Study Population

Briefly, 1,441 subjects with type 1 diabetes were recruited into the DCCT between 1983 and 1989; they were between 13 and 39 years of age, and 53% were male. The primary prevention cohort consisted of 726 subjects with no retinopathy, an albumin excretion rate < 28 µg per minute, and diabetes duration 1 to 5 years. The secondary intervention cohort consisted of 715 subjects who had non-proliferative retinopathy, a urinary albumin excretion rate < 140 µg per minute, and diabetes duration of 1 to 15 years. Subjects in the primary prevention and secondary intervention cohorts were randomized to either intensive or conventional treatment arms, and assessed for complications at frequent follow up visits. The DCCT was terminated in 1993 when the principal study question concerning treatment effects had been answered [1]. The mean duration of follow-up in the DCCT was 6.5 years. In 1994, 1375 subjects (95%) of the original cohort, 687 from the intensive arm and 688 from the conventional arm, agreed to participate in the EDIC follow up study which included annual examinations for complication status [2]. DNA was obtained from 1418 consenting subjects.

Renal Function

Renal function was assessed at yearly visits in the DCCT and on alternate years (based on randomization in the DCCT) during EDIC with approximately half of the cohort evaluated in even EDIC follow-up years and half in the odd EDIC follow up years. The assessment included measurement of urinary albumin excretion based on 4 hour timed urine collections. Antihypertensive medications including ACE inhibitors or angiotensin II receptor blockers were not discontinued. Urine albumin was measured by

fluoroimmunoassay [3] with coefficients of variation and coefficients of reliability of 14% and 95% respectively[4].

Glycosylated hemoglobin was measured by high performance liquid chromatography in a central laboratory at baseline and then quarterly during the DCCT and annually in the EDIC study. Time dependent updated mean HbA_{1c} levels were weighted according to the time a subject was followed in the DCCT and EDIC. Annual medical history, physical exam including body mass index (BMI), blood pressure measured by sphygmomanometer, and laboratory testing were performed as described in the DCCT [1].

Determination of ACE genotypes

DNA was extracted from peripheral blood lymphocytes using sodium dodecylsulfate cell lysis followed by a salt precipitation using Puregene reagents (Gentra System, Inc., Minneapolis, MN). A PCR method [5] was used to determine the ACE InDel polymorphism. The sense and anti-sense primers were CTGGAGACCACTCCCATCCTTTCT and GATGTGGCCATCACATTCGTCAGAT, respectively. The anti-sense primer was labeled with the fluorescent dye 6-FAM (Applied Biosystems) and fragment analysis performed on an ABI 3700 DNA analyzer. The PCR reaction included 12.5 ng genomic DNA, 6 pmoles primer, 2µl DMSO, 0.1 mM $MgCL_2$ 2.6 µl Buffer II (Applied Biosystems), 0.8 mM dNTPs and 0.2 units of *AmpliTaq* DNA polymerase(Applied Biosystems) in a final volume of 25 µl. PCR conditions included an initial denaturation at 93° C for 3 min, and then DNA was amplified using 32 cycles at 92°C for 1 min, 57°C for 1 min, 72°C for 1 min, and a final extension at 72°C for 10 min. The PCR products were then run on an ABI 3700 DNA

analyzer and the deletion allele migrated at 191 bp and the insertion allele at 475 bp corresponding to the insertion of the 287 bp Alu sequence in intron 16 of ACE. Because the D allele in heterozygous samples may be preferentially amplified leading to the mistyping of a D/I genotype as a D/D genotype, all subjects with DD genotype were subjected to a second independent PCR amplification using primer pairs that recognize the insertion sequence and are therefore specific for the insertion allele [6]. The sense primer was replaced with an insertion specific primer 5'TTTGAGACGGAGTCTCGCTC and the PCR reaction mixture included 25 ng genomic DNA, 0.1mM $MnCL_2$, 0.2mM dNTPs, 1.25 µl DMSO, 5 units Taq polymerase, 6 pM each of primer pairs. PCR was carried out as previously outlined with an annealing temperature of 67°C and run for 40 cycles. The 408 bp insertion specific PCR fragment was identified on 10% polyacrylamide gels with ethidium bromide staining.

The two SNPs, rs1800764 and rs9896208 were also genotyped. The following primers 5'GAGATTGTGCCACTGCACTCC, 5'GAGAGTCTTGGAATGTA CCCACTGA and TaqMan MGB probes: 6FAMTGTACAGCAACCCCMGBNFQ, VICATGTACAGCAGCCCCMGBNFQ were used to genotype rs1800764. The primers 5'TATCTGTCCCCCTTAGTCCACA and 5' ACTTCCACCCTCTTTCAATCCT as well as the following TaqMan probes: 6FAMTCTGTGTGTGCGTTGTAMGBNFQ, VICTCTGTGTGTGTGTTGTACMGBNFQ were used to genotype rs9896208. Standard PCR TaqMan reaction mix was used according to manufacturer's specifications.

Genotype Quality Control

To ensure quality control (QC), four CEPH (Centre d'Etude du Polymorphisme Humain) control samples in duplicate and one negative control were included in each 96

well plate. Allele calls were made independently by 2 individuals to ensure accuracy and consistency of genotypes and five percent of the sample was regenotyped to monitor quality assurance. Allele frequencies were determined and genotypes were tested for departure from Hardy-Weinberg equilibrium.

Haplotype Construction

In the 1365 unrelated subjects, linkage disequilibrium between marker pairs was determined using the ldmax option in the GOLD software[10]. The linkage disequilibrium (LD) estimates Δ^2 and D' (in the 5' to 3' direction) between marker rs1800764 and InDel are 0.58 and 0.88 ($p<10^{-5}$), respectively. Between InDel and rs9896208, Δ^2 and D' are 0.46 and 0.93 ($p<10^{-5}$), respectively. The estimates of Δ^2 and D' between the two flanking markers rs1800764 and rs9896208, are 0.34 and 0.69 ($p<10^{-5}$) respectively. The LD of each of the flanking markers to the InDel polymorphism was therefore strongest, and as would be expected, there was less LD between the flanking markers rs1800764 and rs9896208, given that the genomic distance was greater and intragenic recombination more likely.

In the DCCT/EDIC Caucasian subjects three locus haplotype frequencies (frequency $\pm$ SE) for rs1800764 (T/C), InDel (I/D), and rs9896208 (T/C) markers were estimated via Bayesian algorithms implemented in the program PHASE (version 2.1) [11] to be TIC (0.45 $\pm$ 0.001), CDT (0.30 $\pm$ 0.002), CDC (0.12 $\pm$ 0.002), TDT (0.053 $\pm$ 0.002). These frequency estimates differ at the most common haplotype but are similar to the other 3 haplotypes reported in 151 French Caucasian families by Soubrier et al. 2002: TIC (0.32), CDT (0.27), CDC (0.10), TDT (0.07) [7]. Haplotype pairs with frequencies <5% were pooled into a single category, "other" and included CIC, CIT,

TDC, TDT, TIT. The following are the common haplotype pair frequency estimates for the DCCT/EDIC Caucasian subjects: CDT/TIC (0.26), TIC/TIC (0.19), CDC/TIC (0.13), TIC/other (0.12) CDT/other (0.10) CDT/CDT (0.096), CDC/CDT (0.065).

Statistical Analysis

To identify potential confounding variables, we descriptively examined the association between each ACE marker and DCCT baseline variables using Pearson chi-square for categorical variables and Kruskal-Wallis tests for quantitative ones. Initially, non-parametric Kaplan-Meier plots and log-rank tests were used to verify the proportional hazards assumptions and to assess covariate effects on complication-free survival; continuous variables were categorized by tertiles or quartiles. Subsequent analyses of the risk of complications employed discrete-time proportional hazards models[8], corresponding to the annual visit schedule, stratified by the year of DCCT entry (1983-1989); continuous covariates were log transformed as appropriate (details are provided in table 1) and centered.

Preliminary proportional hazards models examining gene effects were also adjusted for cohort, treatment and their interaction. Each of the homozygous genotypes was compared to the more frequent heterozygous genotype via a 1df Wald test. Final proportional hazards models were adjusted for baseline DCCT covariates and time-dependent HbA1c updated means. For each gene locus, the Maddala R^2 statistic that describes the explained proportion of outcome variability was computed using the likelihood ratio statistic from final multivariate models with and without the respective genotype indicators[9].

References

1. The Diabetes Control and Complications Trial Research Group.. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* **329**, 977-86 (1993).
2. Epidemiology of Diabetes Interventions and Complications (EDIC). Design, implementation, and preliminary results of a long-term follow-up of the Diabetes Control and Complications Trial cohort. *Diabetes Care* **22**, 99-111 (1999).
3. Effect of intensive therapy on the development and progression of diabetic nephropathy in the Diabetes Control and Complications Trial. The Diabetes Control and Complications (DCCT) Research Group. *Kidney Int* **47**, 1703-20 (1995).
4. Sustained effect of intensive treatment of type 1 diabetes mellitus on development and progression of diabetic nephropathy: the Epidemiology of Diabetes Interventions and Complications (EDIC) study. *Jama* **290**, 2159-67 (2003).
5. Rigat, B. et al. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *J Clin Invest* **86**, 1343-6 (1990).
6. Shanmugam, V., Sell, K.W. & Saha, B.K. Mistyping ACE heterozygotes. *PCR Methods Appl* **3**, 120-1 (1993).
7. Soubrier, F. et al. High-resolution genetic mapping of the ACE-linked QTL influencing circulating ACE activity. *Eur J Hum Genet* **10**, 553-61 (2002).
8. Cox, D. Regression models and life tables(with discussion). *Journal of the Royal Statistical Society* **B34**, 187-220 (1972).
9. Schemper, M. Further results on the explained variation in proportional hazards regression. *Biometrika* **79**, 202-4 (1992).
10. Abecasis, G.R. & Cookson, W.O. GOLD--graphical overview of linkage disequilibrium. *Bioinformatics* **16**, 182-3 (2000).
11. Stephens, M., Smith, N.J. & Donnelly, P. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet* **68**, 978-89 (2001).