (2001) *Bioinformatics* **17,** 654–655.
12. Croucher, P. J., Mascheretti, S., Hampe, J., Huse, K., Frenzel, H., Stoll, M., Lu, T., Nikolaus, S., Yang, S. K., Krawczak, M., *et al.* (2003) *Eur. J. Hum. Genet.* **11,** 6–16.
13. Ardlie, K. G., Lunetta, K. L. & Seielstad, M. (2002) *Am. J. Hum. Genet.* **71,** 304–311.
14. Freedman, M. L., Reich, D., Penney, K. L., McDonald, G. J., Mignault, A. A., Patterson, N., Gabriel, S. B., Topol, E. J., Smoller, J. W., Pato, C. N., *et al.* (2004) *Nat. Genet.* **36,** 388–393.
15. Marchini, J., Cardon, L. R., Phillips, M. S. & Donnelly, P. (2004) *Nat. Genet.* **36,** 512–517.
16. Gillespie, J. H. (1998) *Population Genetics: A Concise Guide* (Johns Hopkins Univ. Press, Baltimore).