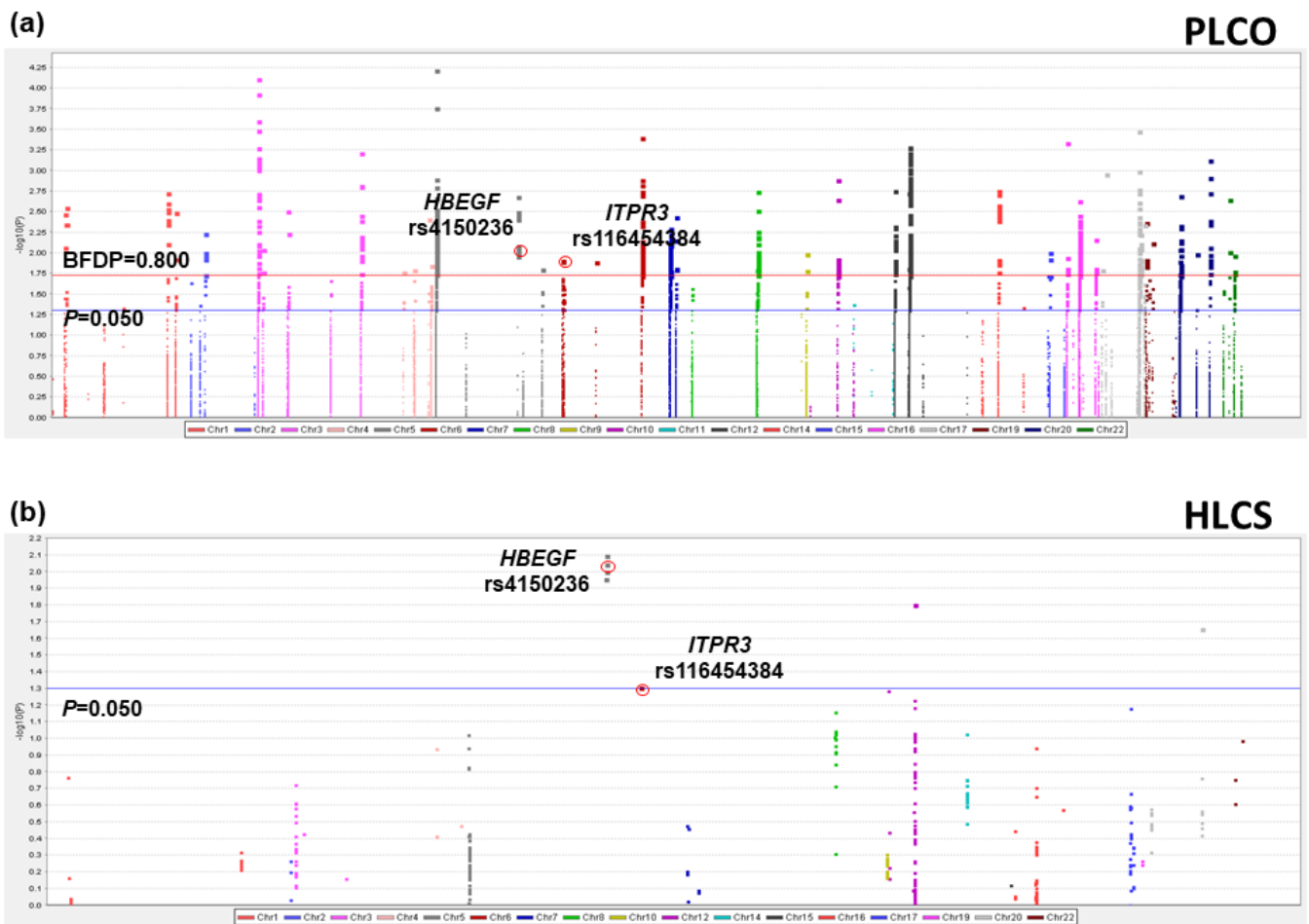


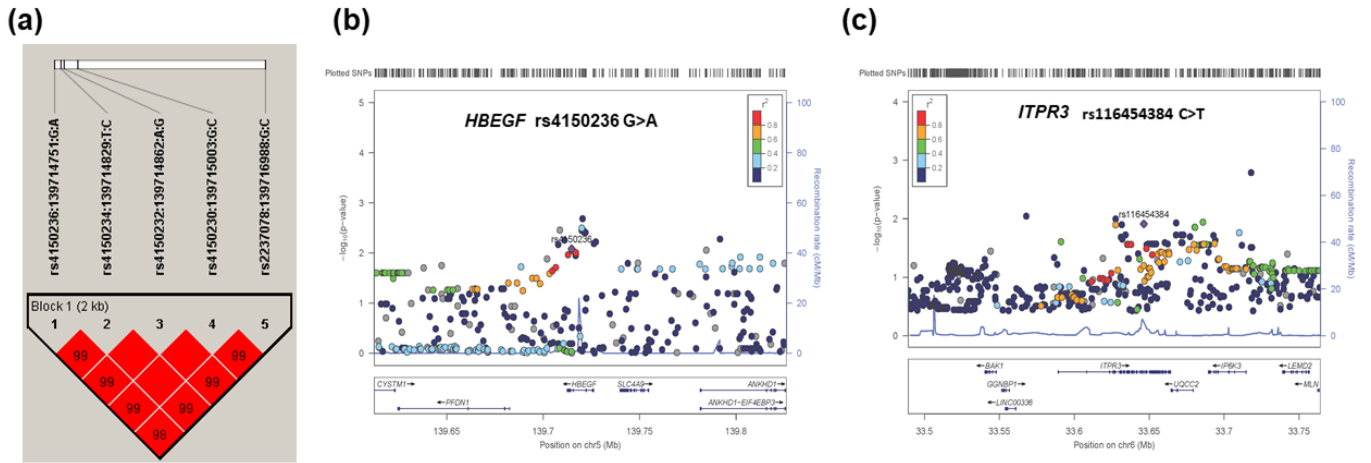
Supplementary Figure 1. The distribution of the imputation info score of the present study.



Supplementary Figure 2. Manhattan plot of the discovery set in genotype data of the PLCO trial and HLCS study.

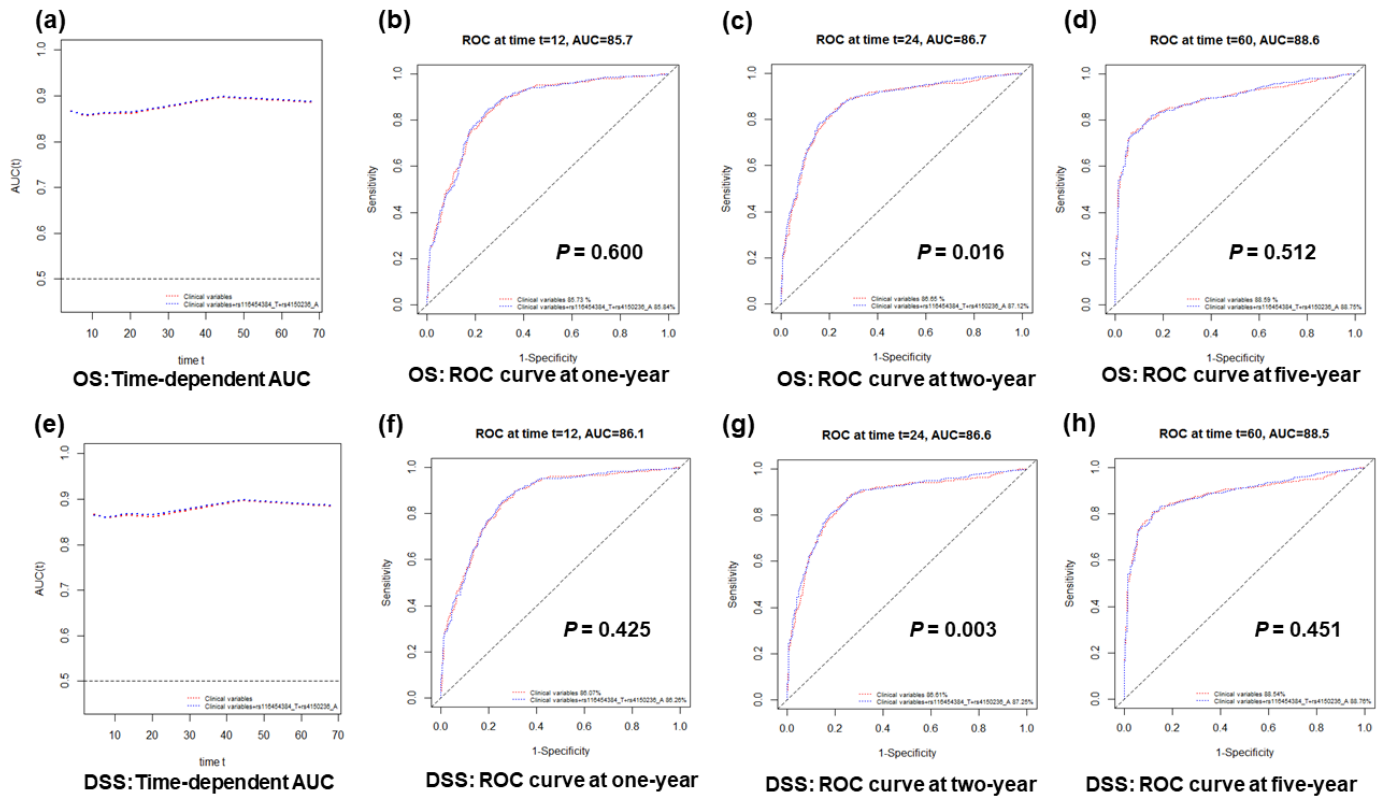
(a) Manhattan plot of the discovery set in genotype data of the PLCO trial and **(b)** HLCS. The statistical values across the autosomes of associations between 22,528 SNPs and overall survival are plotted as $-\log_{10} P$ values. The blue horizontal line indicates $P = 0.05$ and the red line indicates BFDP = 0.80.

Abbreviations: NSCLC, non-small cell lung cancer; SNPs, single nucleotide polymorphism; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening trial; OS, overall survival; BFDP, Bayesian false discovery probability.



Supplementary Figure 3. Regional association plots for the two independent SNPs in the immunoregulatory interaction pathway genes in the 1000 Genome Projects.

(a) Linkage disequilibrium (LD) plot between six SNPs in *HBEGF*. **(b)** and **(c)** Regional association plots for the two independent SNPs in the GnRH pathway genes, **(b)** rs4150236 in *HBEGF* with 150 kb up-down-stream of the gene region; **(c)** rs116454384 in *IP3R3* with 150 kb up-down-stream of the gene region. Data points are colored according to the level of linkage disequilibrium of each pair of SNPs based on the hg19/1000 Genomes European population. The left-hand y-axis shows the association P-value of individual SNPs in the discovery dataset, which is plotted as $-\log_{10}(P)$ against chromosomal base-pair position. The right-hand y-axis shows the recombination rate estimated from HapMap Data Rel 22/phase II European population.

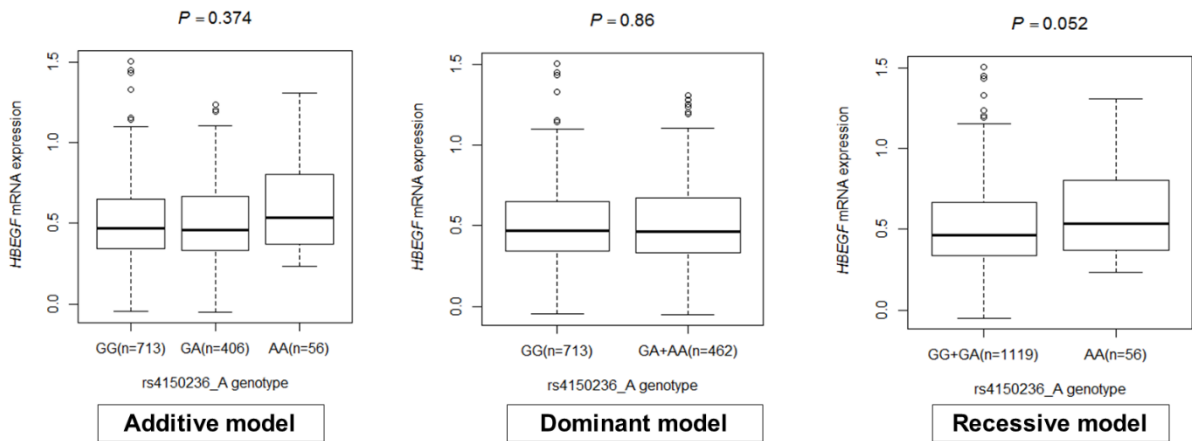


Supplementary Figure 4. NSCLC survival prediction of the two SNPs by ROC curve in PLCO dataset.

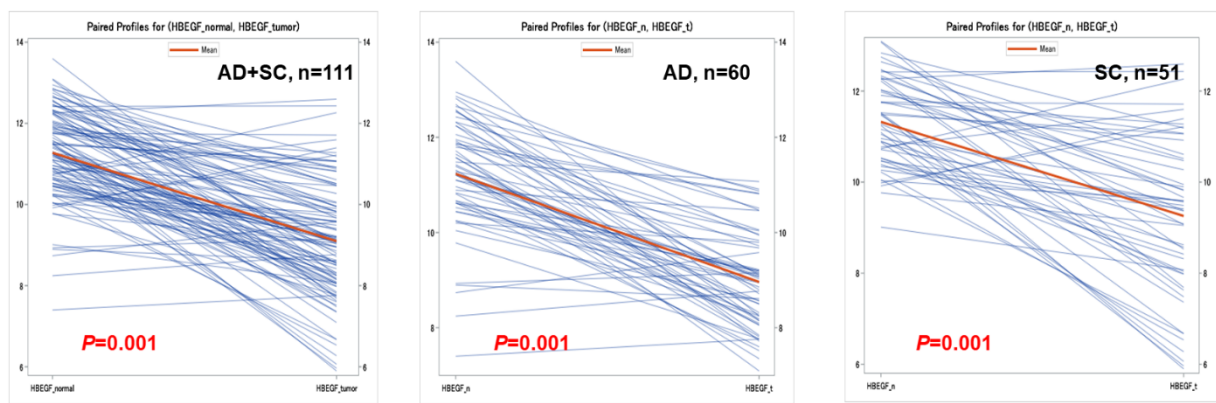
(a, e) Time-dependent AUC estimation for OS and DSS: based on age, sex, smoking status, histology, tumor stage, chemotherapy, surgery, principal component and SNPs. **(b, c, d)** 1,2 and 5-year NSCLC OS prediction by ROC curve, **(f, g, h)** 1,2 and 5-year NSCLC DSS prediction by ROC curve.

Abbreviations: ROC, receiver operating characteristic curve; AUC, area under curve; OS, overall survival; DSS, disease-specific survival; PLCO, The Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial.

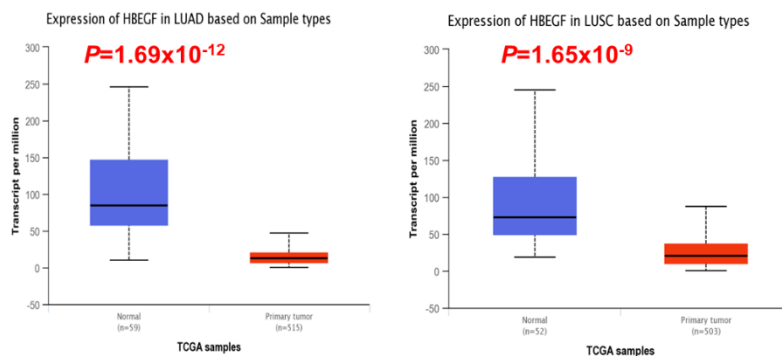
(a) *HBEGF* rs4150236 G>A in the 1000 Genomes Project



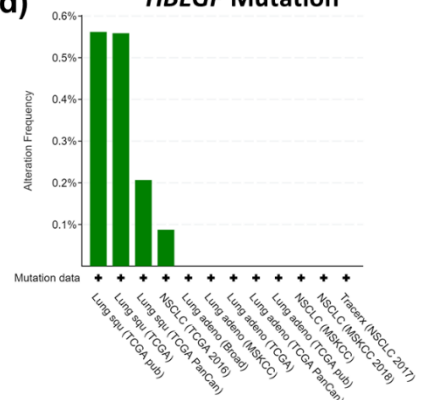
(b) Paired Test – *HBEGF* mRNA expression analysis in the TCGA dataset



(c) *HBEGF* mRNA expression in the TCGA dataset



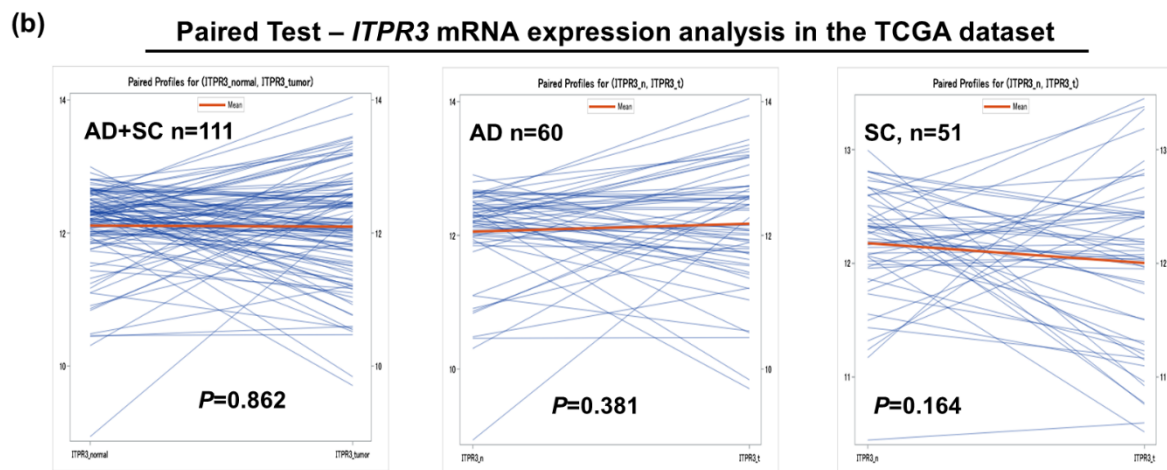
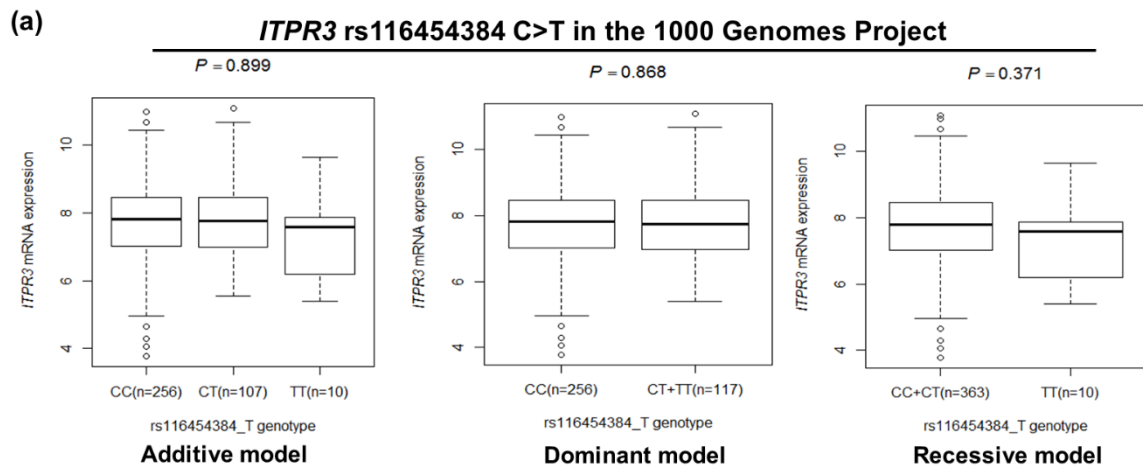
(d) *HBEGF* Mutation



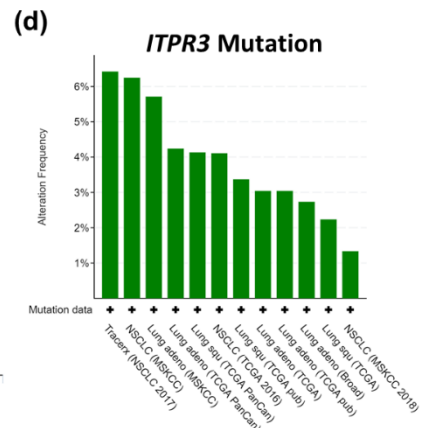
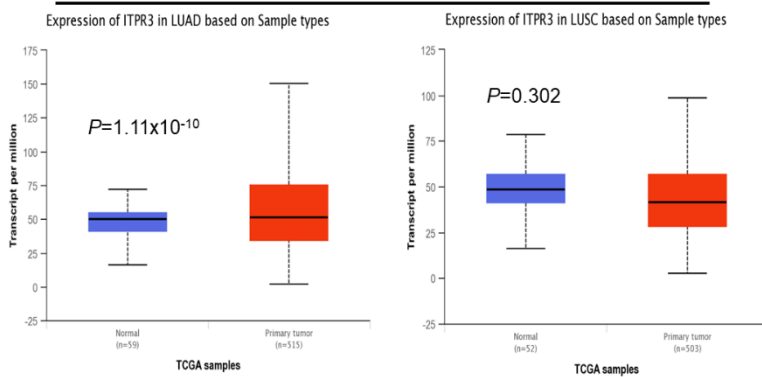
Supplementary Figure 5. Correlation of genotypes of *HBEGF* rs4150236 with their corresponding mRNA expression levels of the *ITPR3* in the 1000 Genomes Project and in the TCGA database.

(a) rs4150236 variant AA genotype was borderline significantly correlated with an increased expression level of HB-EGF mRNA, while this was not the case in the additive model and dominant model, compared with the GG/GA genotypes in the recessive model in the 1000 Genomes Project. **(b)** higher *HBEGF* mRNA expression were found in the adjacent normal tissues than that in the tumor tissues of paired LUAD and paired LUSC. **(c)** higher *HBEGF* mRNA expression were also found in the non-paired normal tissues than tumor tissues in LUAD and (h) LUSC (<http://ualcan.path.uab.edu>). **(d)** Mutation frequency of *HBEGF* in NSCLC (<http://www.cbioportal.org>).

Abbreviation: eQTL, expression quantitative trait loci analysis; TCGA, The Cancer Genome Atlas; NSCLC, non-small cell lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.



(c) *ITPR3* mRNA expression in the TCGA dataset

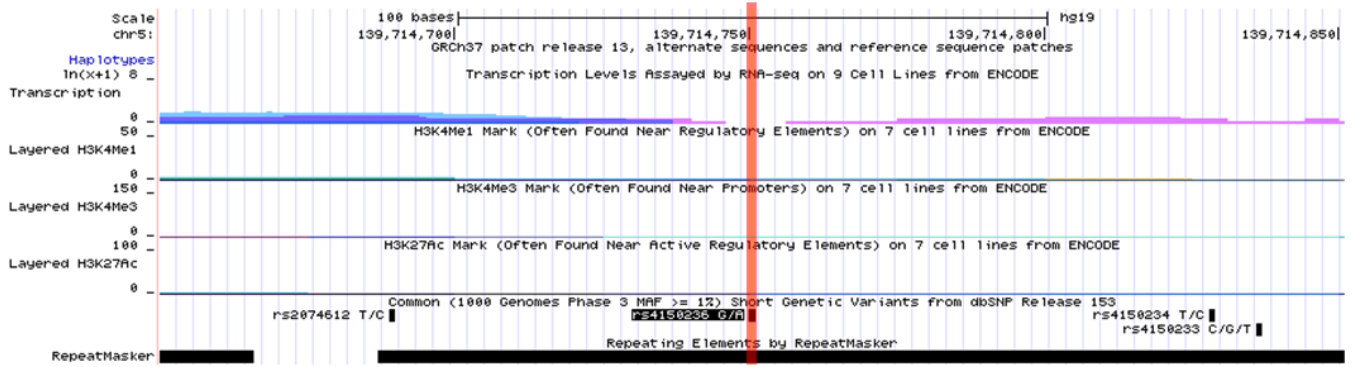


Supplementary Figure 6. Correlation of genotypes of *ITPR3* rs116454384 with their corresponding mRNA expression levels of the *ITPR3* in the 1000 Genomes Project and in the TCGA database.

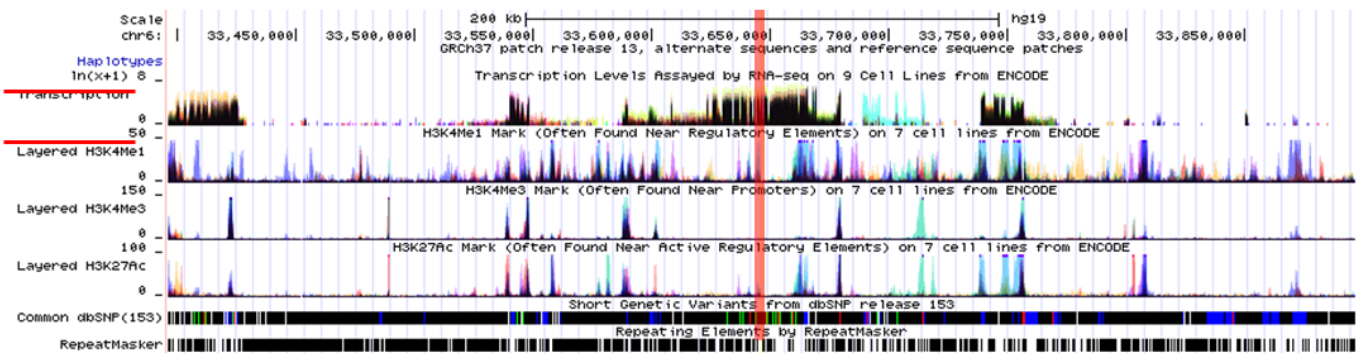
(a) The eQTL for *ITPR3* rs116454384 in 373 Europeans from the 1000 Genomes Project; *PVR* mRNA expression levels has no significant were found in the adjacent normal tissues of 111 NSCLC and in the 51 paired LUSC tissues but no in LUAD; **(b)** the mRNA expression levels of *ITPR3* were not significantly different in in the tumor tissues of 111 paired NSCLC, 60 paired LUSC tissues and 51 paired LUSC tissues. **(c)** Higher *ITPR3* mRNA expression levels were found in the non-paired normal tissues than tumor tissues in LUAD but not in the LUSC (<http://ualcan.path.uab.edu>). **(d)** Mutation frequency of *ITPR3* in NSCLC (<http://www.cbioportal.org>).

Abbreviation: eQTL, expression quantitative trait loci analysis; TCGA, The Cancer Genome Atlas; NSCLC, non-small cell lung cancer; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.

HBEGF rs4150236



ITPR3 rs116454384



Supplementary Figure 7. Expanded view of the ENCODE data for the rs4150236 and rs116454384.

The H3K27Ac, H3K4Me1, and H3K4Me3 tracks showed the genome-wide levels of enrichment of acetylation of lysine 27, the mono-methylation of lysine 4, and tri-methylation of lysine 4 of the H3 histone protein, as determined by the ChIP-seq assays. These levels are thought to be associated with the promoter and enhancer regions. DNase clusters track showed DNase hypersensitivity areas. Tnx factor track showed regions of transcription factor binding of DNA, as assayed by ChIP-seq experiments. Transcription showed target genes of transcription factors from transcription factor binding site profiles. Abbreviation: ChIP, Chromatin Immunoprecipitation Sequencing.