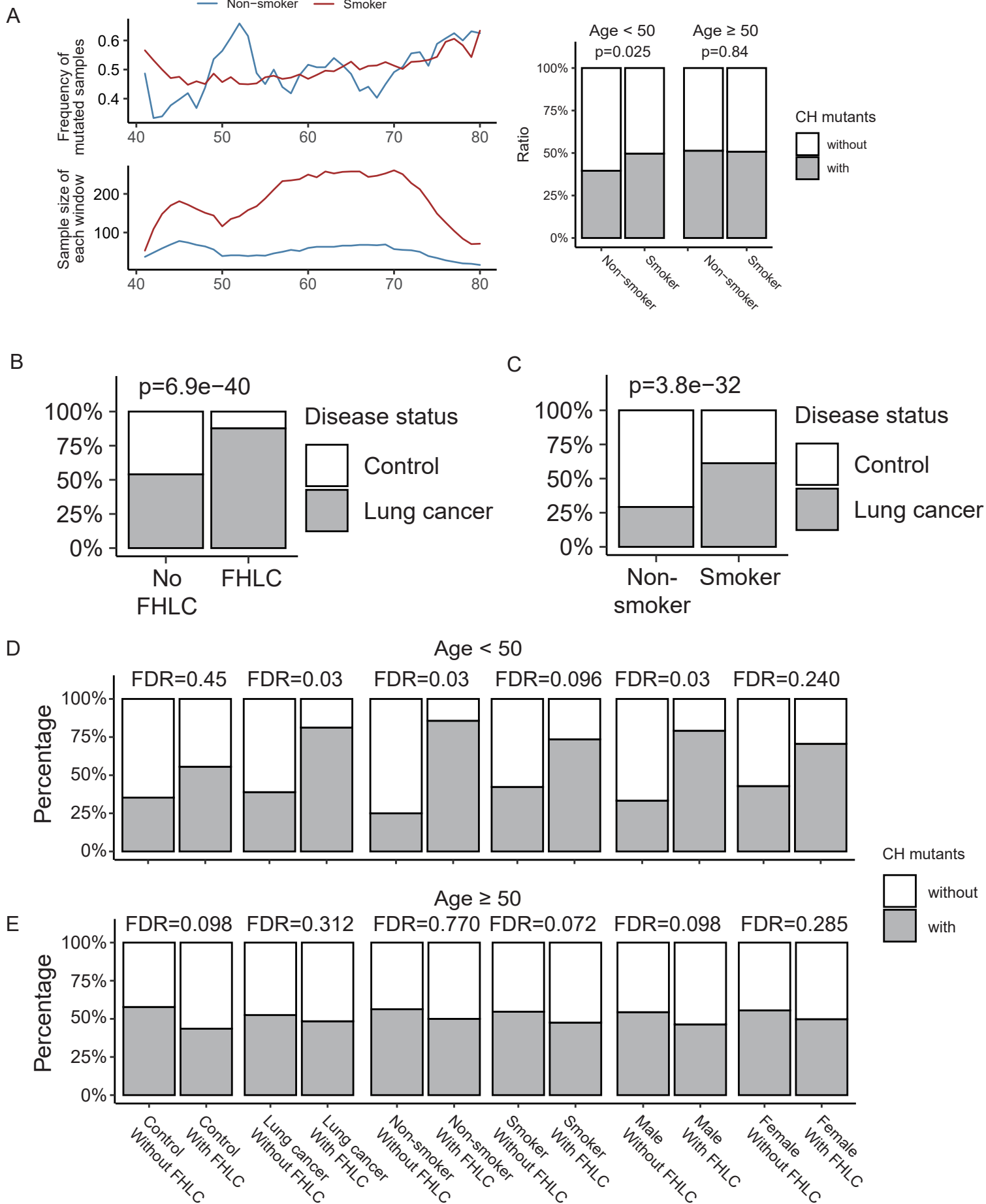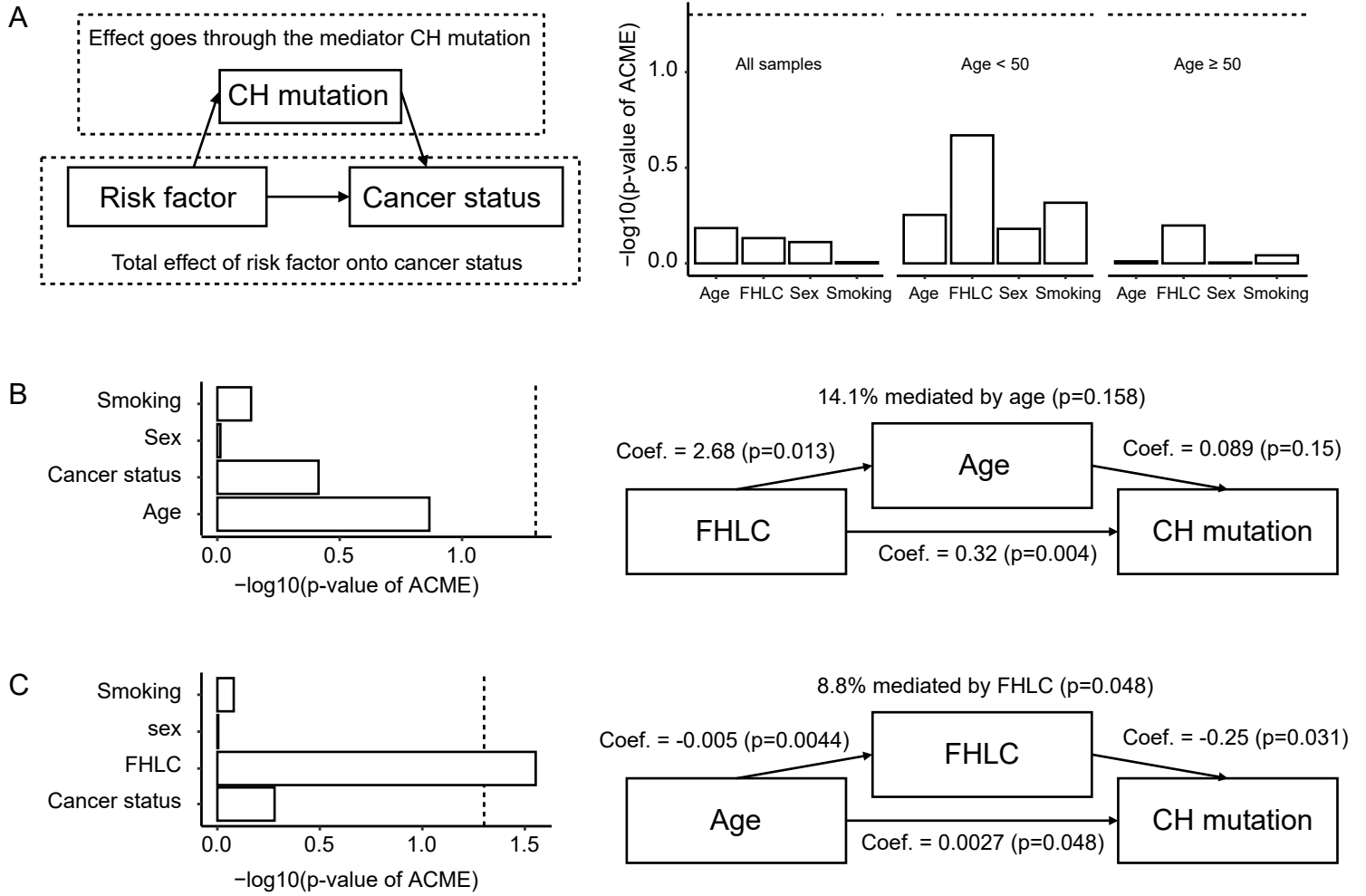# Figure S1

1   **Figure S1 CH mutations associated with smoking and FHLC.** (A) Smokers had more

2   CH mutations than non-smokers in younger age samples (age<50). Samples with either (B)

3   FHLC or (C) smoking history were more likely to be lung cancer patients. (D) Younger age

4   group subjects with family history of lung cancer tend to have significantly fewer CH

5   mutations than those without. (E) In contrast, the opposite trend was observed in the older

6   age group.

7

8

Figure S2

A

B

C

9   **Figure S2 Mediation test.** (A) CH as a mediator between a risk factor and cancer. Bar plot

10   showed the significance of average causal mediation effects (ACME) of CH in all the samples

11   and in young/old groups. Dashed line denoted significance cutoff 0.05. None of the correlation

12   between risk factor and cancer showed significant ACME of CH. (B) In young samples with

13   age < 50, age had the lowest p-values (p=0.156) of ACME than all the other risk factors, with

14   14.1% of correlation effects between FHLC and CH mutations mediated by age (p=0.158). (C)

15   In old samples with age ≥ 50, FHLC had significant ACME (p=0.028), with 8.8% of correlation

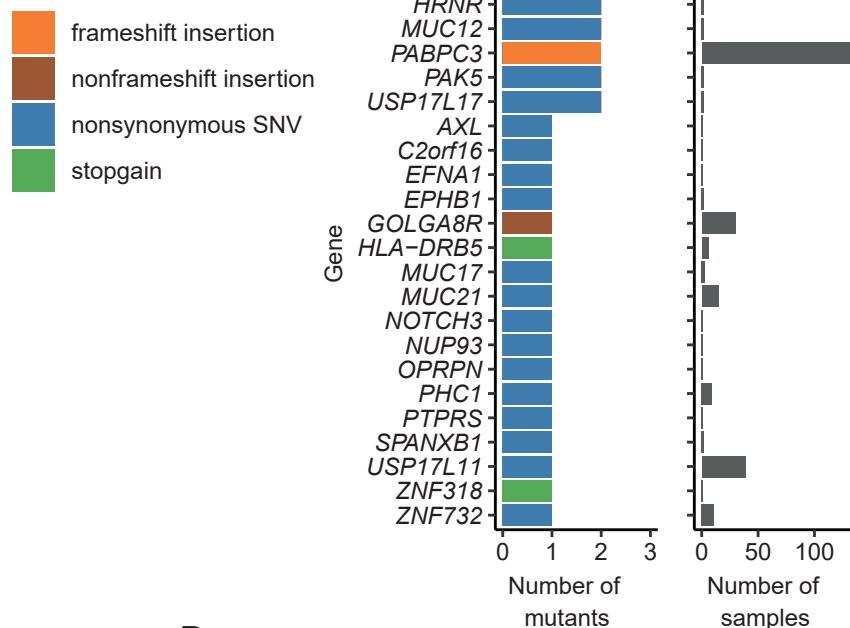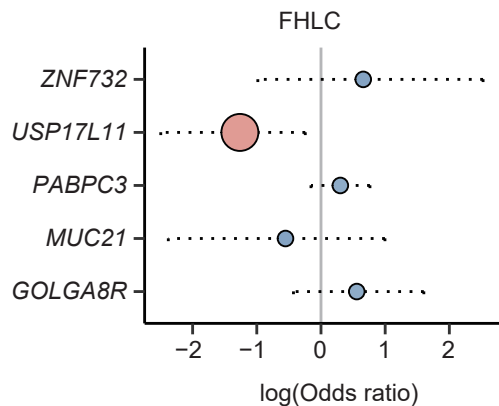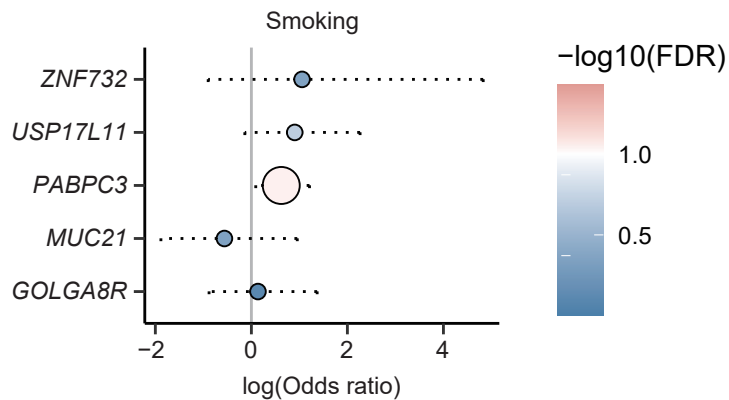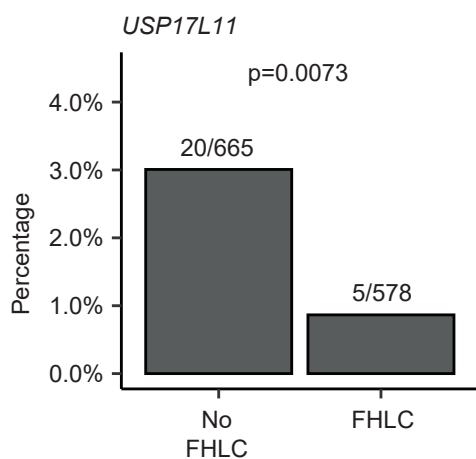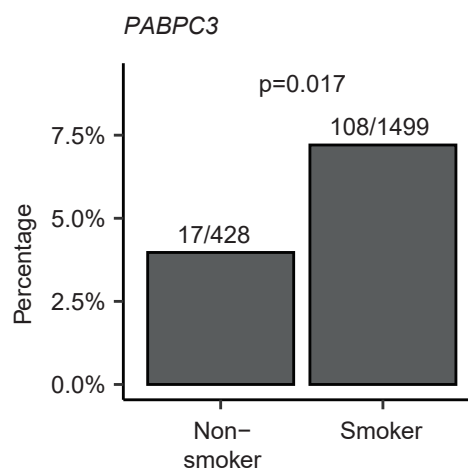16   effects between age and CH mediated by FHLC (p=0.048).

17

18

Figure S3
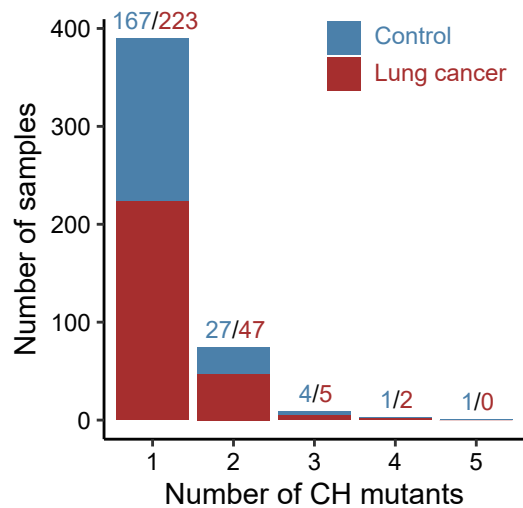
19 **Figure S3 Potential novel CH mutations in genes other than leukemia genes and**

20 **COSMIC cancer genes.** (A) Distribution of number of CH mutations in each sample. Most of

21 the samples have 1 CH mutations; none of the samples have more than 5 CH mutations in

22 these genes. Red and blue denoted numbers of CH mutations in lung cancer patients and

23 controls respectively. (B) Number and type of CH mutations in those genes. Most genes had

24 only one CH mutation in a few samples. PABPC3 and USP17L11 had the largest number

25 and highest frequency of CH mutations. Correlation between CH and (C) FHLC or (D)

26 smoking were tested in genes which were mutated in more than 10 samples. (E) Samples

27 without FHLC had more USP17L11 CH mutations. (F) Smoker samples had more PABPC3

28 CH mutations.

29

30

31 **Supplementary tables**

32 **Table S1 Clinical traits and sample size.**

33 **Table S2 SNPs significantly associated with CH mutations.**

34 **Table S3 eQTLs among CH associated SNPs.**

35

36