

Diabetologia, accepted June 2, 2009 (Author manuscript; not copy-edited)

A susceptibility gene for type 2 diabetes confers substantial risk for diabetes complicating cystic fibrosis

S. M. Blackman^{1,2}, S. Hsu¹, S. E. Ritter², K. M. Naughton², F. A. Wright³, M. L. Drumm⁴, M. R. Knowles⁵, G. R. Cutting²

Abstract

Aims/Hypothesis: Insulin-requiring diabetes affects 25-50% of young adults with cystic fibrosis (CF). Although the cause of diabetes in CF is unknown, recent heritability studies in CF twins and siblings indicate that genetic modifiers play a substantial role. We sought to assess whether genes conferring risk for diabetes in the general population may play a modifier role in CF.

Methods: We tested whether a family history of type 2 diabetes affected diabetes risk in CF patients in 539 families in the CF Twin and Sibling family-based study. A type 2 diabetes susceptibility gene (transcription factor 7-like 2, or *TCF7L2*) was evaluated for association with diabetes in CF using 998 patients from the family study and 802 unrelated CF patients in an independent case-control study.

Results: Family history of type 2 diabetes increased the risk of diabetes in CF (odds ratio=3.1; p=0.0009). A variant in *TCF7L2* associated with type 2 diabetes (rs7903146, "T" allele) was associated with diabetes in CF in the family study (p=0.004) and in the case-control study (p=0.02; combined p=0.0002). In the family study, variation in *TCF7L2* increased the risk of diabetes about 3-fold (hazard ratio=1.75 per allele, 95% CI 1.3-2.4; p=0.0006), and decreased the mean age at diabetes diagnosis by 7 years. In CF patients not treated with systemic glucocorticoids, the effect of *TCF7L2* was even greater (hazard ratio=2.9 per allele, 95% CI 1.7-4.9, p=0.00011).

Conclusions: A genetic variant conferring risk for type 2 diabetes in the general population is a modifier of risk for diabetes in CF.

Author affiliations: ¹Division of Pediatric Endocrinology, Johns Hopkins University School of Medicine, Baltimore, MD; ²McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD; ³Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill, Chapel Hill, NC; ⁴Departments of Pediatrics and Genetics, Case Western Reserve University, Cleveland, OH; and ⁵Cystic Fibrosis-Pulmonary Research and Treatment Center, School of Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC

Running title: Association of *TCF7L2* with CF-related diabetes

Corresponding author: Dr. Garry R. Cutting, M.D.

Mailing address: McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, 733 N. Broadway, Broadway Research Building 559, Baltimore, MD 21205

Phone: (410) 614-0212 **Fax:** (410) 614-0213 **E-mail:** gcutting@jhmi.edu

Words: 3180 (abstract: 250); **References:** 39

Keywords: cystic fibrosis, diabetes, CFRD, genetics, association, glucocorticoid, corticosteroid, *TCF7L2*, modifier gene

Introduction

Diabetes is an important complication of cystic fibrosis (CF), a multisystem, genetic disease caused by mutations in the cystic fibrosis transmembrane conductance regulator (*CFTR*) gene [1]. Diabetes risk increases with age, affecting about 25% of adolescents and 40-50% of adults with CF [2;3]. As CF patients are now living longer, diabetes has become the most common systemic complication of CF after lung disease. Diabetes is associated with a significantly worse CF prognosis [4], although treatment of diabetes improves nutritional status and pulmonary function [5]. Among CF patients, the prevalence of diabetes is approximately 10-fold greater (by about one-third the age) than is seen for type 2 diabetes in the general population [6].

About 90% of patients with CF have exocrine pancreatic insufficiency (PI), which is determined largely by *CFTR* mutation [7], and PI is a requirement for developing diabetes in CF [2;3]. However, *CFTR* mutations do not otherwise explain diabetes risk; rather, we recently demonstrated that variation in other genes accounts for most of the risk for diabetes in CF [8]. Although type 1 and type 2 diabetes also have a strong genetic basis, it is unknown whether susceptibility genes for type 1 or type 2 diabetes might also play a role in CF. On one hand, there appears to be less similarity with type 1 diabetes, considering that diabetes incidence increases steadily with age rather than peaking in childhood, and that islet autoimmunity characteristic of type 1A diabetes plays little or no role in CF according to most (e.g., [9;10]) but not all (e.g., [11]) reports. On the other hand, type 2 diabetes, like diabetes in CF, shows a monotonic increase in prevalence with age [6], a progressive defect in insulin secretion [12;13], and an accumulation of amyloid polypeptide in pancreatic islets [14;15]. These similarities suggest that disease mechanisms of diabetes in CF could be shared

with type 2 diabetes seen in the general population.

In this study, we evaluated the relatives of CF patients to determine if diabetes in CF and type 2 diabetes cluster in the same families. We tested whether the rs7903146 variant in *TCF7L2*, the type 2 diabetes susceptibility gene whose common variant confers the greatest risk [16;17], is a risk factor for diabetes in CF and whether treatment with systemic glucocorticoids, an important non-genetic risk factor for diabetes in CF, affected the risk conferred by *TCF7L2*.

Methods

Study subjects. Clinical data and DNA samples were collected by the Cystic Fibrosis Twin and Sibling family-based study [18;19], with additional clinical data provided by the U.S. CF Foundation Patient Registry (Bethesda, MD). Enrolment was based on conclusive diagnosis of CF [20] and having an affected twin and/or sibling. As diabetes occurs almost exclusively in CF patients who are pancreatic insufficient [3], we selected 998 CF patients with pancreatic exocrine insufficiency or a *CFTR* genotype predictive for severe pancreatic exocrine insufficiency (as described in [18]) from 1154 patients (86%) in the CF Twin and Sibling study for analysis (Table 1). Monozygous (MZ) twins were excluded to avoid oversampling of MZ transmission patterns. Each CF patient was classified using longitudinal clinical and laboratory data (described in [8]) as Euglycaemic (at least 2 diabetes screens done; all within normal limits), Hyperglycaemic (milder glucose abnormalities such as impaired glucose tolerance), Possible Diabetes (e.g., intermittent or resolved diabetes), or Diabetes (chronic, insulin-treated diabetes for ≥ 1 year). Fifty-two were excluded (49 for insufficient information, 3 for antibody-positive type 1 diabetes).

Family history obtained by questionnaire [21] (available upon request) was stratified by degree of relatedness (e.g., first-degree, such as a parent or sibling, vs. second-degree, such as a grandparent, aunt, or uncle) and was categorized as positive (at least 1 first-degree or 2 second-degree relatives on the same side of the family giving a “yes” response for “Diabetes-Adult onset,” recognizing that about 90% of diabetic adults have type 2 diabetes [22]), possible (any other combination of second-degree relatives), or negative (no first- or second-degree relatives). A negative family history required completion of both nuclear and extended pedigree and no reported first or second-degree relatives with adult-onset diabetes. The “Matched Euglycaemic” group of pancreatic insufficient Euglycaemic patients was selected to match the Diabetes patients by age and sex as described previously [8]. Other phenotypes were obtained from medical records, by questionnaire, and from the CF Foundation Patient Registry. Systemic glucocorticoid treatment was defined as oral or intravenous glucocorticoids (i.e., not intranasal or inhaled). Definitions of meconium ileus [18], lung function [19] and nutritional [23] phenotypes are described elsewhere.

The case-control study of unrelated CF patients used for replication is described elsewhere [24]. Briefly, the UNC/CWRU Genetic Modifiers Study included 167 CF patients with diabetes and 635 CF patients without diabetes, all homozygous for the $\Delta F508$ mutation in *CFTR* and affected with pancreatic exocrine insufficiency. Diabetes was defined as a diagnosis by a physician plus insulin use. This CF clinic sample was originally ascertained based on age and lung function criteria: severe lung disease (260 patients of mean age 15.8 years; 38 diabetic), mild lung disease in an older age group (246 patients of mean age 37.7 years; 88 diabetic) or mild lung disease in a younger age group (296 patients of mean age 20.6 years; 41 diabetic) [24]. These studies were approved by the Institutional Review Boards of all

participating institutions, and informed consent was obtained from all subjects before enrolment.

DNA isolation and genotyping. DNA isolation and identification of *CFTR* mutations have been previously described [18]. TaqMan® Assays-on-Demand (Applied Biosystems) were used to genotype *TCF7L2* single nucleotide polymorphism rs7903146 as described elsewhere [23]. Genotype repeatability was 99.7% in 1133 duplicated samples. One Mendelian error was identified in 678 families in the Family study and was corrected by DNA sequencing. DNA from those with absent or erroneous genotypes was sequenced using the BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Genotypes were verified by sequencing in 2% of the samples; no discrepancies were detected.

Statistical Analysis. Genetic association was deemed present when the frequency of allele inheritance differed significantly from random Mendelian transmission of alleles. Transmission distortion was quantified using the FBAT statistic as implemented in the software packages FBAT 1.7.3 [25] and PBAT 3.5 [26] with an additive genetic model and an optimal phenotypic offset [27]. Essentially identical results were obtained using an offset equal to the prevalence rate of diabetes in the analyzed sample and offsets of 0 (considering only affected individuals) or 1 (considering only unaffected individuals). Age of onset of diabetes was analyzed as a censored trait using PBAT. Genetic association was also assessed after disregarding family structure by comparison of genotype and allele frequencies with diabetes status (case-control association analysis) and with diabetes-free survival rates (association with cumulative incidence of disease). The same results were obtained when analysis was restricted to individuals who were not informative in the family-based analysis (not

shown). Population-attributable risk was calculated from hazard ratios (e.g., [28]).

Statistical calculations were performed using Intercooled Stata 10 (StataCorp, College Station, TX). In case of small cell sizes, the Fisher exact test was used. The Student t test was used to compare normally distributed continuous data. The log rank test was used to compare cumulative incidence rates. Significance levels from analysis of the primary (Family) and replication (Case-control) studies were combined calculating an average of the Z-statistics weighted by the square root of the number of individuals in each study (e.g. [29]).

Results

Diabetes in a first-degree relative correlated with diabetes in the CF patient (OR=2.8, 95% CI 1.2-6.1, $p=0.01$; Table 2). Diabetes in at least two second-degree relatives on the same side of the family trended toward correlation (OR=2.2, 95% CI 0.93-5.1, $p=0.07$). CF patients meeting either of these two conditions had an increased risk for diabetes in CF (OR=2.8, 95% CI 1.5-5.3, $p=0.001$). Family history was known for about half the study group and did not depend on the CF patient's diabetes status. Because our CF patients with diabetes were, on average, older than other CF patients in this study, we examined whether the greater risk for familial diabetes might be due to older family members. This was not the case, as the correlation between a positive family history of diabetes (especially in a first-degree relative) and diabetes in CF remained substantial when diabetic CF patients were compared with euglycaemic CF patients matched for age, sex, and pancreatic status ("Matched Euglycaemic"; Table 2).

Genotypes of the *TCF7L2* SNP (rs7903146) were obtained for 943 of the 998 available CF patients, providing 533 parent-parent-child trios for analysis. Parental genotypes were in Hardy-Weinberg equilibrium ($p=0.7$), and their minor allele frequency (28%)

was in the range reported for Caucasian populations [30]. In families, the "T" allele of the *TCF7L2* SNP was transmitted to CF patients with diabetes significantly more often than expected by chance, demonstrating association with increased risk of diabetes (200 informative families, $p=0.005$; FBAT). Distorted transmission was also seen in separate analyses of the Diabetic ("T" allele overtransmitted, 54 informative families, $p=0.04$) and Euglycaemic ("C" allele overtransmitted, 146 informative families, $p=0.01$; FBAT) subgroups, demonstrating that variation in the *TCF7L2* SNP operates at both phenotypic extremes. Finally, diabetic CF patients displayed substantial differences from Euglycaemic CF patients in the frequency of the high-risk *TCF7L2* SNP. As shown in ESM Table 1, differences were seen in allele frequency ($p=0.007$) and genotype frequency ($p=0.02$ for additive model, $p=0.009$ for dominant model) that yielded 1.7-fold increase in disease odds per allele (95% CI 1.5-2.4, $p=0.007$).

Variation at *TCF7L2* also affected the age at which diabetes was diagnosed. Treating the diabetes diagnosis as a time-dependent event (by censoring at the present age), we noted that CF patients in the family study who had inherited the "T" allele of *TCF7L2* developed diabetes at a younger age than those who inherited the C allele (211 informative families, log rank $p=0.01$; Wilcoxon $p=0.002$, PBAT). The magnitude of this effect is estimated by comparing diabetes incidence with *TCF7L2* genotype for all CF patients in the family study, disregarding family structure (Figure 1(a)). Each high risk "T" allele of *TCF7L2* conferred a significant increase in the cumulative incidence of diabetes (log rank $p=0.001$; hazard ratio (HR) = 1.75, 95% CI 1.3-2.4; $p=0.0006$). While it is possible that specific mutations of the *CFTR* gene could affect diabetes risk, similar results were obtained when analysis was restricted to the 518 individuals homozygous for $\Delta F508$, the most common *CFTR* genotype (HR=1.5 per allele, 95% CI 1.0-2.3, $p=0.049$) or the other

439 individuals (not homozygous for $\Delta F508$) with two mutations causing severe *CFTR* dysfunction (e.g., frameshift or nonsense mutations; HR=2.17, 95% CI 1.3-3.6, $p=0.003$; difference in HR not significant, $p=0.2$).

Association of the *TCF7L2* SNP with diabetes in CF was also seen in an independent collection of unrelated CF patients recruited originally based on age and lung disease by the UNC/CWRU Genetic Modifiers Study [24]. In that study, those at greatest risk for diabetes (due to increased age) are a subgroup of 246 patients recruited for older age and milder lung disease (88 with diabetes, 158 without). In the older group, a significant increase in risk for those carrying a T allele was evident at both the allele (OR=1.6, 95% CI 1.1-2.4, $p=0.02$) and genotype levels (OR = 1.9, 95% CI 1.1-3.6, $p=0.016$; dominant model) and by regression analysis (per-allele OR=1.92, 95% CI 1.1-3.3, $p=0.015$) as shown in ESM Table 1. Genotype frequencies in two subgroups of younger patients (with mild or severe lung disease) from the UNC/CWRU study that are predicted to be at lower risk for diabetes did not show association with diabetes. When age of onset of diabetes was considered, there was an increase in diabetes cumulative incidence with one or two *TCF7L2* T alleles in both the older age subgroup (log rank $p=0.0035$; HR = 1.9, 95% CI 1.2-2.9, $p=0.004$, dominant model), and when all 802 unrelated patients were considered (log rank $p=0.02$; HR = 1.4, 95% CI, 1.1-2.0, $p=0.02$, Figure 1(b)). Thus, the T allele of the *TCF7L2* SNP conferred increased risk and earlier age at diagnosis for diabetes in the family-based and case-control studies. Considered jointly, the two studies support association of the rs7903146 SNP of *TCF7L2* with diabetes in CF ($p=2.0 \times 10^{-4}$).

We tested whether an important non-genetic risk factor for diabetes in CF, treatment with systemic glucocorticoids ([8;31;32] and Table 1), may interact with *TCF7L2*. Regression modelling in the Family study suggested an interaction between steroid treatment and

TCF7L2 genotype (ESM Table 2). That is, *TCF7L2* did not significantly affect diabetes risk in the 405 CF patients receiving glucocorticoid treatment in the last year (HR = 1.4 per allele, 95% CI 0.9-2.1, $p=0.14$; ESM Figure 1(a)) or in the 103 CF patients treated with glucocorticoids for at least 30 days (HR = 1.1 per allele, 95% CI 0.6-2.0, $p=0.8$), whereas the risk ratio for *TCF7L2* was greater in the other 485 patients who were not treated with glucocorticoids in the last year (HR = 2.9 per allele, 95% CI 1.7-4.9, $p=1.1 \times 10^{-4}$; ESM Figure 1(b)). The *TCF7L2* SNP remained an independent risk factor with essentially the same effect size after adjusting for cross-sectional or longitudinal measures of lung function, body mass index, and family history of type 2 diabetes (data not shown). Data were not available to perform a similar analysis in the replication (case-control) study. In summary, adjusting for steroid exposure, an important clinical risk factor for diabetes in CF, reveals a subgroup in which the effect size of *TCF7L2* association is substantially increased.

Discussion

Diabetes in CF patients and type 2 diabetes in the general population share some pathophysiologic features, and both have a strong genetic basis. This multicenter study demonstrates, with independent replication, that a susceptibility gene for type 2 diabetes (*TCF7L2*) plays a role in modifying the risk for diabetes in CF. Family history data suggest that other susceptibility genes for type 2 diabetes may also play a role. Finally, a non-genetic factor, treatment with systemic glucocorticoids, modifies the risk conferred by *TCF7L2*. In the first analysis (family history), the increased diabetes risk to the CF patient depending on the number of and genetic distance to affected relatives, as is the case for type 2 diabetes [33], suggesting that familial sharing of genes is more important than sharing of environment. Two prior studies found no significant correlation between diabetes in CF and diabetes in family

members without CF [9;10]. Neither study was powered to detect effects of the magnitude measured in the current study, although one found a trend toward correlation [9].

From the validated susceptibility genes for type 2 diabetes, we selected *TCF7L2* for testing based on its having the greatest effect size and its validation in the largest number of patients [34]. Greater diabetes risk was associated with the same *TCF7L2* allele in CF as in the general population [16], indicating that this genetic risk factor is shared by these two forms of diabetes. Decreased pancreatic β -cell function is a hallmark of diabetes in CF (e.g., [2;13]), so the involvement of *TCF7L2* in CF supports roles hypothesized for *TCF7L2* in beta cell proliferation and/or insulin secretion (e.g., [35;36]). We hypothesize that variation in *TCF7L2* affects beta cell function in CF leading to earlier development of diabetes, as is predicted to occur in type 2 diabetes [37]. However, the effect of *TCF7L2* (relative risk of 1.42-1.51 per allele in the general population [30]) may be greater in some pancreatic insufficient CF patients, as suggested by the greater risk ratio seen in those patients not treated with systemic steroids (95% CI, 1.7-4.9). Using the hazard ratio in Fig. 1 to estimate effect size, variation at *TCF7L2* is estimated to account for a substantial fraction of the diabetes cases in the current CF study (population-attributable risk 32%, 95% CI: 14-49%). The cumulative incidence analyses may underestimate the magnitude of *TCF7L2* effect, because Hyperglycaemic and Possible Diabetes patients were grouped with Euglycaemic patients due to lack of age-of-onset data for these intermediate phenotypes. To illustrate the effect of combining this *TCF7L2* variant and *CFTR* dysfunction, we note that a 14 year old CF patient with two high-risk *TCF7L2* alleles has the same predicted risk for diabetes as a 60 year old individual in the general population with no *TCF7L2* risk alleles [6], a difference of 46 years attributable to variation in both *CFTR* and *TCF7L2*.

In the general population, the risk of type 2 diabetes conferred by *TCF7L2* appears to be reduced by interaction with environmental factors such as lifestyle intervention [37]. In CF, the current study found evidence that treatment with systemic steroids influences the effect of *TCF7L2* on risk for diabetes by 1.7-fold. Accounting for sources of environmental variation such as steroid treatment made the remaining sources of variation in CF patients (such as from genetic modifiers) more prominent. While there has been debate as to the utility of stratifying risk for type 2 diabetes by typing individuals for susceptibility genes (e.g., [38;39]), in the case of CF patients not treated with steroids, a substantially earlier predicted onset for diabetes is likely to be clinically relevant, and could be used to individualize the timing of diabetes screening and initiation of treatment.

Strengths of the current study include the use of family-based association testing which is insensitive to artefacts generated by population stratification [25], and the depth and breadth of longitudinal clinical data (collected as part of CF care and largely unrelated to whether patients were diagnosed with diabetes). A limitation is that not all patients received regular oral glucose tolerance testing. We sought to minimize the effect of this limitation by excluding individuals with insufficient data to classify their phenotype, and by using additional information such as postprandial glucose or haemoglobin A1c data. Because the highly detailed phenotype information was not available for the replication (case-control) study, some individuals may have been misclassified as not having diabetes in that study, reducing the apparent effect of *TCF7L2* in that sample. This consideration and the targeted recruitment strategy used in the latter study, limit the degree to which quantitative comparisons of *TCF7L2* effects may be made between the primary and replication studies. Finally, it should be recognized that this study's findings are limited to the ~90% of CF patients who have severe

exocrine pancreatic insufficiency, as the ~10% of CF patients with significant pancreatic exocrine function are not at high risk for developing diabetes.

This study demonstrates diabetes in CF and type 2 diabetes have at least one pathologic mechanism in common. The higher frequency and earlier development of diabetes in CF patients present an opportunity to evaluate and optimize gene-based tests or therapies that may have implications for individuals with type 2 diabetes. On the other hand, the clinical differences exemplified by the underweight young adult with CF and diabetes and the older, often overweight adult with type 2 diabetes may be exploited to investigate how variation in genes such as *TCF7L2* lead to diabetes in different nutritional and inflammatory contexts. Thus, characterization of genetic modifiers of monogenic disorders like CF may help inform the mechanisms by which common polymorphisms cause common diseases such as type 2 diabetes.

Acknowledgements. The authors gratefully acknowledge the participation of the many CF patients, families, research coordinators, and clinicians in the CF Twin and Sibling Study, the Genetic Modifiers of CF Study, and the Canadian Consortium for CF Genetic Studies. Portions of this work were previously published in abstract form at the 2007 North American Cystic Fibrosis Conference and the 2007 meeting of the American Society of Human

Genetics. The JHU CF Twin and Sibling Study is supported by NIH grant DK076446 (S.B.), DK044003 and HL068927 (G.R.C.), and CF Foundation funding CUTTIN06P0 (G.R.C.). The UNC/CWRU Genetic Modifiers Study is supported by the CF Foundation (KNOWLE00A0) and NIH grants HL068890, DK066368 (M.R.K.) and RR000046. The funding sources had no role in the study design, collection, analysis, and interpretation of data, writing of the report, or the decision to submit for publication. The authors thank Lori L. Vanscoy, MD, J. Michael Collaco MD, Lindsay B. Henderson and Deanna Green, MD for assistance in defining important clinical covariates in the JHU study, Patrick Sosnay, MD and Kathryn McDougal, PhD for helpful discussions, Patricia Cornwall for administrative help, Keluo Yao for aid in SNP testing, and David J. Cutler, PhD for guidance in data analysis. The authors also thank Rhonda Pace for coordination of the UNC/CWRU molecular biology of the study; Karen Mohlke, PhD for advice about SNP selection; Jason Luo, Director of the Human Molecular Genotyping Core, for SNP testing; Sally Wood and Leia Charnin for coordination of the clinical phenotyping; Hollin Calloway for assistance in data analysis and illustrations; David Fargo, PhD and Airong Xu, MD for Bioinformatics support; and the Gene Modifier Study Group. The authors also thank Peter Durie, M.D., Ruslan Dorfman, Ph.D., Julian Zielenski, Ph.D., and Mary Corey, Ph.D. for reading the manuscript.

References

1. Welsh MJ, Ramsey BW, Accurso FJ, Cutting GR (2001) Cystic Fibrosis. In: Scriver CR, Beaudet AL, Valle D, Sly WS (eds) *The Metabolic and Molecular Bases of Inherited Disease*. Eighth edn. McGraw-Hill, Inc., New York, pp 5121-5188
2. Moran A, Hardin D, Rodman D, et al (1999) Diagnosis, screening and management of cystic fibrosis related diabetes mellitus: a consensus conference report. *Diabetes Res Clin Pract* 45: 61-73
3. Mackie AD, Thornton SJ, Edenborough FP (2003) Cystic fibrosis-related diabetes. *Diabet Med* 20: 425-436

4. Milla CE, Warwick WJ, Moran A (2000) Trends in pulmonary function in patients with cystic fibrosis correlate with the degree of glucose intolerance at baseline. *Am J Respir Crit Care Med* 162: 891-895
5. Mohan K, Israel KL, Miller H, Grainger R, Ledson MJ, Walshaw MJ (2007) Long-Term Effect of Insulin Treatment in Cystic Fibrosis-Related Diabetes. *Respiration*
6. Cowie CC, Rust KF, Byrd-Holt DD, et al (2006) Prevalence of diabetes and impaired fasting glucose in adults in the U.S. population: National Health And Nutrition Examination Survey 1999-2002. *Diabetes Care* 29: 1263-1268
7. Ahmed N, Corey M, Forstner G, et al (2003) Molecular consequences of cystic fibrosis transmembrane regulator (CFTR) gene mutations in the exocrine pancreas. *Gut* 52: 1159-1164
8. Blackman SM, Hsu S, Vanscoy LL, et al (2009) Genetic modifiers play a substantial role in diabetes complicating cystic fibrosis. *J Clin Endocrinol Metab* 94: 1302-1309
9. Minicucci L, Cotellessa M, Pittaluga L, et al (2005) Beta-cell autoantibodies and diabetes mellitus family history in cystic fibrosis. *J Pediatr Endocrinol Metab* 18: 755-760
10. Lanng S, Thorsteinsson B, Pociot F, et al (1993) Diabetes mellitus in cystic fibrosis: genetic and immunological markers. *Acta Paediatr* 82: 150-154
11. Nousia-Arvanitakis S, Galli-Tsinopoulou A, Dracoulacos D, Karamouzis M, Demitriadou A (2000) Islet autoantibodies and insulin dependent diabetes mellitus in cystic fibrosis. *J Pediatr Endocrinol Metab* 13: 319-324
12. Kahn SE (2001) Clinical review 135: The importance of beta-cell failure in the development and progression of type 2 diabetes. *J Clin Endocrinol Metab* 86: 4047-4058
13. Bismuth E, Laborde K, Taupin P, et al (2008) Glucose tolerance and insulin secretion, morbidity, and death in patients with cystic fibrosis. *J Pediatr* 152: 540-5, 545
14. Hull RL, Westermarck GT, Westermarck P, Kahn SE (2004) Islet amyloid: a critical entity in the pathogenesis of type 2 diabetes. *J Clin Endocrinol Metab* 89: 3629-3643
15. Couce M, O'Brien TD, Moran A, Roche PC, Butler PC (1996) Diabetes mellitus in cystic fibrosis is characterized by islet amyloidosis. *J Clin Endocrinol Metab* 81: 1267-1272
16. Grant SF, Thorleifsson G, Reynisdottir I, et al (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet* 38: 320-323
17. Weedon MN (2007) The importance of TCF7L2. *Diabet Med* 24: 1062-1066
18. Blackman SM, Deering-Brose R, McWilliams R, et al (2006) Relative contribution of genetic and nongenetic modifiers to intestinal obstruction in cystic fibrosis. *Gastroenterology* 131: 1030-1039
19. Vanscoy LL, Blackman SM, Collaco JM, et al (2007) Heritability of lung disease severity in cystic fibrosis. *Am J Respir Crit Care Med* 175: 1036-1043
20. Rosenstein BJ, Cutting GR (1998) The diagnosis of cystic fibrosis: A consensus statement. *J Pediatr* 132: 589-595
21. Kahn LB, Marshall JA, Baxter J, Shetterly SM, Hamman RF (1990) Accuracy of reported family history of diabetes mellitus. Results from San Luis Valley Diabetes Study. *Diabetes Care* 13: 796-798
22. Zimmet P, Alberti KG, Shaw J (2001) Global and societal implications of the diabetes epidemic. *Nature* 414: 782-787
23. Collaco JM, Vanscoy L, Bremer L, et al (2008) Interactions between secondhand smoke and genes that affect cystic fibrosis lung disease. *JAMA* 299: 417-424

24. Drumm ML, Konstan MW, Schluchter MD, et al (2005) Genetic modifiers of lung disease in cystic fibrosis. *N Engl J Med* 353: 1443-1453
25. Laird NM, Horvath S, Xu X (2000) Implementing a unified approach to family-based tests of association. *Genet Epidemiol* 19 Suppl 1: S36-S42
26. Steen KV, Lange C (2005) PBAT: A comprehensive software package for genome-wide association analysis of complex family-based studies. *Hum Genomics* 2: 67-69
27. Lunetta KL, Faraone SV, Biederman J, Laird NM (2000) Family-based tests of association and linkage that use unaffected sibs, covariates, and interactions. *Am J Hum Genet* 66: 605-614
28. Altshuler D, Hirschhorn JN, Klannemark M, et al (2000) The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet* 26: 76-80
29. Gaulton KJ, Willer CJ, Li Y, et al (2008) Comprehensive association study of type 2 diabetes and related quantitative traits with 222 candidate genes. *Diabetes* 57: 3136-3144
30. Cauchi S, El Achhab Y, Choquet H, et al (2007) TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med* 85: 777-782
31. Marshall BC, Butler SM, Stoddard M, Moran AM, Liou TG, Morgan WJ (2005) Epidemiology of cystic fibrosis-related diabetes. *J Pediatr* 146: 681-687
32. Adler AI, Gunn E, Haworth CS, Bilton D (2007) Characteristics of adults with and without cystic fibrosis-related diabetes. *Diabet Med* 24: 1143-1148
33. Weijnen CF, Rich SS, Meigs JB, Krolewski AS, Warram JH (2002) Risk of diabetes in siblings of index cases with Type 2 diabetes: implications for genetic studies. *Diabet Med* 19: 41-50
34. Helgason A, Palsson S, Thorleifsson G, et al (2007) Refining the impact of TCF7L2 gene variants on type 2 diabetes and adaptive evolution. *Nat Genet* 39: 218-225
35. Liu Z, Habener JF (2008) Glucagon-like peptide-1 activation of TCF7L2-dependent Wnt signaling enhances pancreatic beta-cell proliferation. *J Biol Chem*
36. Wegner L, Hussain MS, Pilgaard K, et al (2008) Impact of TCF7L2 rs7903146 on insulin secretion and action in young and elderly Danish twins. *J Clin Endocrinol Metab* 93: 4013-4019
37. Florez JC, Jablonski KA, Bayley N, et al (2006) TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med* 355: 241-250
38. Meigs JB, Shrader P, Sullivan LM, et al (2008) Genotype score in addition to common risk factors for prediction of type 2 diabetes. *N Engl J Med* 359: 2208-2219
39. Lyssenko V, Jonsson A, Almgren P, et al (2008) Clinical risk factors, DNA variants, and the development of type 2 diabetes. *N Engl J Med* 359: 2220-2232

Table 1. Comparison of clinical features of pancreatic insufficient CF patients (excluding MZ twins) in the twin/sibling study group with euglycaemia, hyperglycaemia or diabetes

	Euglycaemia	Hyperglycaemia or Possible Diabetes	Diabetes
Number of subjects	509	387	102
Mean age (years)	12.5 ^a	15.9 ^a	23.6
Caucasian	476/509 (94%)	353/387 (91%)	96/102 (94%)
Female	237/509 (47%)	175/387 (45%)	55/102 (54%)
CFTR genotype ΔF508 ×2^b	273/507 (54%)	218/386 (56%)	59/100 (59%)
Severe CFTR ×2^c	443/489 (91%)	356/369 (95%)	88/93 (95%)
Lung function Percentile§ Cross-sectional	0.73 (n=391) ^a	0.67 (n=357) ^a	0.55 (n=98)
Longitudinal	0.63 (n=253) ^a	0.58 (n=299) ^d	0.51 (n=94)
Nutritional status^e Cross-sectional	-0.08 (n=500) ^a	-0.49 (n=387) ^f	-0.76 (n=100)
Longitudinal	-0.04 (n=492) ^a	-0.29 (n=382)	-0.47 (n=94)
Steroid treatment (≥30 days/yr)	90/509 (18%) ^a	78/387 (20%) ^a	46/102 (45%)
Steroid treatment (days/yr)	19 (n=304) ^a	31 (n=268) ^a	80 (n=59)
Positive test for P. Aeruginosa	344/462 (74%) ^a	310/366 (85%) ^g	90/92 (98%)
Meconium Ileus	85/444 (19%) ^f	76/342 (22%) ^d	8/86 (9%)
Lung transplantation	1/451 (0.2%) ^a	4/361 (1%) ^g	9/89 (10%)

^aP<0.0001 vs. Diabetes

^bHomozygous for ΔF508

^cBoth *CFTR* alleles with severe (e.g., frameshift or nonsense) mutations (list available upon request).

§Cross-sectional and longitudinal measures of lung-function were based on forced expiratory volume in 1 second as defined previously [19].

^dP<0.01 vs. Diabetes

^eNutritional status was assayed by body mass index standard deviation score (BMI SDS), with longitudinal averages as defined previously [23]. Cross-sectional BMI SDS was calculated by averaging over the year prior to enrolment.

^fP<0.05 vs. Diabetes

^gP<0.001 vs. Diabetes

Table 2. Influence of a family history of type 2 diabetes on diabetes prevalence in pancreatic insufficient CF patients.

	Diabetic	Euglycaemic	Odds Ratio (95% CI)^a	Matched Euglycaemic^b	Odds Ratio (95% CI)^c
Both first- and second-degree family history known	65/128 (51%)	274/500 (55%)	0.85 (0.6-1.3)	68/128 (53%)	0.91 (0.5-1.4)
First-degree relative with diabetes	11/78 (14%)	18/320 (6%)	2.8 (1.2-6.1) ^d	2/83 (2%)	6.6 (1.4-31) ^d
≥2 second-degree relatives on the same side with diabetes	9/65 (14%)	19/276 (7%)	2.2 (0.93-5.1)	5/68 (7%)	2.0 (0.6-6.4)
≥1 first- or ≥2 second-degree relatives on the same side with diabetes	19/78 (24%)	33/321 (10%)	2.8 (1.5-5.3) ^e	7/83 (8%)	3.5 (1.4-9) ^f

^aOdds ratio (95% confidence interval) comparing family history in diabetic group vs. euglycaemic group

^bDiabetes patients were matched by age, sex, and number of $\Delta F508$ alleles with euglycaemic patients.

^cOdds ratio (95% confidence interval) comparing family history in diabetic group vs. matched euglycaemic group

^dP<0.05 vs. diabetic group

^eP=0.001 vs. diabetic group

^fP<0.01 vs. diabetic group

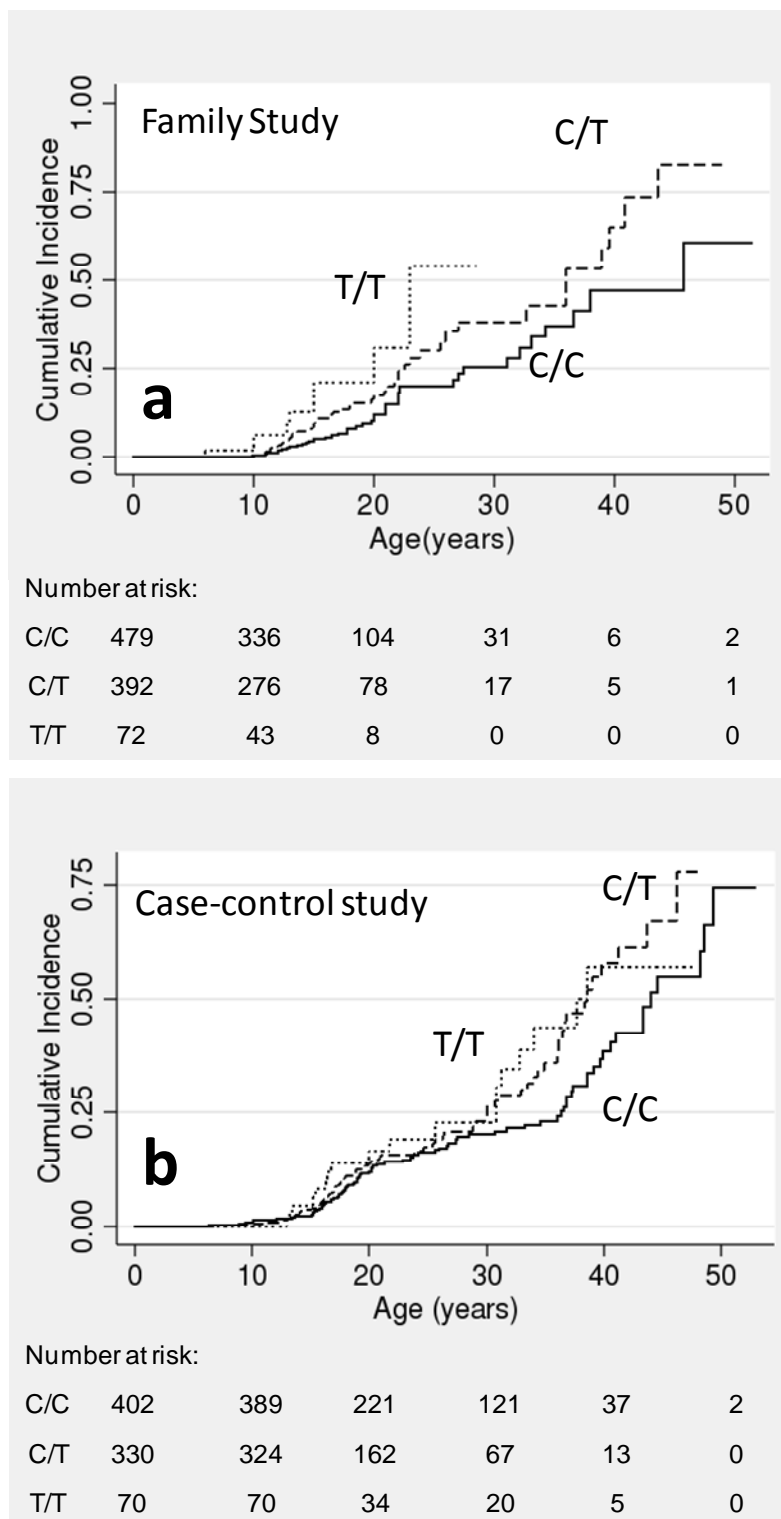


Figure 1: Cumulative incidence of diabetes grouped by *TCF7L2* SNP genotype. Individuals in the primary (family) study (**a**) who were homozygous for the major allele (C/C genotype, solid line) developed diabetes at a significantly lower rate than patients heterozygous (C/T genotype, dashed line) or homozygous for the minor allele (T/T genotype, dotted line). Results were similar when comparing all 3 genotypes separately as shown (additive genetic model; $p=0.001$, log rank) or when grouping

genotypes according to dominant ($p=0.003$) or recessive ($p=0.007$) genetic models. Modelling of these data indicated that each T allele conferred about a 75% increase in risk of developing diabetes compared to PI CF patients of the same age without a T allele (hazard ratio 1.75, 95% CI 1.3-2.4; $p=0.0006$). The age at which the cumulative incidence reached 25% varied by 7 years with the number of high-risk “T” alleles (C/C genotype: 27.4 years; C/T genotype: 22.6 years; T/T genotype: 20.0 years). Individuals in the replication (case-control) study (**b**) who were homozygous for the major allele (C/C genotype, solid line) developed diabetes at a significantly lower rate than patients with one or two copies of the minor allele (C/T genotype; dashed line; or T/T genotype: dotted line). In this study (which had a targeted recruitment strategy based on age and lung function) the data were more consistent with a dominant genetic model, with incidence rates for all 3 genotypes considered separately being less distinct ($p=0.06$, log rank), and rates for those with one or two copies of the T allele being significantly greater than for those with none ($p=0.02$, log rank). Modelling of these data showed association under dominant (HR = 1.43, 95% CI 1.1-2.0; $p=0.02$) or additive (HR = 1.3, 95% CI 1.04-1.61, $p=0.02$) genetic models.

Electronic Supplementary Material

ESM Table 1. Distribution of *TCF7L2* genotypes (SNP rs7903146) among CF patients stratified by diabetes phenotype^a.

	<i>TCF7L2</i> SNP Genotype			OR (per allele) ^b	P (Genotype) ^c	P (Allele) ^c
	CC	CT	TT			
Primary (Family) study, all patients	498 (51%)	405 (41%)	78 (8.0%)			
Euglycaemic (age ≥ 10 yr)	150 (57%)	98 (37%)	14 (5.3%)	1.7 (1.5-2.4), $p=0.007$	0.02 (Add.) 0.009 (Dom.)	0.007
Diabetes	41 (41%)	49 (49%)	9 (9.1%)			
Replication (Case-control) study, all patients	402 (50%)	330 (41%)	70 (8.7%)			
Old Mild (≥34 yrs), non-diabetes	99 (63%)	46 (29%)	13 (8.2%)	1.92 (1.13-3.26) $p=0.015$	0.052 (Add.) 0.016 (Dom.)	0.02
Old Mild (≥34 yrs), diabetes	41 (47%)	37 (42%)	10 (11%)			

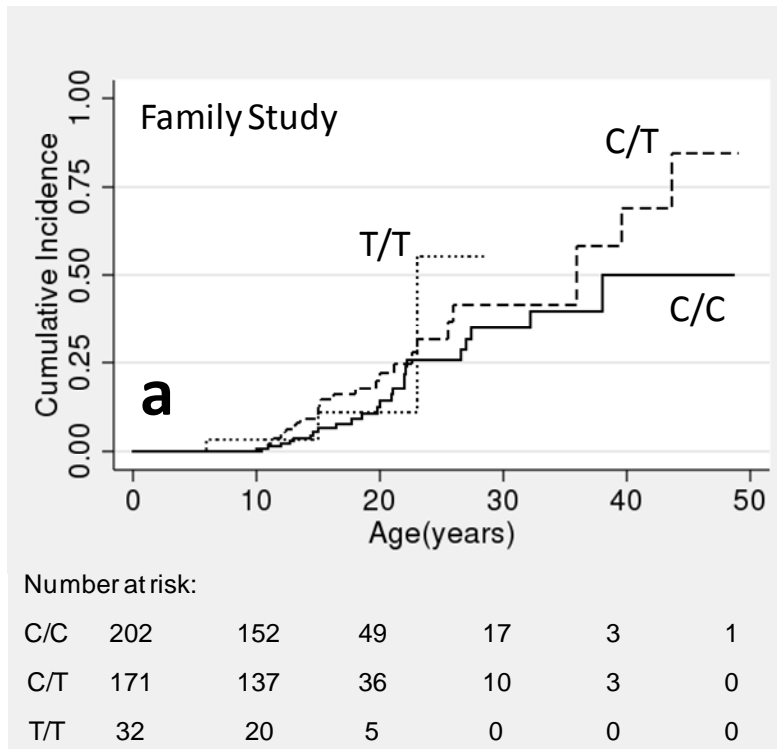
^aMonozygous twins excluded. Both study groups did not deviate from Hardy-Weinberg equilibrium (Family study, $p=0.7$, calculated using founder genotypes; Case-control study, $p=1.0$). As noted, individuals without diabetes younger than 10 years of age were excluded due to low age-dependent risk for diabetes.

^bPer-allele odds ratio and P-value determined by logistic regression in comparison with at-risk (age \geq 10) Euglycaemic controls from the Family study, or the non-diabetes Old Mild (\geq 34 yrs) group from the Case-control study.

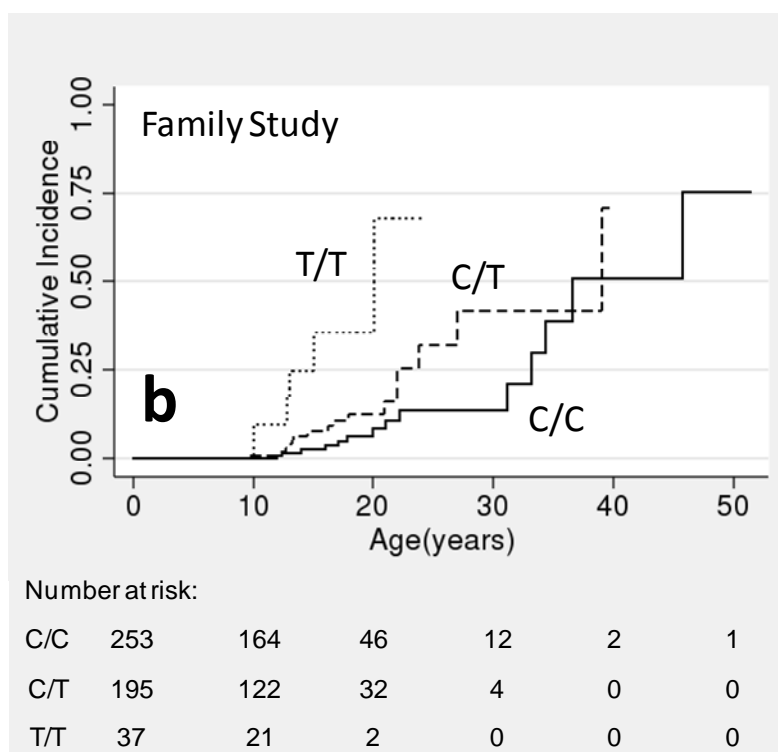
^cP-value for difference calculated by Fisher exact test for genotype counts (grouped according to additive (Add.) or dominant (Dom.) genetic models) and allele counts. Genotype counts were not significantly different for any group using recessive models. Counts in the Family study are compared to at-risk Euglycaemic (age \geq 10 years); counts in the Case-control study counts are compared to non-diabetes Old Mild (\geq 34 yrs) group.

ESM Table 2. Association of *TCF7L2* with diabetes after adjusting for steroid treatment (n=794).

Covariate	OR	95% CI	P-value
<i>TCF7L2</i> genotype (OR per allele)	2.6	1.5-4.4	0.001
Steroid treatment in the last year	2.0	1.0-4.0	0.05
Interaction term (steroids)	0.52	0.3-1.06	0.07



ESM Figure 1(a). Cumulative incidence of diabetes in 405 patients treated with systemic steroids within the last year, grouped by *TCF7L2* SNP genotype. Variation of *TCF7L2* did not significantly associate with the cumulative incidence of diabetes in the subset of patients in the primary (family) study who were treated with systemic steroids ($P=0.2$, log rank). No significant effect was seen by regression for an additive genetic model (hazard ratio = 1.4 per allele, 95% CI, 0.9-2.1; $p=0.1$), dominant model (HR = 1.6, 95% CI, 0.9-2.7, $p=0.09$), or recessive model (HR = 1.1, 95% CI 0.3-3.6, $p=0.8$).



ESM Figure 1(b). Cumulative incidence of diabetes in 485 patients not treated with systemic steroids, grouped by *TCF7L2* SNP genotype. Variation of *TCF7L2* was correlated with increased cumulative incidence of diabetes in the primary (family) study. Individuals homozygous for the major allele (C/C genotype, solid line) developed diabetes at a significantly lower rate than patients heterozygous (C/T genotype, dashed line) or homozygous for the minor allele (T/T genotype, dotted line). Results were similar when comparing all 3 genotypes separately as shown (additive genetic model; $p=3.0 \times 10^{-7}$, log rank) or when grouping genotypes according to dominant ($p=0.005$) or recessive ($p=1.8 \times 10^{-6}$) genetic models. Modelling of these data indicated that each T allele conferred about a 190% increase in risk of developing diabetes compared to PI CF patients of the same age without a T allele (HR = 2.9, 95% CI 1.7-4.9; $p=1.1 \times 10^{-4}$).