A comparison of methods for identification of genetic variants related to age-of-onset of cystic fibrosis related diabetes

Hua Ling[1], Peng Zhang[1], Elizabeth W. Pugh[1], Melis Atalar[2], Scott M. Blackman[3]

[1]*Center for Inherited Disease Research, Johns Hopkins University, United States of America*
[2]*The McKusick-Nathan Institute of Genetic Medicine, Johns Hopkins University, United States of America*
[3]*Division of Pediatric Endocrinology, Johns Hopkins University, United States of America*

Introduction

Cystic Fibrosis (CF) is a monogenic disorder that affect more than 80,000 people worldwide. Diabetes (CF-related diabetes or **CFRD**) is the most common extrapulmonary complication of CF and affects > 40% people with CF by adulthood with a broad range of age of onset (AOO). Variation in CFRD risk has been shown to be heritable and *Blackman, et al*[1] identified five loci associated with CFRD onset, analyzed as time to event while excluding related individuals in a total of 3,059 individuals. To better account for relatedness in survival analyses, we investigated different strategies using a family subset of the above data, the CF Twin and Sibling study, which includes 396 samples from 288 small families (siblings and half siblings) genotyped on the Illumina 610-Quad. The sample size and statistical power were deliberately limited in these analyses in order to test adequate control of type 1 error.

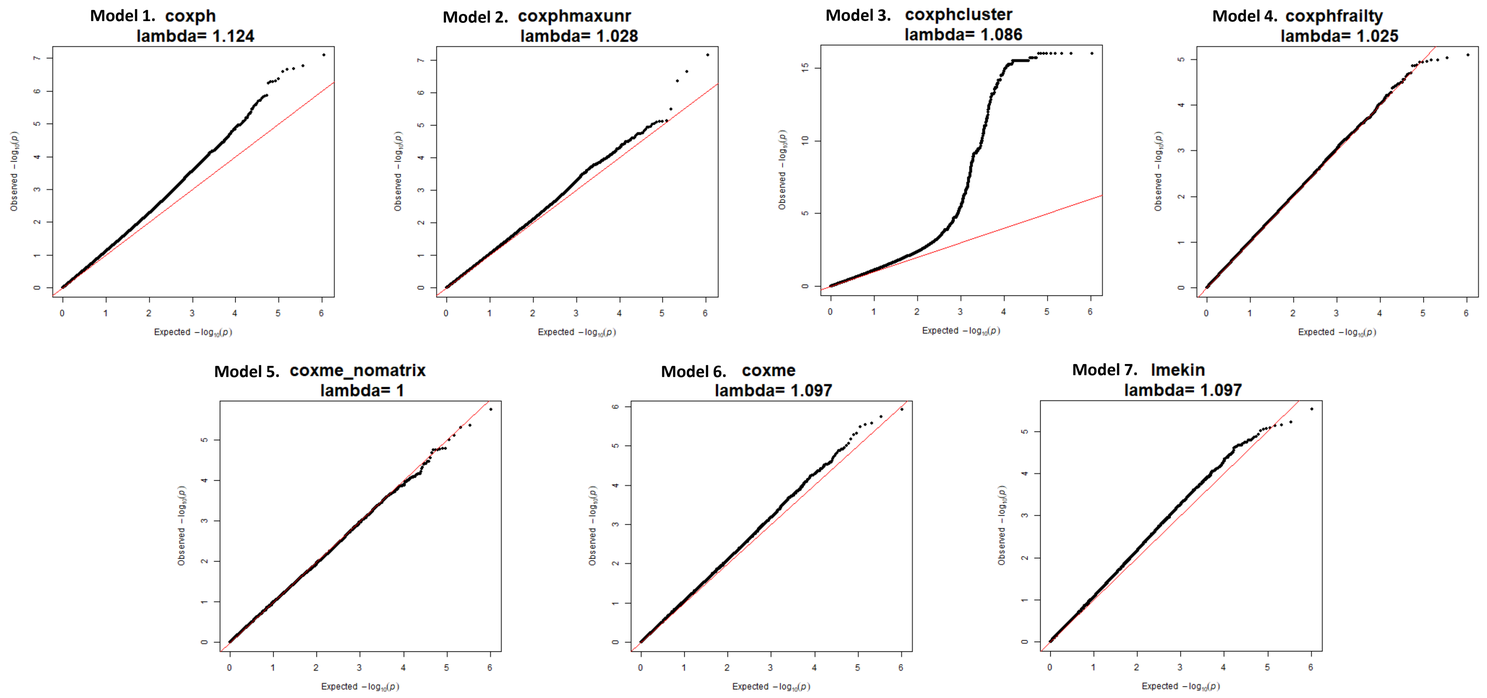
Methods

Study samples: 396 samples (288 families)
Genotyping Array: Illumina 610-Quad
Dependent variables: Time to event (CFRD) (Model 1-6) **OR** Martingale residual w/8PCs (Model 7)
Covariates: 8 PCs
Kinship coefficient: KING [2]
Statistical models : AOO for CFRD ~ SNP + 8PCs + …

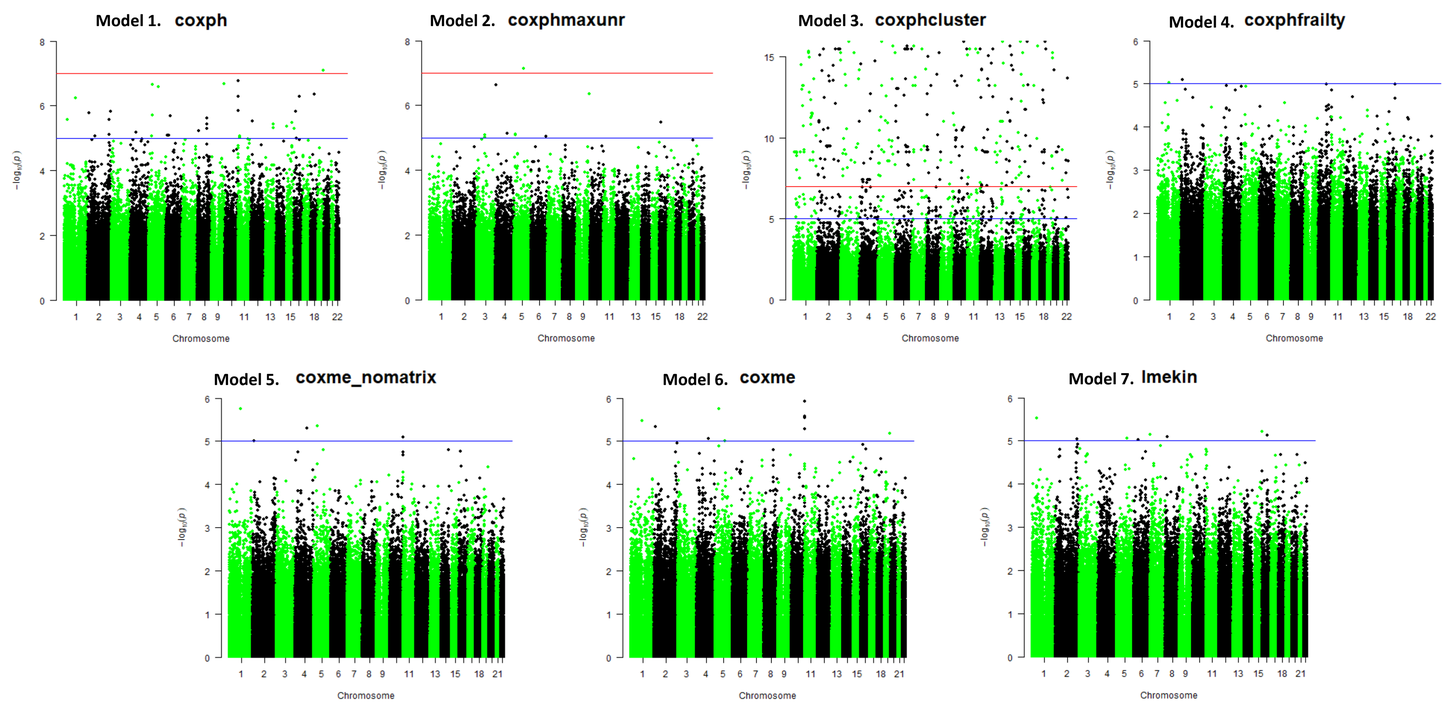
1. Coxph with everybody
2. Coxph with maximum unrelated
3. Coxph with marginal model (cluster)
4. Coxph with frailty model
5. Mixed effects Cox model with family-specific random intercept [3]
6. Mixed effects Cox model with kinship coefficient matrix as random intercept
7. Mixed effects model using martingale residual with kinship coefficient as random intercept

Results

QQ Plots



Manhattan Plots



Summary

1. Mixed effect Cox model with family-specific random effect(model 5) and Coxph with frailty (model 4) perform best in controlling false discovery rate.
2. For mixed Cox PH model, using calculated kinship coefficient did not show an improved lambda (model 6 vs 5) in this twin/sibling dataset.
3. Cox PH with marginal model showed huge inflation and is not recommended.
4. Using Martingale residuals in mixed model with kinship coefficient matrix showed equivalent level of lambda as mixed model with AOO CFRD.

Reference:

1. Blackman SM, et al. Genetic modifiers of cystic fibrosis-related diabetes. *Diabetes*. 2013 Oct;**62**(10):3627-35. PMID:23670970
2. Manichaikul A, et al Robust relationship inference in genome-wide association studies. *Bioinformatics*. 2010; **26**(22):2867-2873 PMID: 20926424
3. Therneau T, *https://CRAN.R-project.org/package=coxme*

For suggestions or questions, please email hling1@jhmi.edu.