Exome Sequencing for Discovery of Genetic Modifiers of Susceptibility to Pseudomonas Infection in Cystic Fibrosis

Presented by Mary Emond

Exome Sequencing Project Meeting

March 28, 2012

Outline

- Review of Phase I main result
- Validation study results for Phase I top candidate
- Additional Phase I results
- Phase I + Phase II results
- Future directions

Phase I (1)

- Extreme phenotypes: early chronic Pa versus never Pa by age 15 or older
- N=43 and 48, respectively
- Chronic: two consecutive Pa+ quarters
- Early: less the 6 YO at onset (< 10^{th} percentile)
- 472 and 2652 person-years of observation
- Overall Pa+ frequency: > 28% vs. 0.07%

Phase I (2)

- QC: Standard GATK filters (except SB), HWE, Ti/Tv, Het/Hom, ns/ss, missingness, singletons, mean depth, % of sites with ≥8x coverage
- Nothing remarkable (except large number of singletons in one exome)
- No differences between cases and controls.

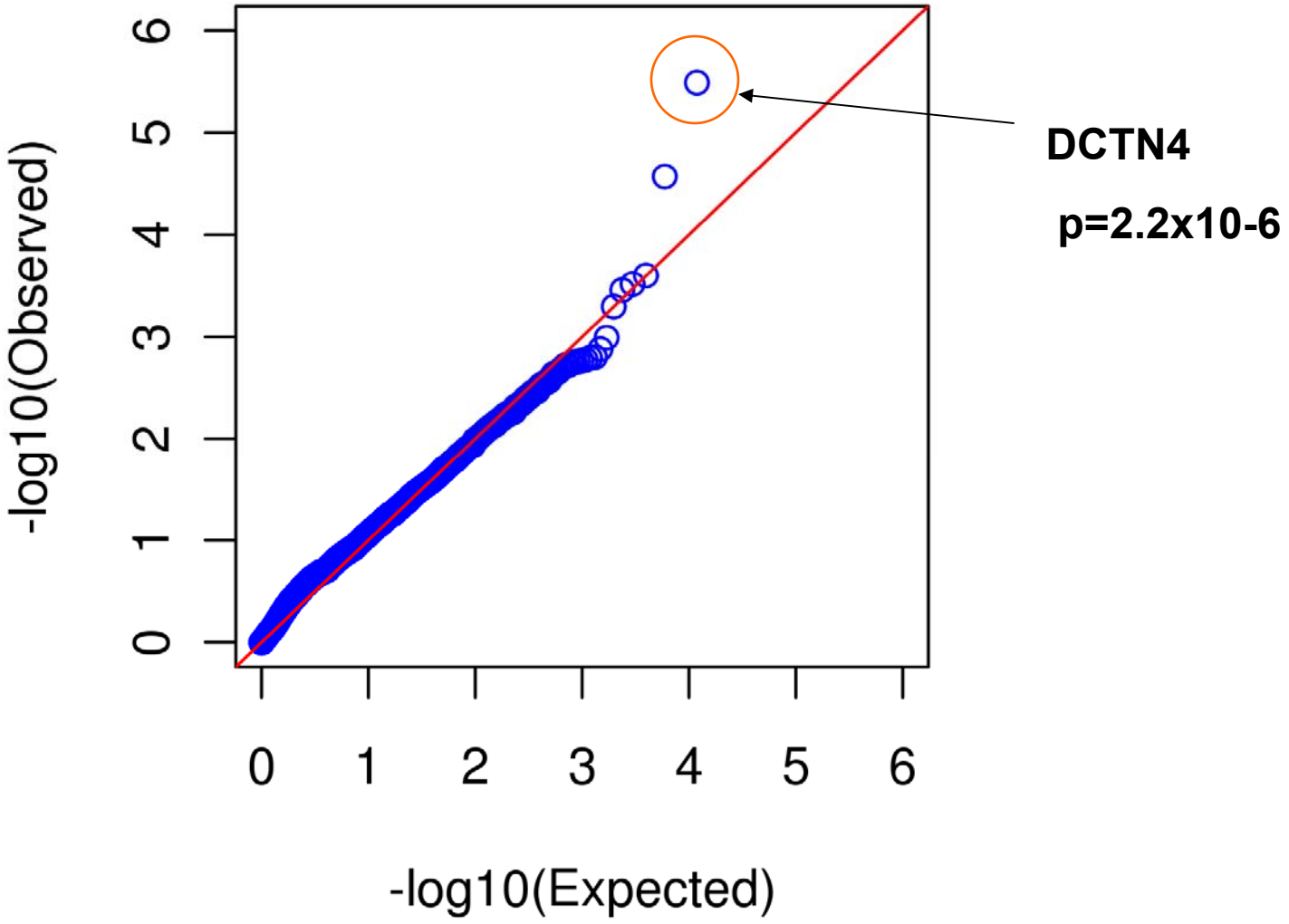
Phase I (3)

- Top hit using MZ method for collapsing and regression: DCTN4 (reported previously)
- Improved adjustments/QQ-plot
- Resampling-based p-value
- Consistency of MZ result with other methods, different MAF threshold

Chronic Pa (n=43) vs. Never Pa (n=48)

MZ logistic regression, T=0.125

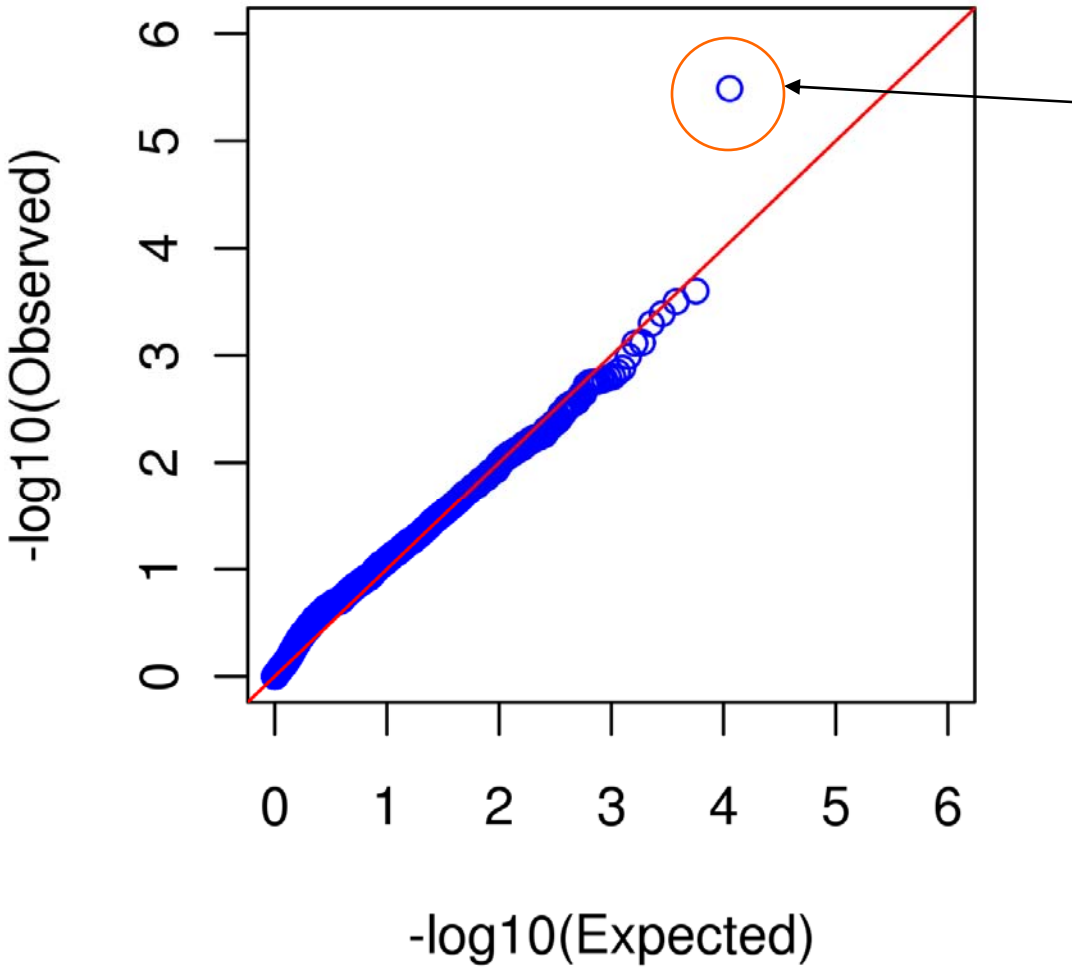
Covariates: PCs 1, 2, 18 and risk group



Chronic Pa (n=43) vs. Never Pa (n=48)

MZ logistic regression, T=0.05

Covariates: PCs 1, 2, 18 and risk group



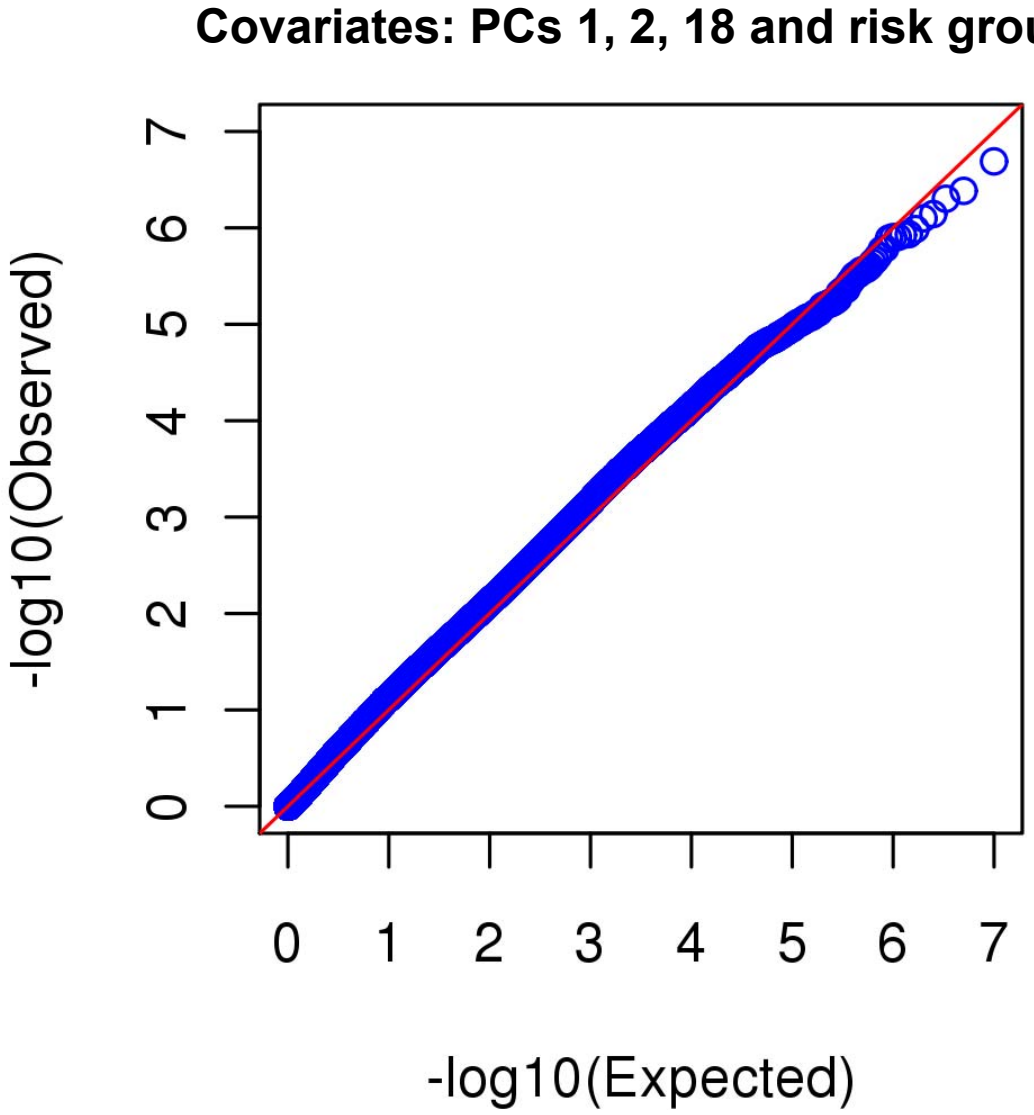
DCTN4

p=2.2x10-6

DCTN4 – sites creating "signal"

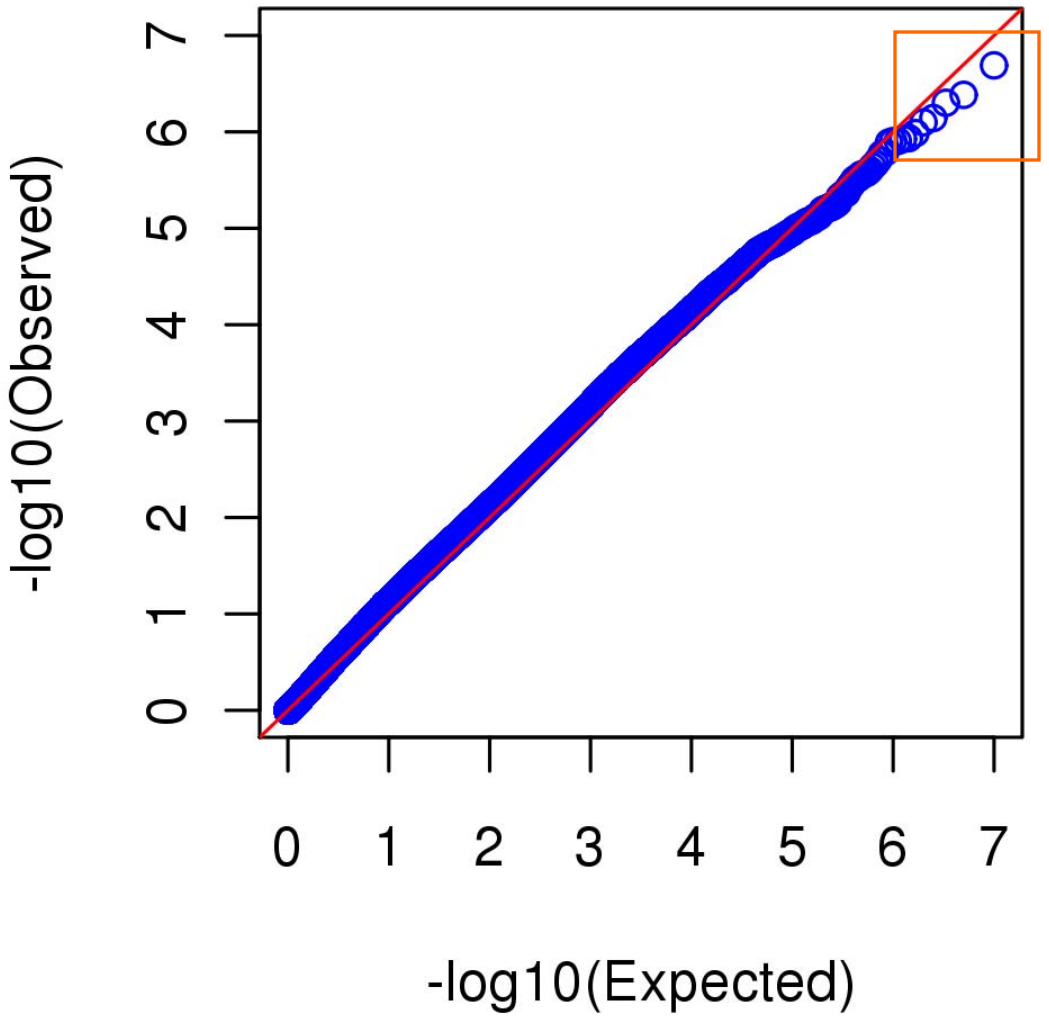
Site	RS	MAF	GERP	type	N cases	N cont
1500090432	35772018	0.02	5.3	missense	3	0
150087076	11954652	0.04	4.1	missense	9	0
				Total variant	12	0
				Total Non-variant	31	48

Distribution of MZ test statistic under the null hypothesis of no association of either SNP with outcome, sampling from the joint distribution of the two SNPs (10M trials)



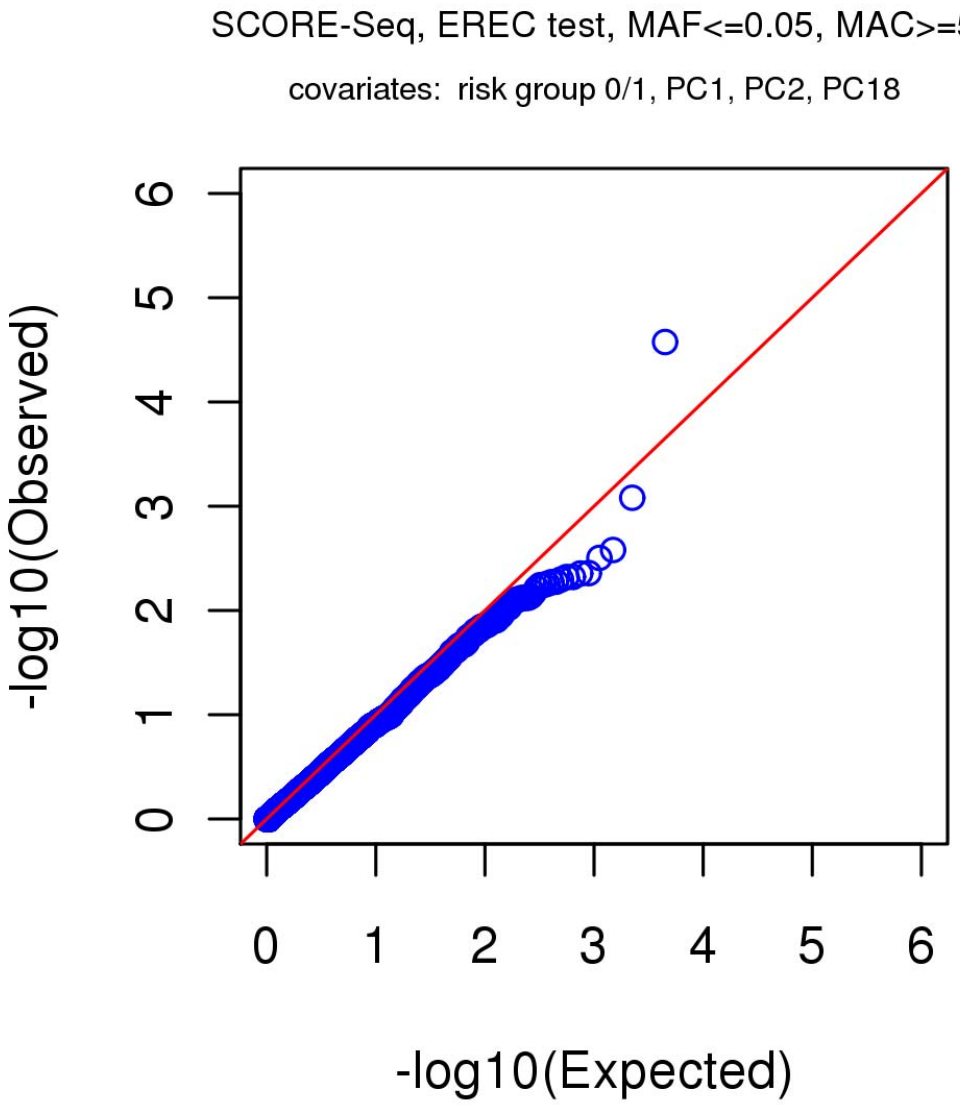
Distribution of MZ test statistic under the null hypothesis of no association of either SNP with outcome, sampling from the joint distribution of the two SNPs (10M trials)

Covariates: PCs 1, 2, 18 and risk group



15 of 10M trials result in $|t| \geq$ observed, or $p=1.5 \times 10^{-6}$

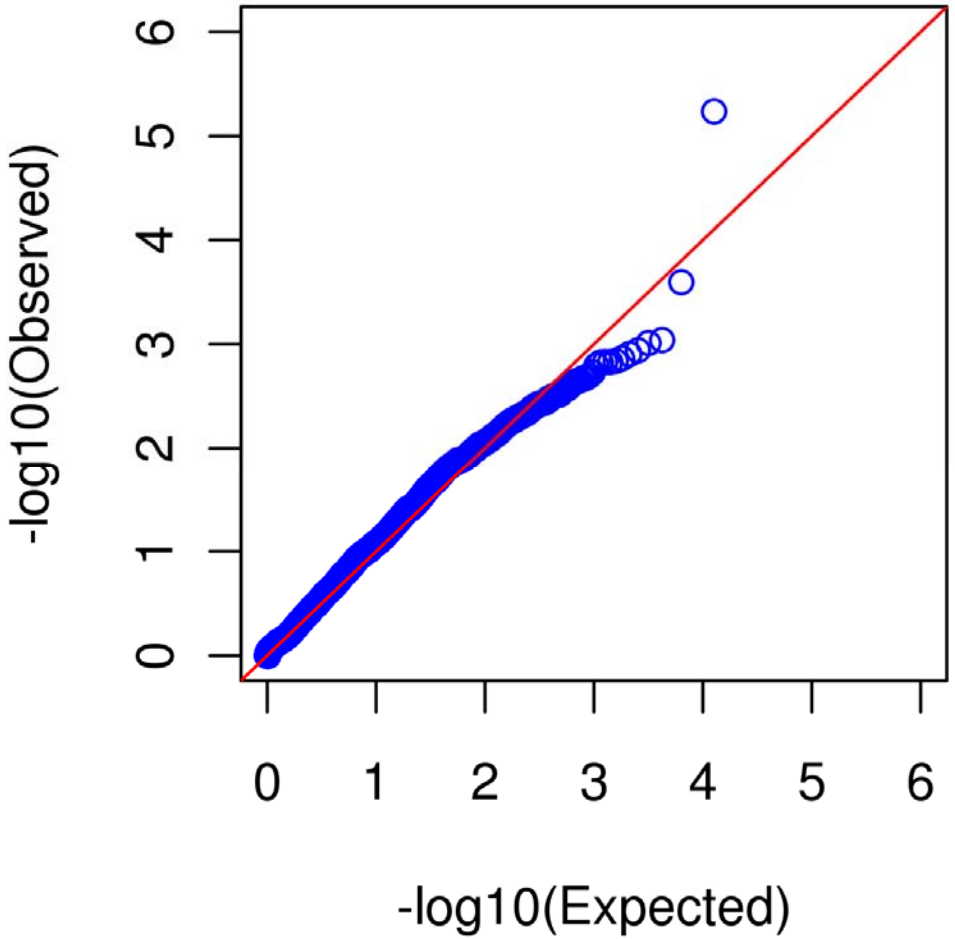
Chronic Pa (n=43) vs. Never Pa (n=48)



Chronic Pa (n=43) vs. Never Pa (n=48)

SKAT-O with small sample adjustment

Covariates: PCs 1, 2, 18 and risk group



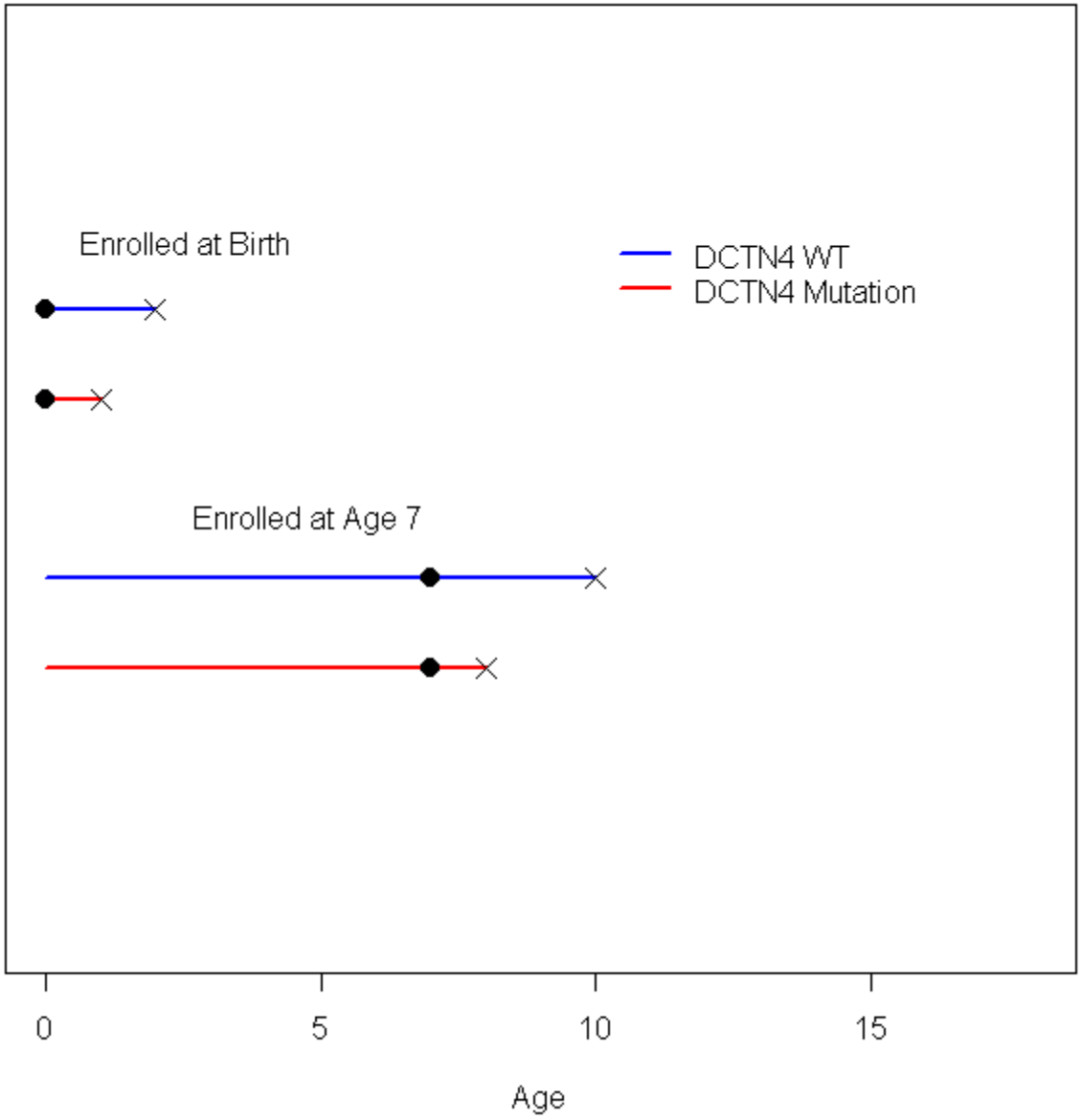
Correction for multiple testing

- Bonferroni: 2.2E-6 x 11,542 = 0.02
- Bonferroni seems conservative; test statistics are not distributed the same for all genes; p-values are not uniform
- There are only 3026 genes with a minor allele count of 8 or more. Should we only test genes that have a minor allele count > some cutoff?

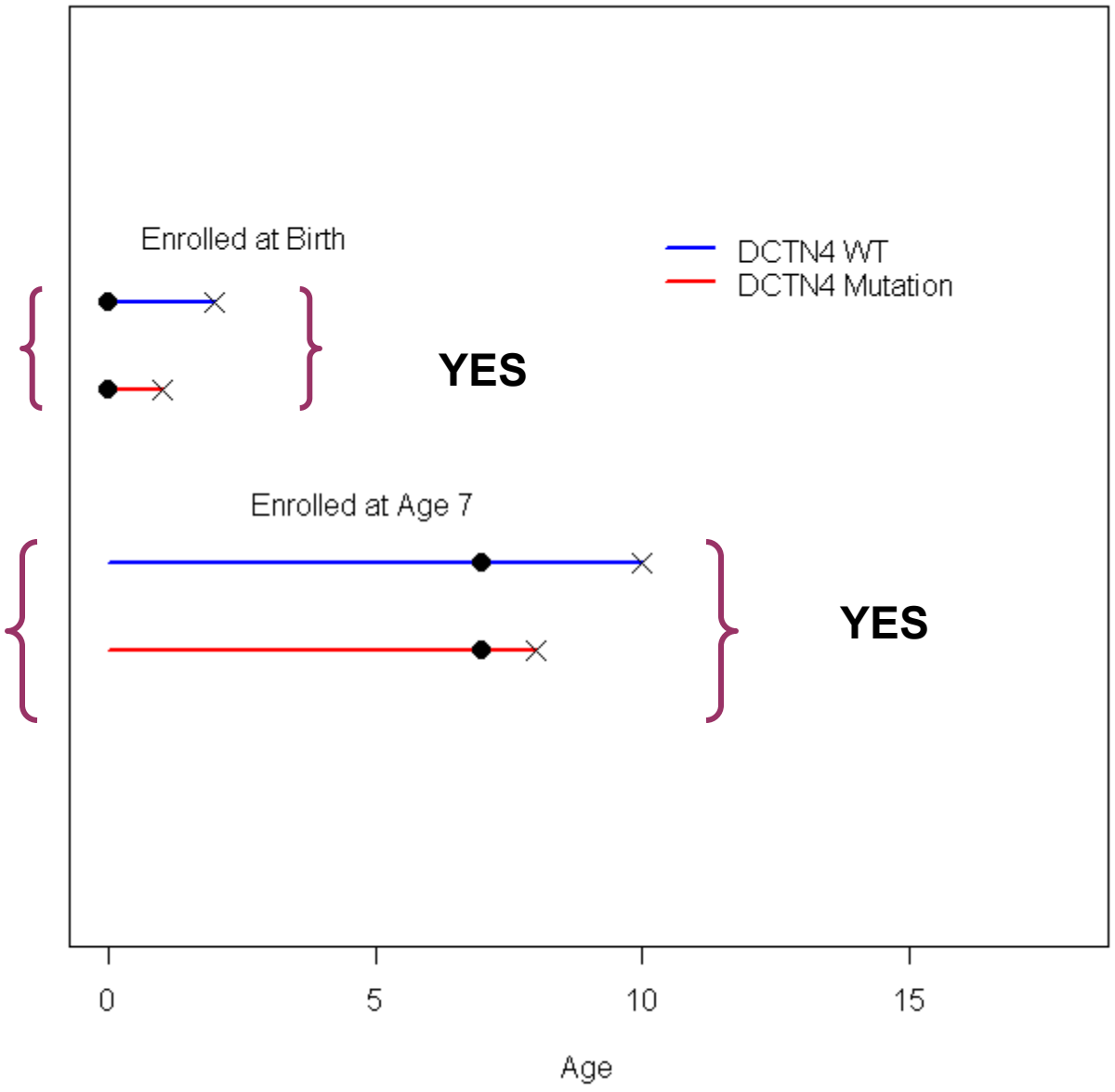
Resequencing of DCTN4 in Validation Sample

- N=676 EPIC study individuals enrolled under the *criterion of no previous Pa+ cx*
- For whom DCTN4 was successfully seq'n at both positions found in exome study
- Analysis: Cox model with Age at Onset of Chronic Pa as primary outcome, stratified on enrollment age
- Genetic variables: (1) indicator of either or both variants; (2) genotype indicators

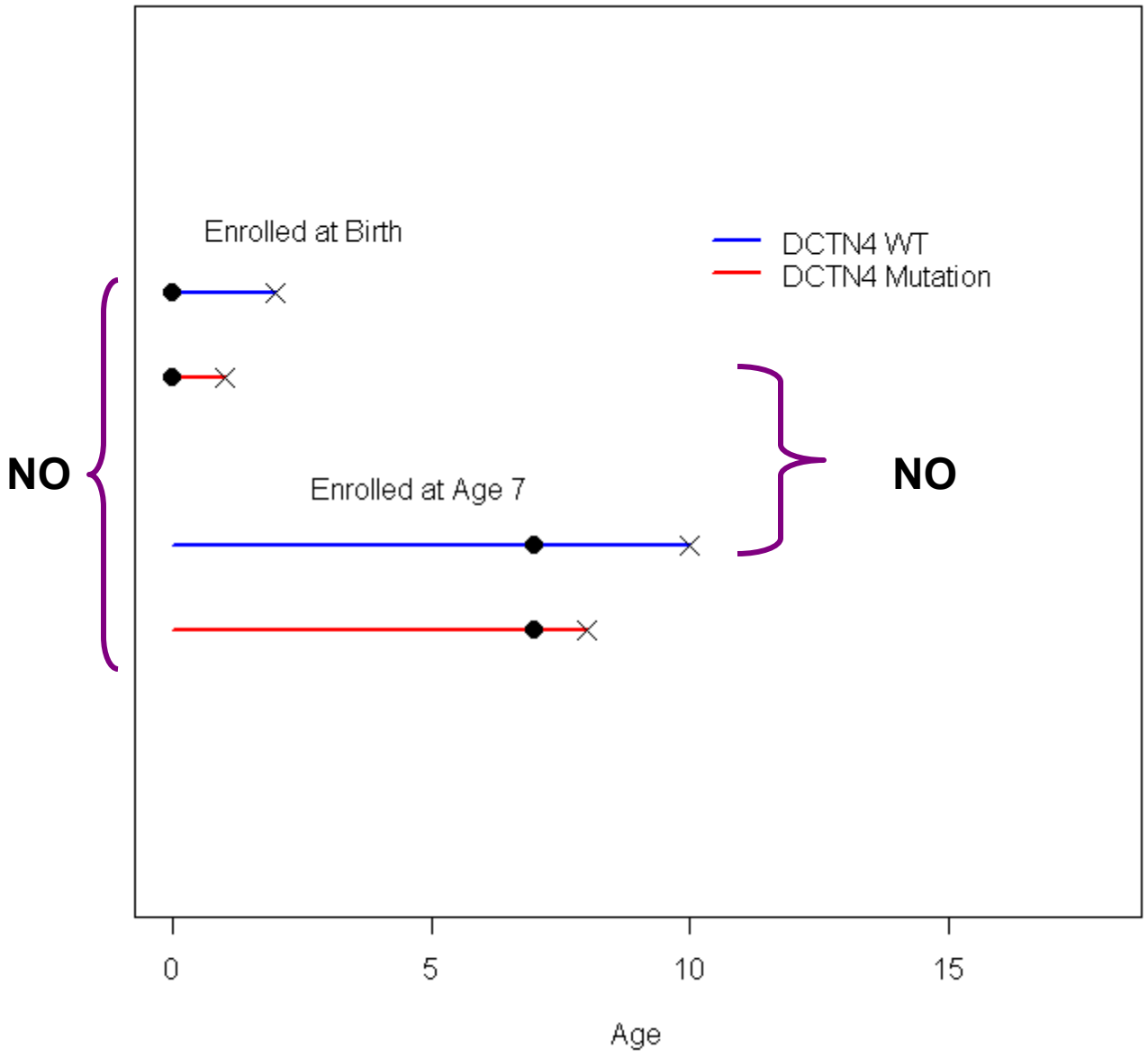
Stratification on Enroll Age Provides the Appropriate Comparisons



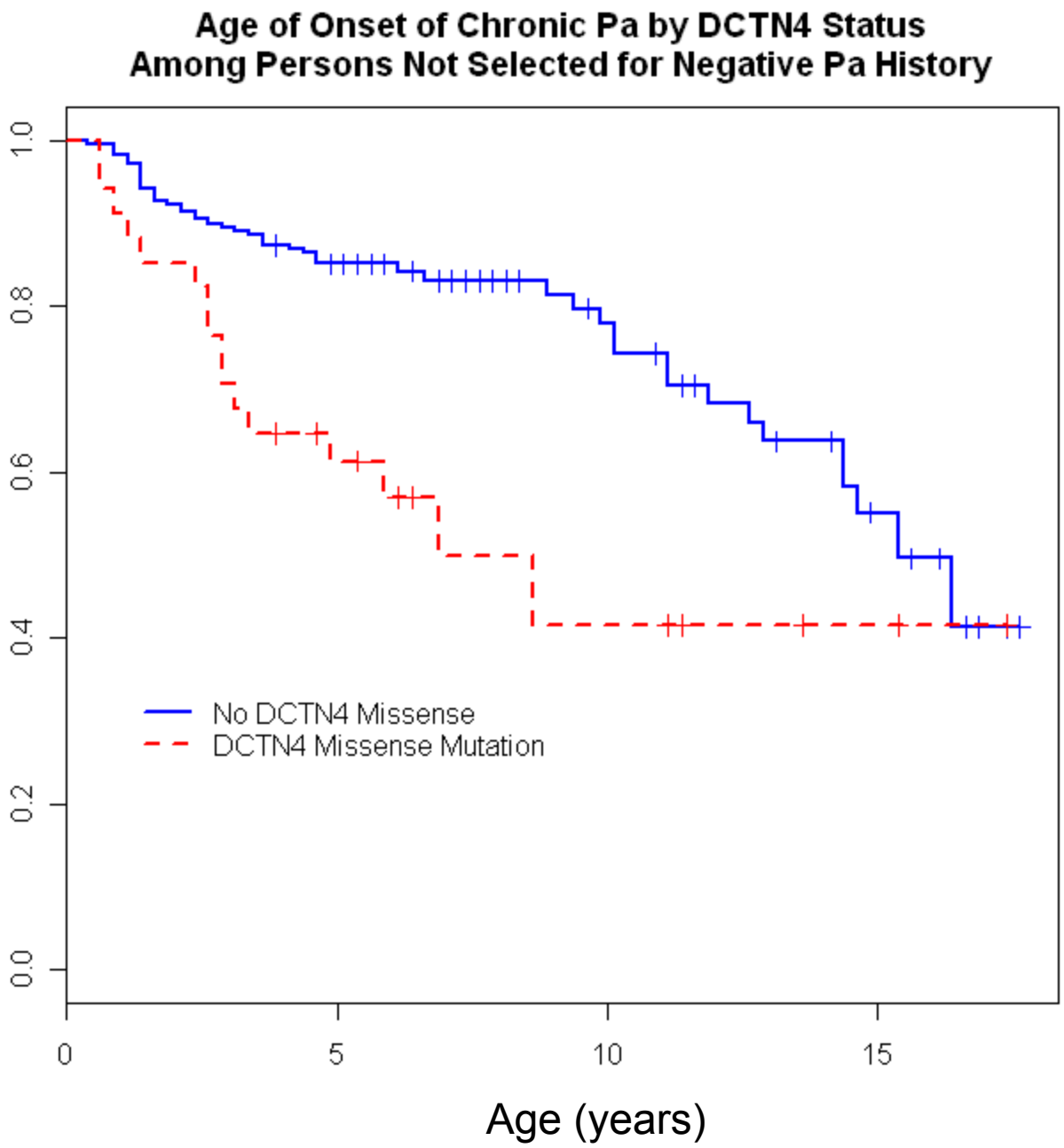
Compare Age at Onset Within Strata



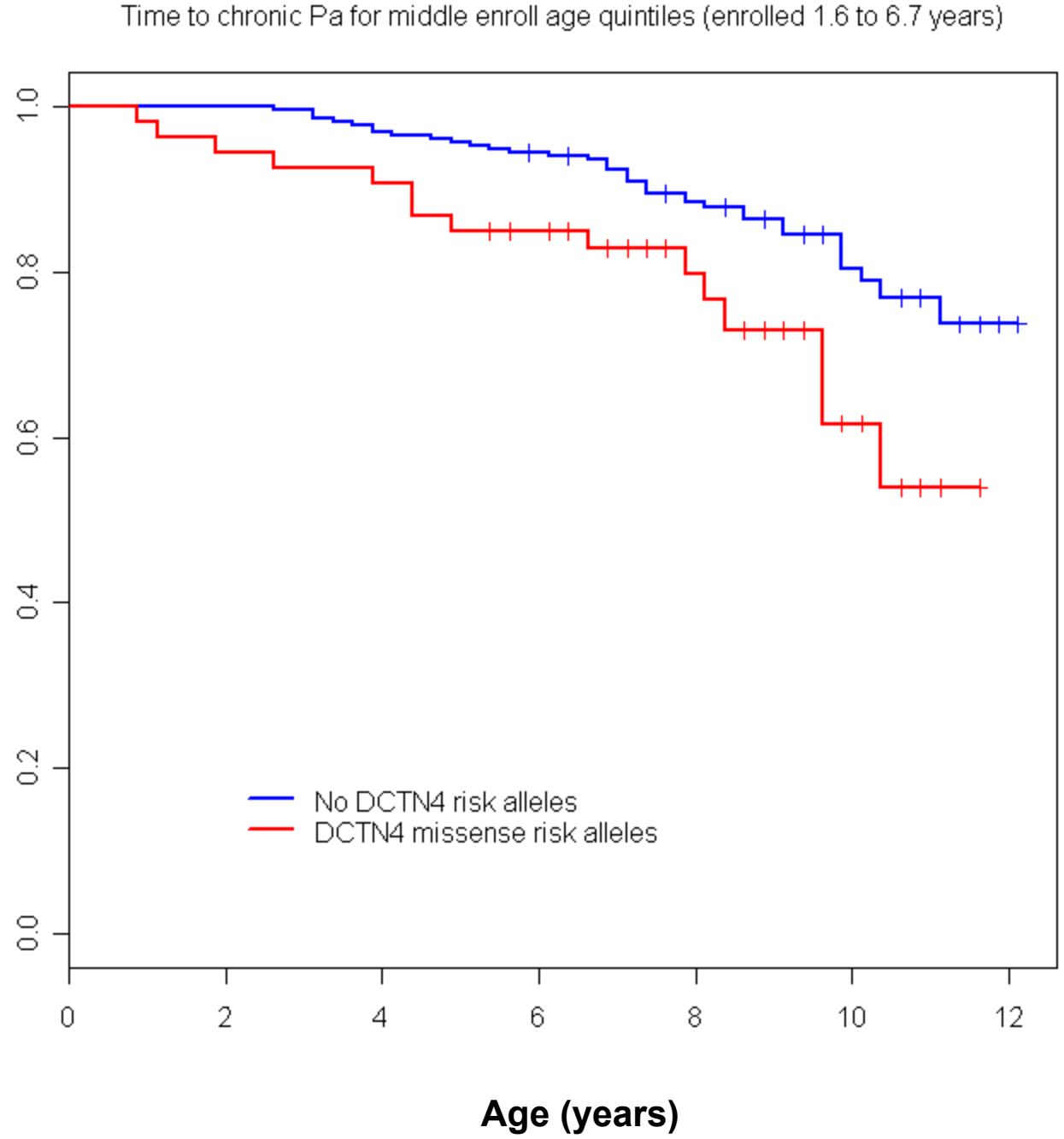
No comparisons across strata

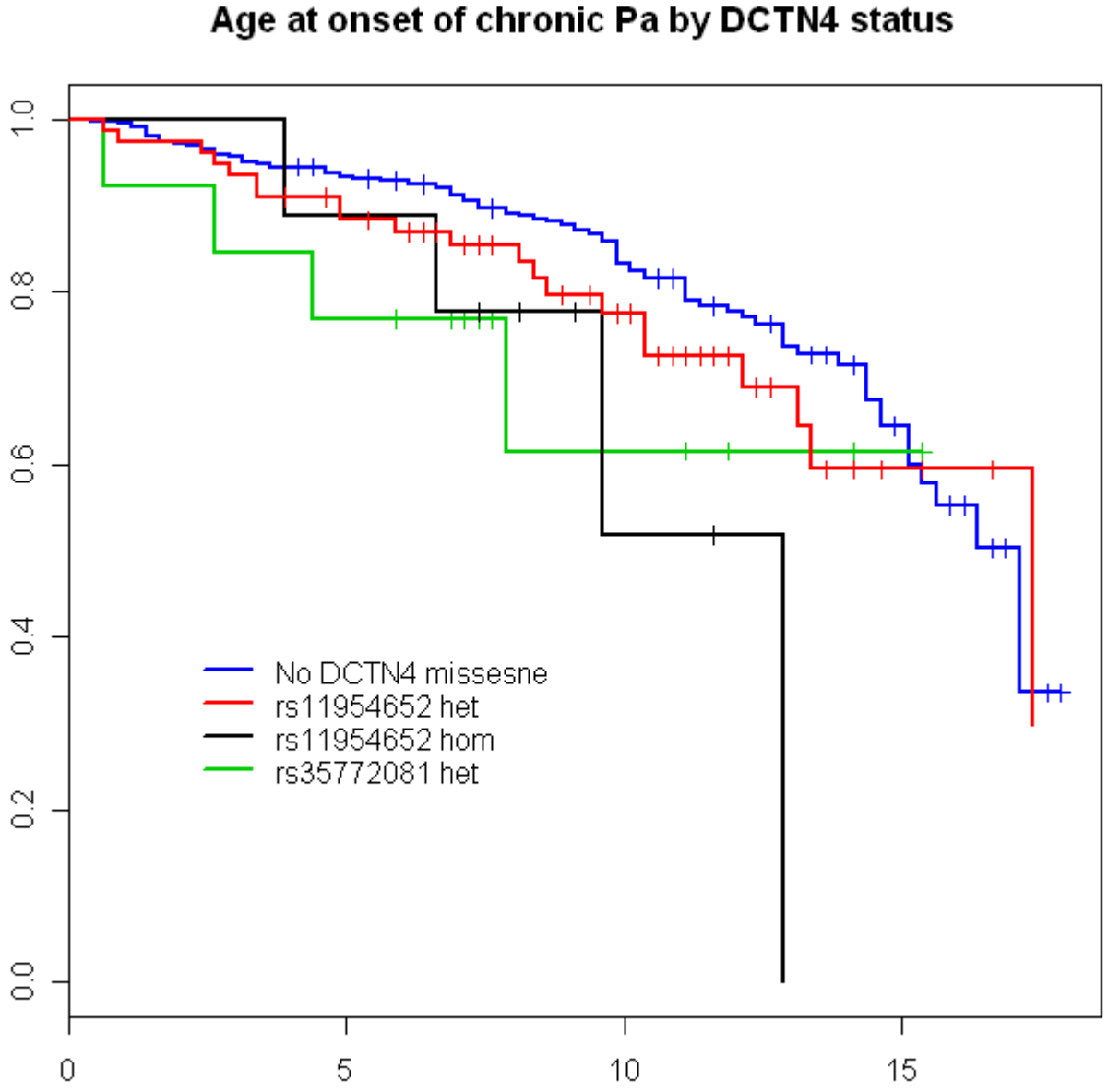


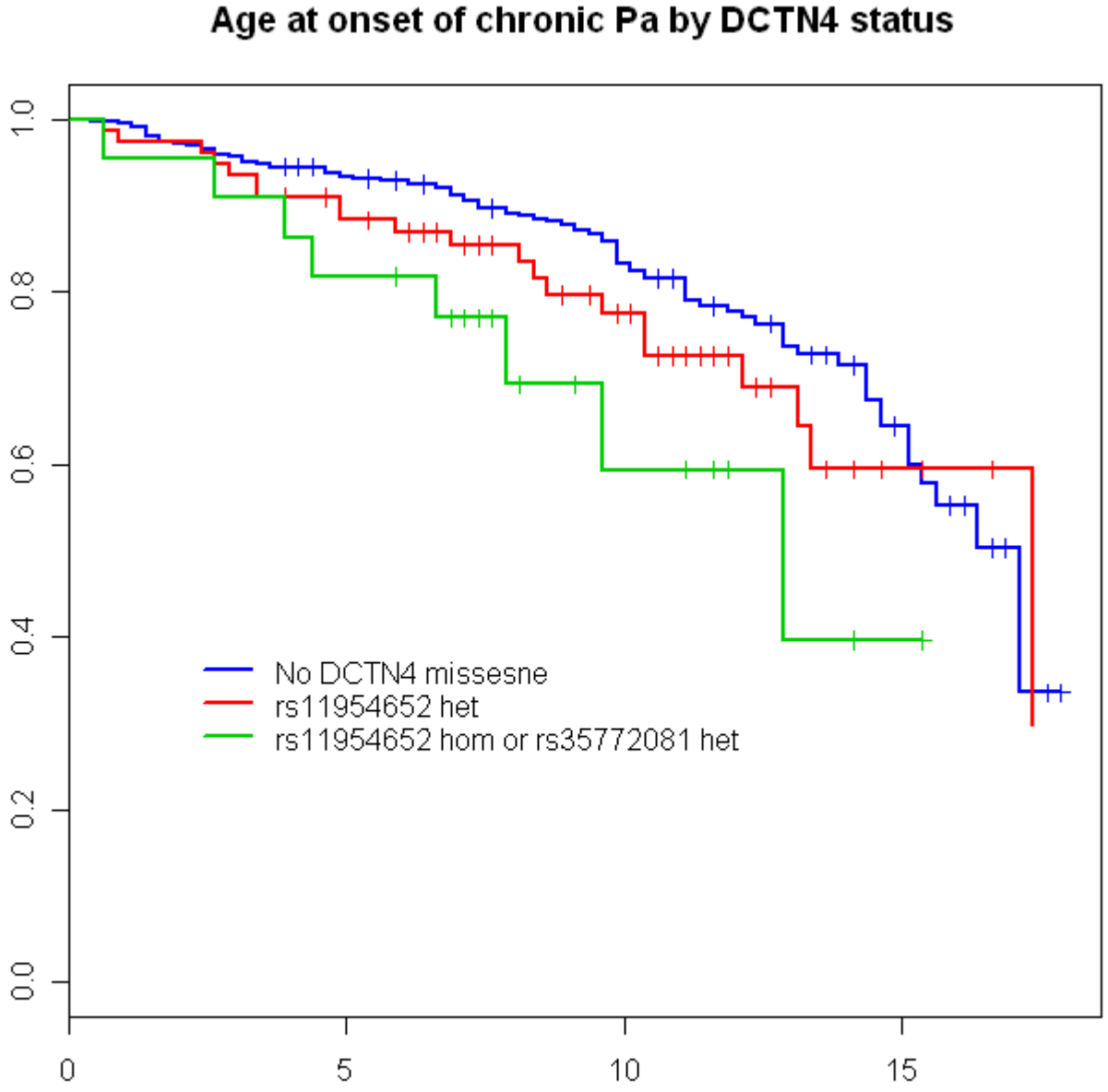
Outcome	Comparison	N	N Fail	HR	P	LB	UB
AOO chronic	no missense	565	93	1.0	--	--	--
	any missense	102	28	1.9	0.004	1.2	2.9
AOO chronic	no missense	565	93	1.0	--	--	--
	rs11954652 het	78	20	1.6	0.050	1.0	2.7
	rs11954652 hom	9	4	5.2	0.002	1.9	15
	rs35772018 het	13	4	15.9	0.002	2.8	91
AOO chronic	no missense	565	93	1.0	1.000	--	--
	rs11954652 het	78	20	1.7	0.050	1	2.7
	rs11954652 hom or rs35772018 het	22	8	3.3	0.002	1.5	6.9



(On study before age 1.6)


Time to chronic Pa for middle enroll age quintiles (enrolled 1.6 to 6.7 years)
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
10
12
No DCTN4 risk alleles
DCTN4 missense risk alleles
Age (years)

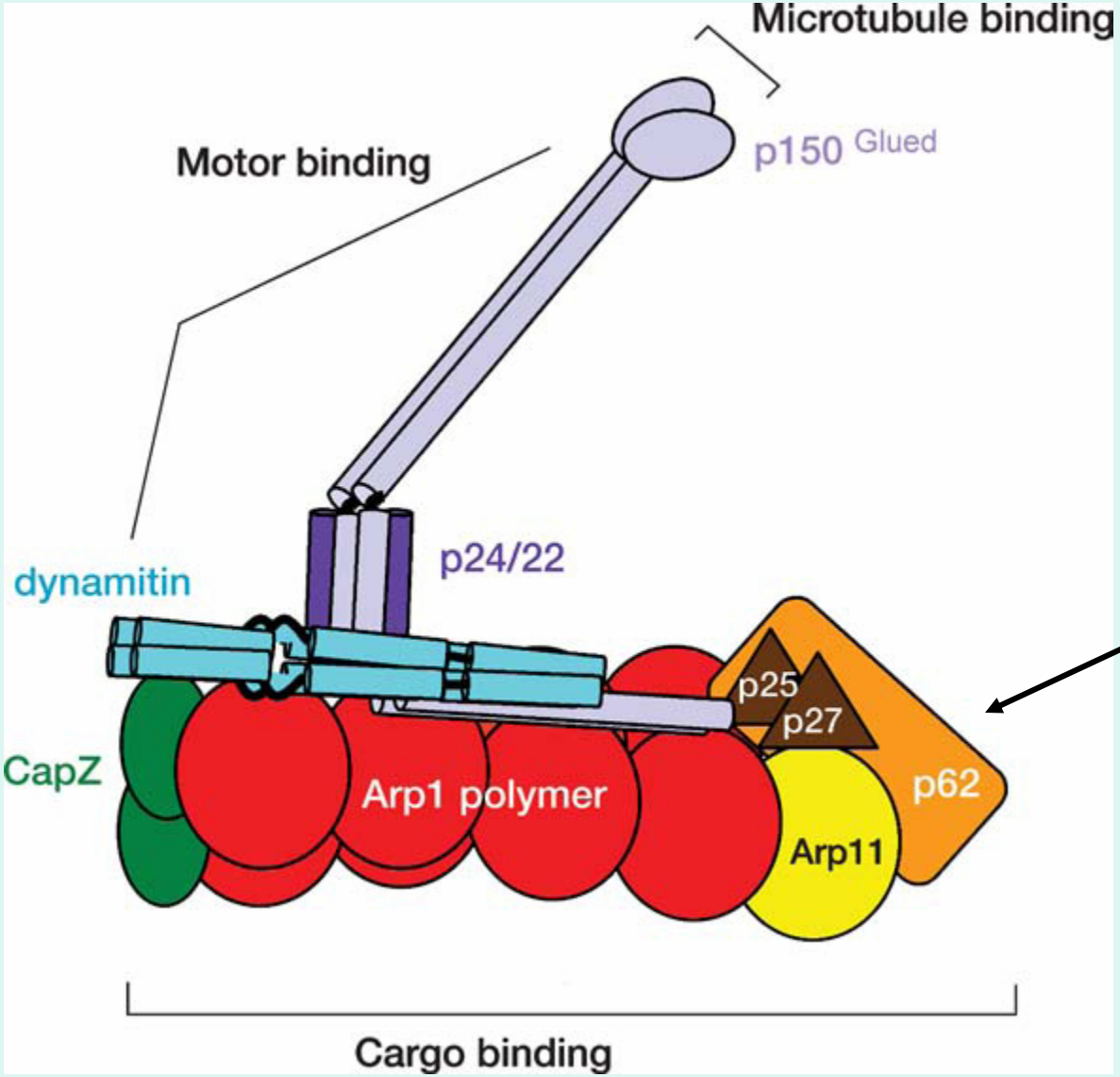

Age at onset of chronic Pa by DCTN4 status
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
No DCTN4 missesne
rs11954652 het
rs11954652 hom
rs35772081 het


Age at onset of chronic Pa by DCTN4 status
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
No DCTN4 missesne
rs11954652 het
rs11954652 hom or rs35772081 het

Association of DCTN4 variants with secondary outcomes

Event	Group	N	# events	HR	95% CI	p
Age at first Pa+ culture	no missense	565	345	1.0	--	--
	any missense	102	70	1.4	(1.1, 1.8)	0.01
Age of onset of mucoid Pa	no missense	565	84	1.0	--	--
	any missense*	102	16	2.6	(1.1, 5.9)	0.026
Time from first Pa to mucoid Pa	no missense	410	55	1.0	--	--
	any missense*	89	17	3.8	(1.4, 10.5)	0.01

Dynactin Complex



DCTN4 protein product

NHLBI Exome Sequencing Project (ESP)

National Heart Lung and Blood Institute
National Institutes of Health

Program Officer: Deborah Applebaum-Bowden

Heart GO	Lung GO	Women's Health Initiative Sequencing Program
PI: Stephen Rich University of Virginia	PIs: Michael Bamshad U. of Washington Kathleen Barnes, Johns Hopkins University	PI: Rebecca Jackson Ohio State University

UNIVERSITY of WASHINGTON

PIs: Debbie Nickerson, Mark Rieder, Jay Shendure, & Phil Green

BROAD INSTITUTE

PIs: Stacey Gabriel & David Altshuler

NHLBI Exome Sequencing Project (ESP)

National Heart Lung and Blood Institute
National Institutes of Health

Program Officer: Deborah Applebaum-Bowden

Heart GO	Lung GO	Women's Health Initiative Sequencing Program
PI: Stephen Rich University of Virginia	PIs: Michael Bamshad U. of Washington Kathleen Barnes, Johns Hopkins University	PI: Rebecca Jackson Ohio State University

UNIVERSITY of WASHINGTON

PIs: Debbie Nickerson, Mark Rieder, Jay Shendure, & Phil Green

BROAD INSTITUTE

PIs: Stacey Gabriel & David Altshuler