Lung Cancer Pharmacogenomics—An Integrative Analysis of Patient Survival, Genetic Variants, and Gene Expression

Chao Agnes Hsiung
Institute of Population Health Sciences
National Health Research Institutes

Health Care 趨勢

- Prevention, Personalization, Prediction, Participation

  三段五級預防 Precision medicine 風險預測 (Risk prediction)

  Pharmacogenomics 預後預測 (Prognostic prediction)

- Big data : volume, velocity, variety, veracity, value

  Integration

- Pharmacogenomics→customized therapy (precision medicine)

  Identify subgroup

  Drug target

Lung Cancer Treatment

- Lung cancer is the leading cause of cancer death in the world
- Early stage lung cancers, treated by resection, have much better survival than late stage
- 85% lung cancer cases are late stage; survival is poor.
- Non-small cell lung cancer (NSCLC) includes lung adenocarcinoma and lung squamous cell lung cancer and accounts for more than 80% of lung cancer.
- Standard care of late stage NSCLC was chemotherapy before 2004, a little better than palliative care.
- The introduction of TKI in 2003 is remarkable for late stage NSCLC.

Gefitinib

- First tyrosine kinase inhibitor (TKI) of *EGFR*.
- Its development was motivated by the observation that EGFR expression is higher in lung cancer tissue than in adjacent normal tissue.
- It was approved by US FDA in May 2003.
- Drug response appeared more frequently in Japan, female, non-smokers, lung ADC (RCTs).
- Expression of EGFR is not a good predictor of drug response.

April 29 2004 (*NEJM*, *Science*)

- *EGFR* mutations in lung cancer: correlation with clinical response to gefitinib therapy
- Mutations in *EGFR* were found in 8/9 vs 0/7, in gefitinib responsive vs nonresponsive lung cancer. (*NEJM*)
- Somatic mutations found in 15/58 from Japan and in 1/61 from the US. (*Science*)
- Treatment with gefitinib causes tumor regression more frequently in Japan than in the US.

Sept. 2004 *PNAS* (Varmus)

- *EGFR* gene mutations are common in lung cancers from never smokers and are associated with sensitivity of tumors to gefitinib and erlotinib.
- Mutations found in 7/10 vs 0/8, according to gefitinib sensitivity.
- Mutations found in 5/7 vs 0/10, according to erlotinib sensitivity.
- Mutations found in 7/15 vs 4/81, according to smoking status. (non-smokers vs smokers)

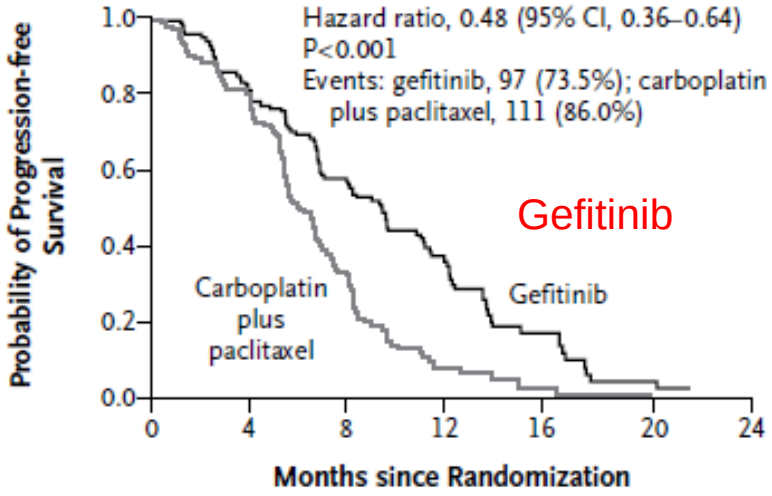
Precision medicine

- Tumors with *EGFR* mutations were sensitive to a tyrosine kinase inhibitor (TKI) (2004).
- From clinical variables to genetic markers (*Nature Genetics* 2005) for TKI response.
- Longer progression-free survival in patients with mutated *EGFR* treated with first-line EGFR-TKIs, compared with those treated with platinum-based chemotherapy. (2010-13, 4 prospective RCTs)
- Testing for *EGFR* mutation is standard in patients with advanced NSCLC. Mutation positive may benefit from TKI.
- First-line use IIIB/IV lung ADC, reimbursement by NHI program, 2011, June

An Important Study in 2009: Gefitinib vs Carboplatin plus Paclitaxel in Lung ADC

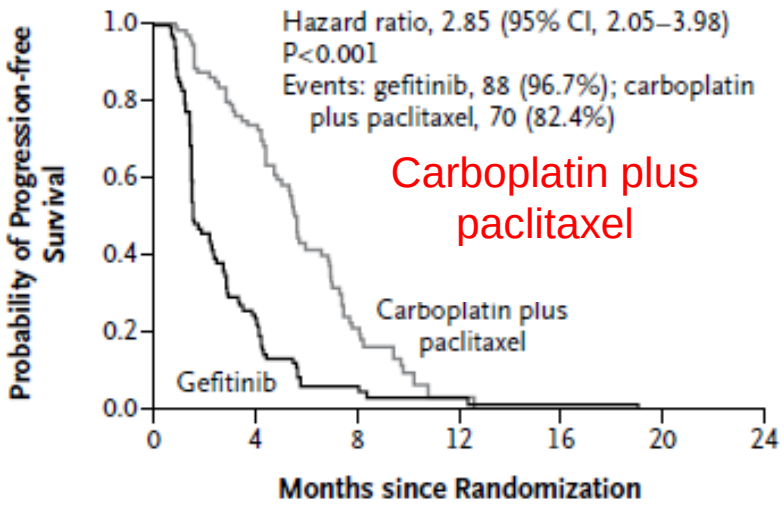
- Progression-free Survival Comparison for Different EGFR Mutation Status

EGFR-Mutation–Positive



No. at Risk							
Gefitinib	132	108	71	31	11	3	0
Carboplatin plus paclitaxel	129	103	37	7	2	1	0

EGFR-Mutation–Negative



No. at Risk							
Gefitinib	91	21	4	2	1	0	0
Carboplatin plus paclitaxel	85	58	14	1	0	0	0

Study subjects

Advanced lung adenocarcinoma

Nonsmokers or former light smokers

Untreated

Mok et al., *N Engl J Med* 2009

Two observations

- Mutation positive patients experienced longer progression-free survival (PFS).
- Many patients with mutated *EGFR* still have short PFS.

Cancer Pharmacogenomics

- For some targeted therapies, specific somatic mutations are predictive of target efficacy.
- In cancer, tumors may have specific disease-defining somatic mutations, germline genetic variation will also affect drug response.

Wheeler et al., *Nature Reviews/Genetics* 2013

Drug	Drug target	Cancer type (or types)	Somatic markers
Cetuximab	EGFR	Colorectal, head and neck	*EGFR* and *KRAS*
Erlotinib	EGFR	Lung, pancreatic	*EGFR*
Exemestane	Aromatase	Breast	*ESR1*, *ESR2* and *PGR*
Gefitinib	EGFR	Lung	*EGFR*
Tamoxifen	Oestrogen receptor	Breast	*ESR1*, *ESR2* and *PGR*

Drug	Mechanism of action	Cancer type (or types)	Genes	Variants	Phenotype	Type of study	Evidence
Tamoxifen	Inhibits the oestrogen receptor	Hormone-receptor-positive breast	*CYP2D6*	rs16947 rs1065852 rs28371706 rs28371725 rs35742686 rs3892097 rs5030655 rs5030656 rs59421388 rs61736512	Tamoxifen metabolism, progression-free and overall survival	Candidate gene	Conflicting results may be due to study design and quality control; studies are ongoing

Lung Cancer Pharmacogenomics Study

- Looking for subgroup for whom TKIs are more beneficial.
- Looking for germline genetic variants associated with PFS. (Pharmacogenomics) Genome-wide association study. Non-invasive.

EGFR mutation
- ＋ 、***EGFR* -TKIs**
- — 、platinum-based doublets

***EGFR*＋** → Risk allele at marker SNPs
- germline genetic variants
- non-invasive markers

→ No, *EGFR* -TKIs

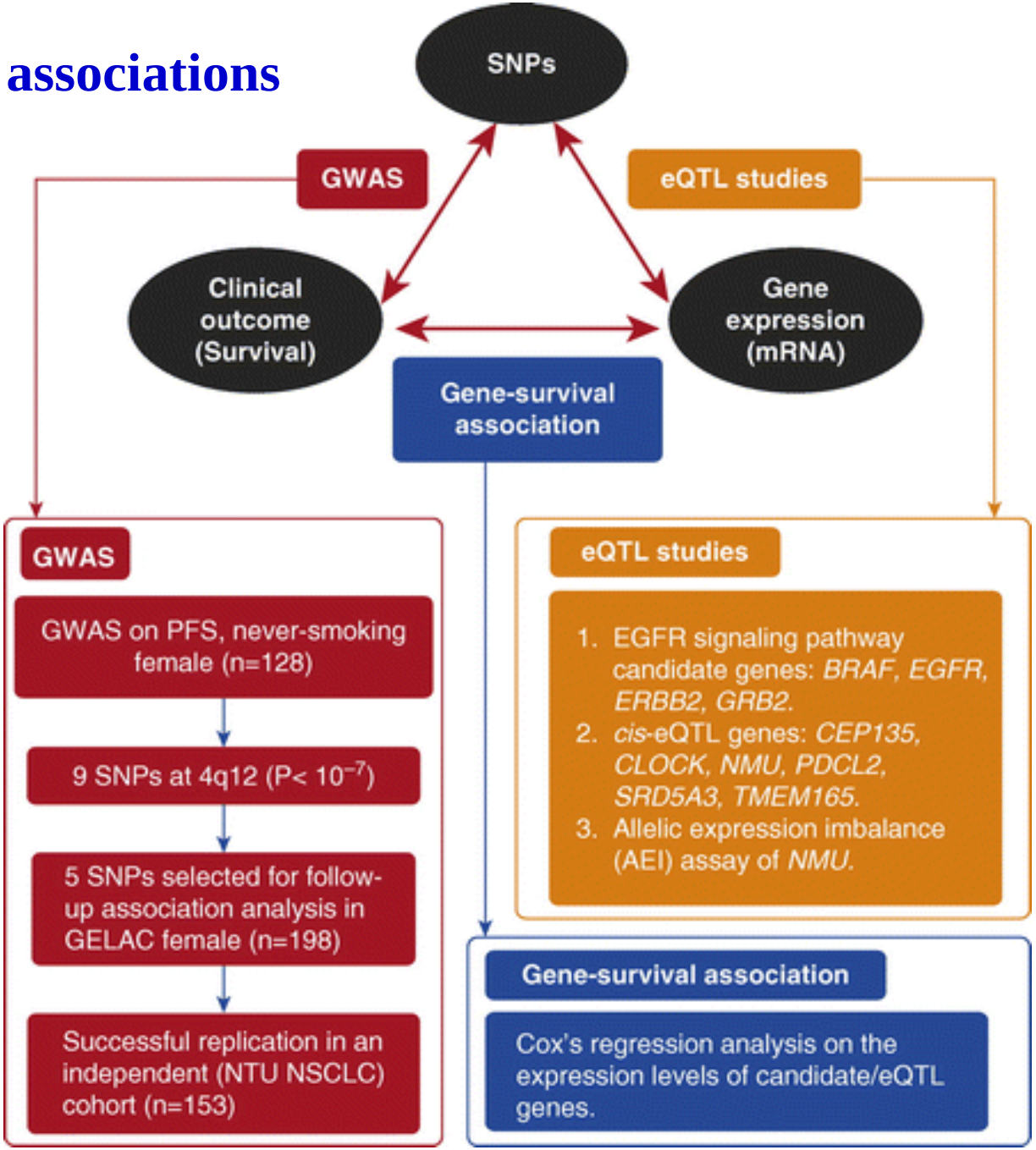
→ Yes, first-line platinum-based doublets immunotherapy

Methods

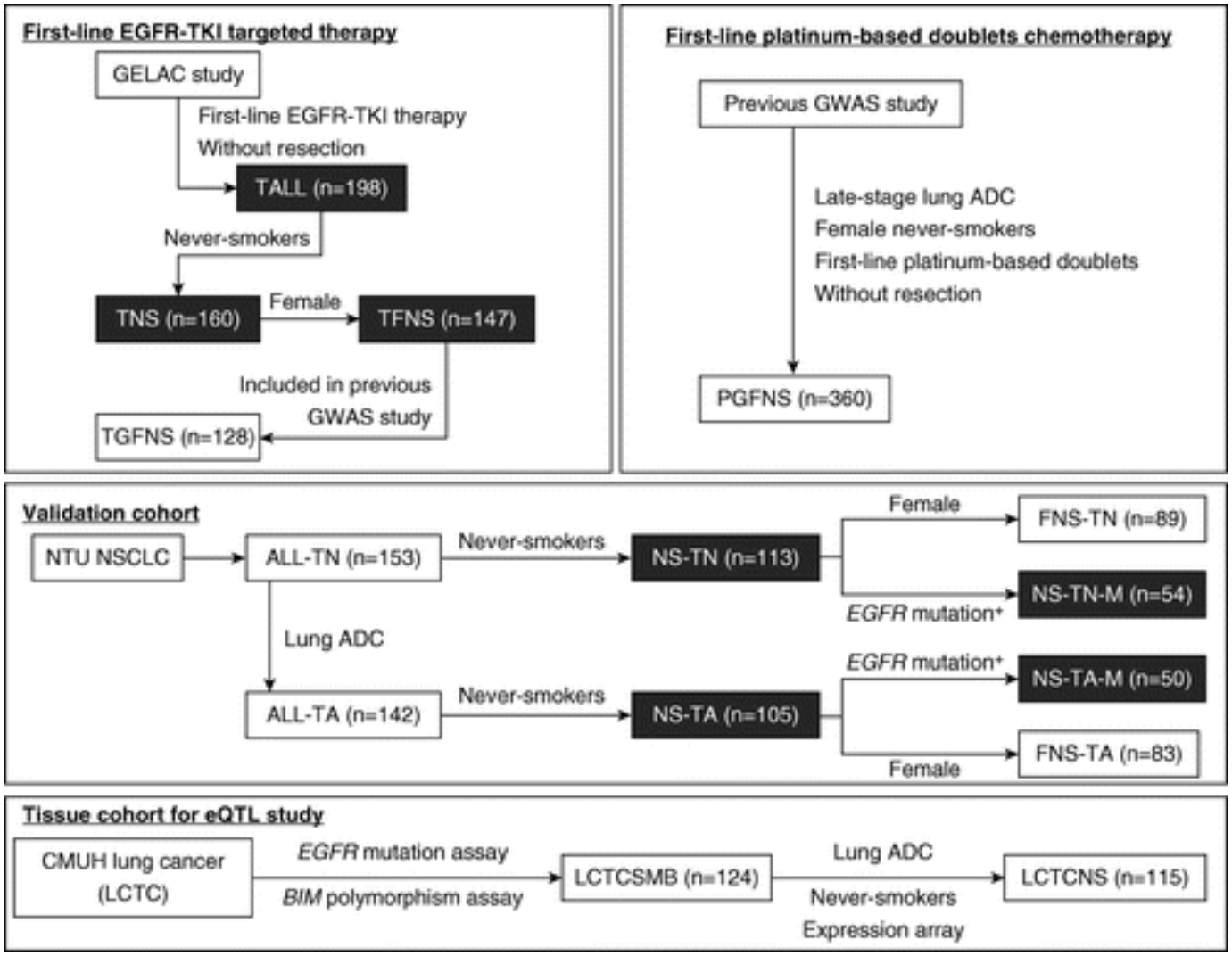
- A genome-wide association study on PFS was performed in never-smoking women diagnosed with lung adenocarcinoma and who were treated with first-line EGFR-TKIs.
- Significant single-nucleotide polymorphisms (SNPs) were selected for follow-up association analysis and for replication assay in another independent cohort.
- Progression free survival (Time to progression) : defined as the earliest time that the tumor metastasis appeared, tumor size enlarged, TKI-treatment terminated or died of lung cancer.

Integration of different associations

Enhancing the biological plausibility of the association between SNPs and PFS by eQTL analyses and associations between gene expression and clinical phenotypes

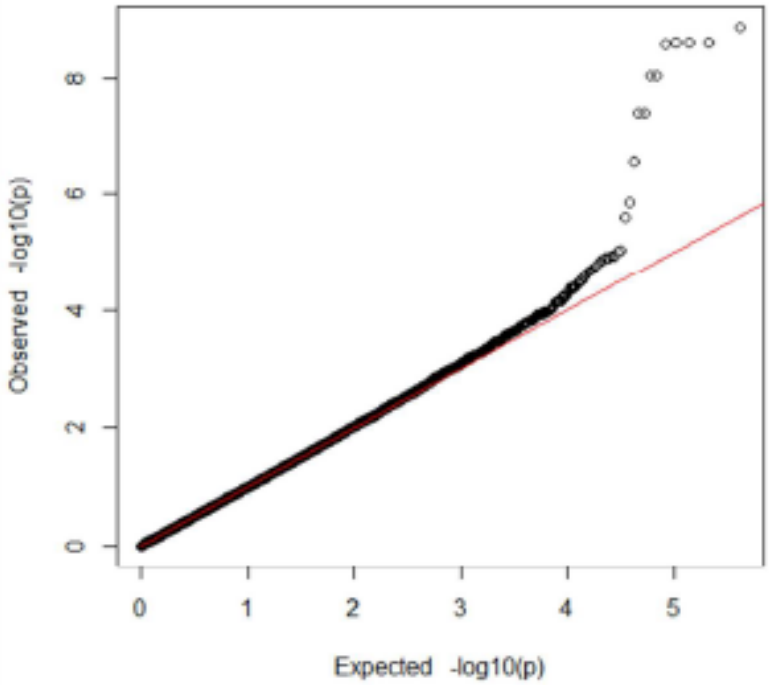


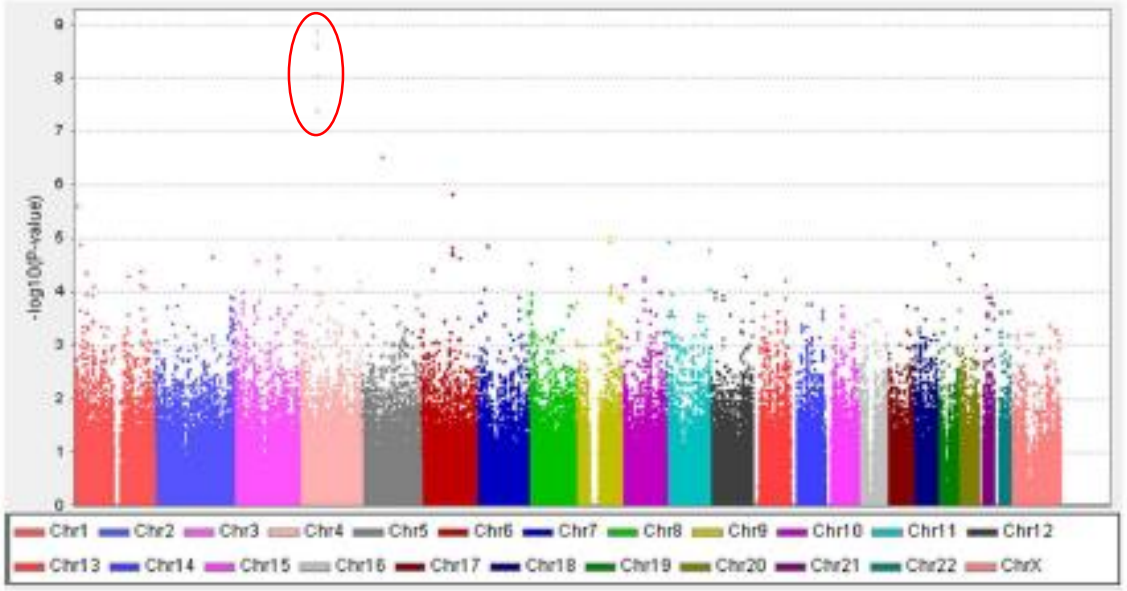
Cohorts and subcohorts used in this study. Cohorts from which association of genome-wide level of significance ($P<10^{-8}$) was generated, or from which association was successfully replicated ($P<0.05$), are highlighted.



Genome-wide association study of progression-free survival for TKI

Quantile-quantile plot for –log P values (A); Manhattan plot of P values in –log scale (B); based on the GWAS of 128 never-smoking female patients with lung ADC treated with first-line TKIs, without resection. Genomic inflation factor lambda was 0.980 (there is little population substructure).





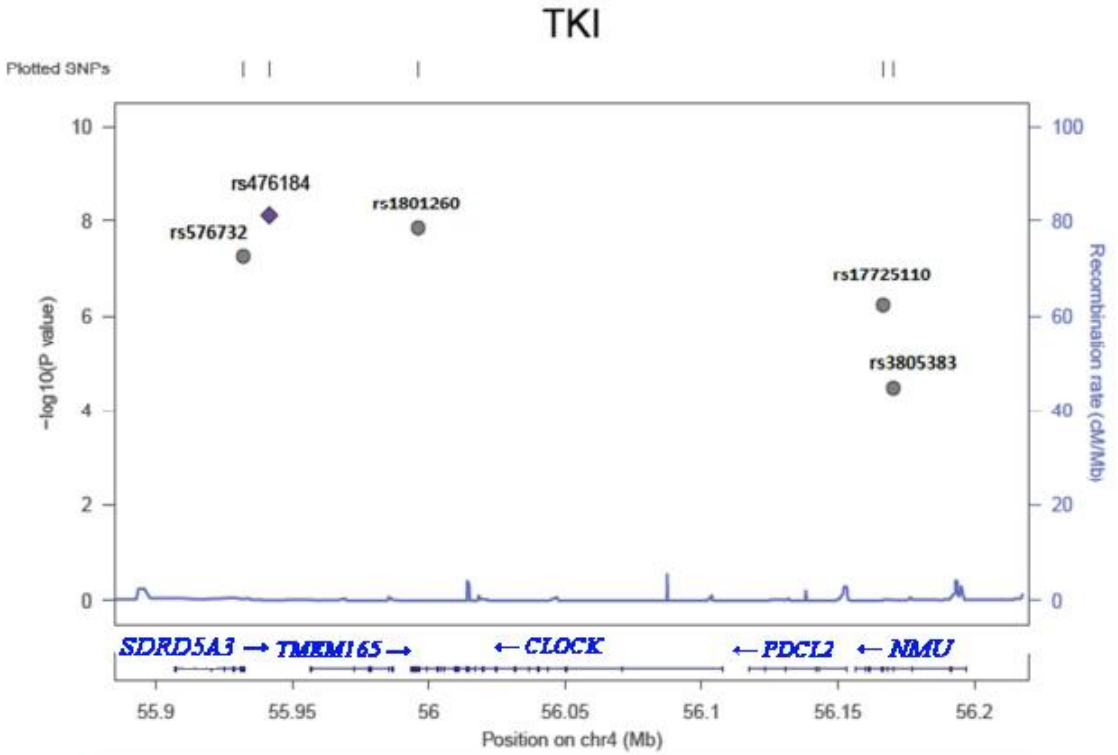
Association between SNPs at 4q12 and PFS Based among TFNS, TNS, and All Patients with Lung ADC Treated with First-Line TKI in GELAC Cohort

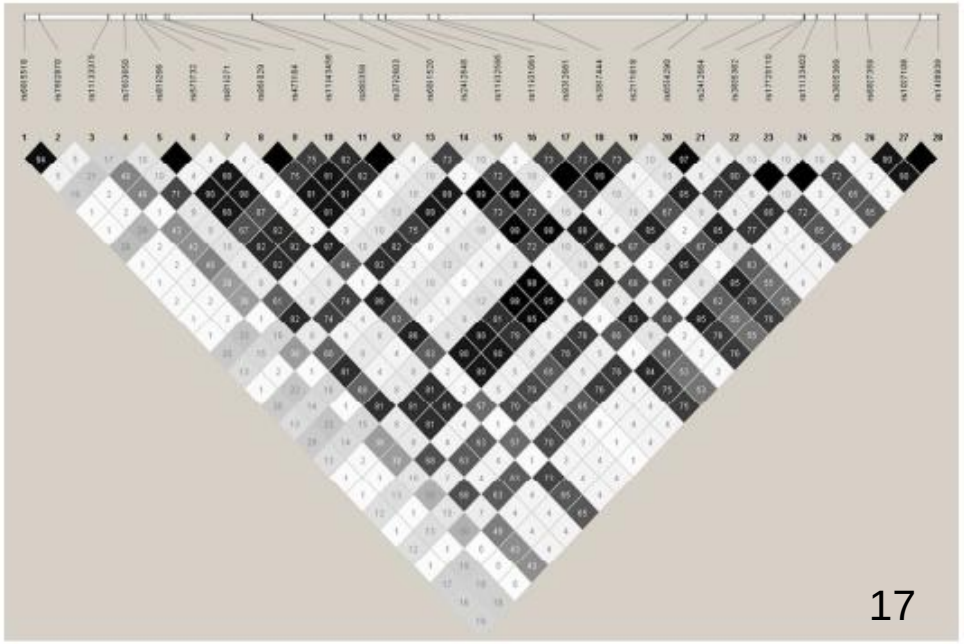
SNP	Chr	Position*	Allele†	MAF‡	*P* Value§	Hazard Ratio§‖ (95% CI)
rs576732	4	55931894	G/A	0.08	4.6×10^{-7}	3.72 (2.23–6.20)
					5.8×10^{-8}	3.97 (2.41–6.52)
					6.3×10^{-8}	3.54 (2.24–5.59)
rs476184	4	55941621	T/C	0.0718	3.2×10^{-8}	4.58 (2.67–7.85)
					3.1×10^{-8}	4.90 (2.90–8.28)
					8.5×10^{-9}	4.05 (2.51–6.51)
rs1801260	4	55996126	T/C	0.0704	2.8×10^{-8}	4.62 (2.69–7.94)
					5.7×10^{-9}	4.91 (2.88–8.39)
					1.6×10^{-8}	4.04 (2.49–6.55)
rs17725110	4	56166779	A/G	0.0631	4.1×10^{-6}	3.64 (2.10–6.31)
					5.4×10^{-7}	3.92 (2.30–6.69)
					7.7×10^{-7}	3.49 (2.13–5.73)
rs3805383	4	56170095	C/T	0.0541	3.5×10^{-4}	2.88 (1.61–5.14)
					8.0×10^{-5}	3.09 (1.76–5.42)
					4.7×10^{-5}	2.90 (1.74–4.84)

Largest hazard ratios appeared in the group of nonsmokers

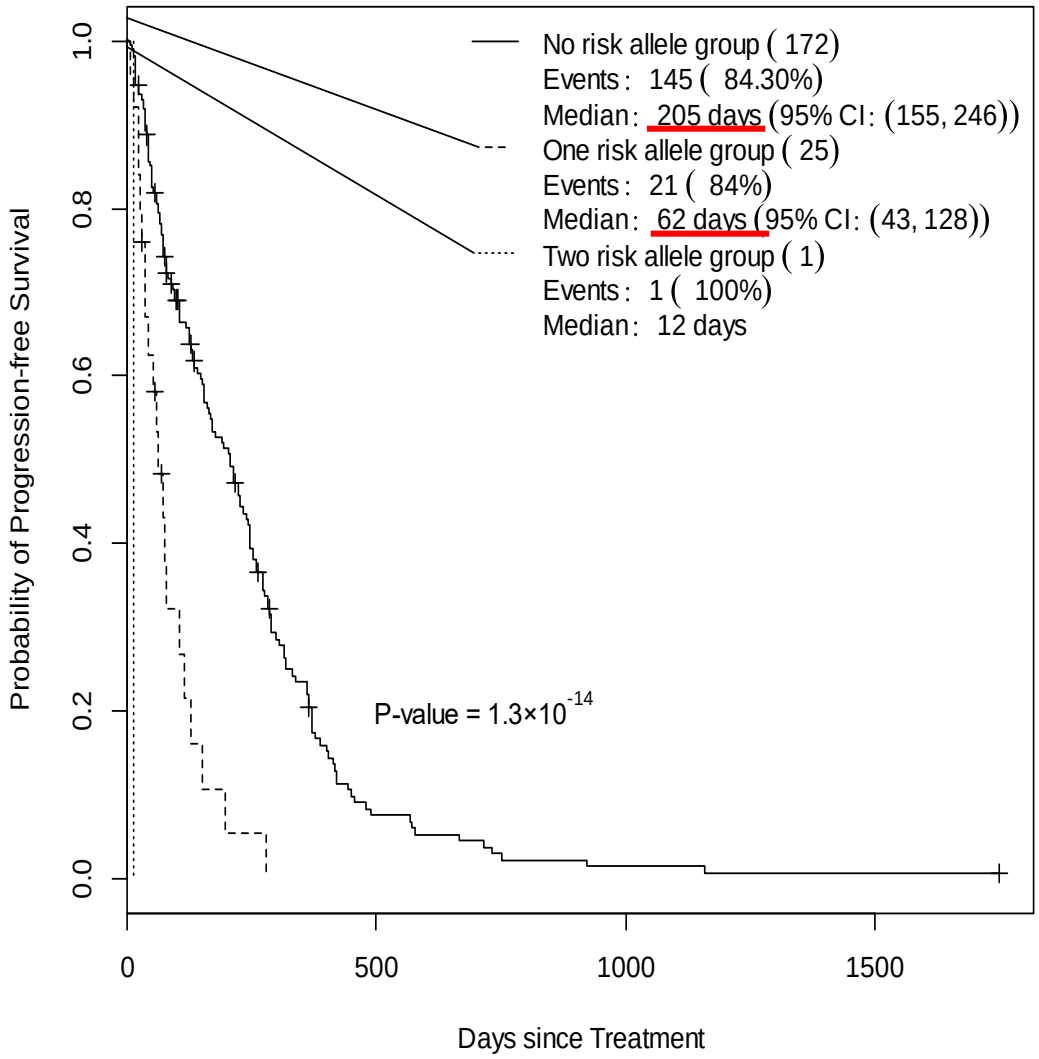
Association results, recombination rate and LD plots for the chromosomal region 4q12 associated with progression-free survival among patients with lung ADC treated with first-line TKIs

Am J Respir Crit Care Med. 201
(IF=13.118, Ranking : 3.4%)





PFS is significantly associated with the genotype at risk variant in patients treated with first-line EGFR-TKIs



Precision Medicine : Lung Cancer Pharmacogenomics for first line TKI

Figure 1a, SNP1_202

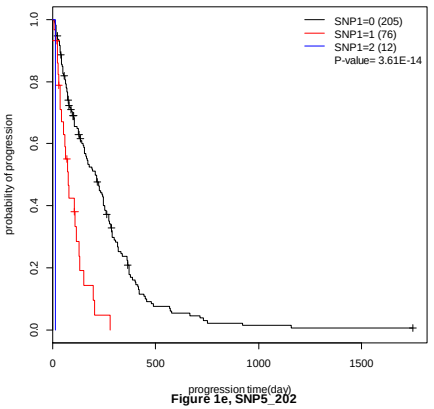


Figure 1b, SNP2_202

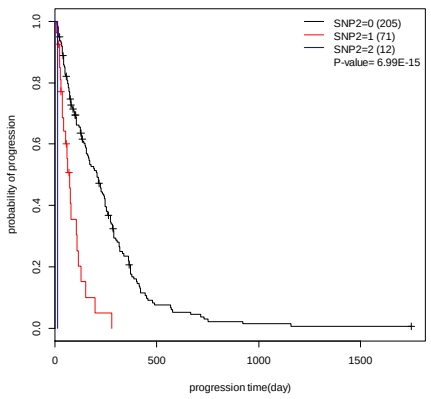


Figure 1c, SNP3_202

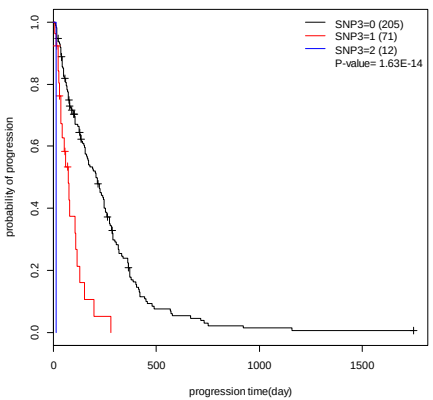


Figure 1d, SNP4_202

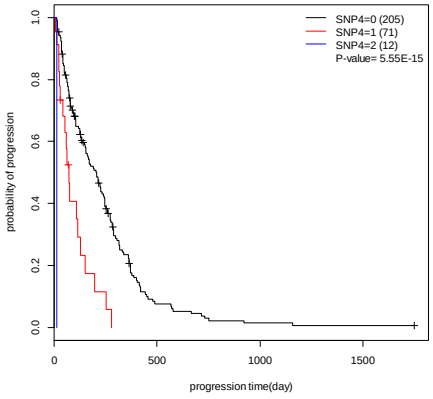


Figure 1e, SNP5_202

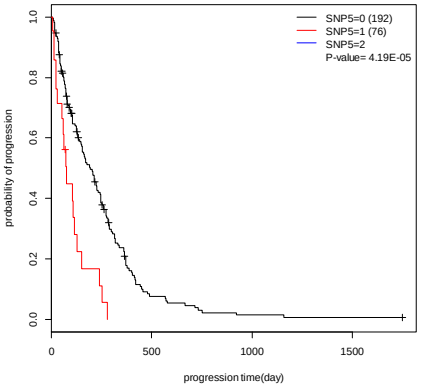
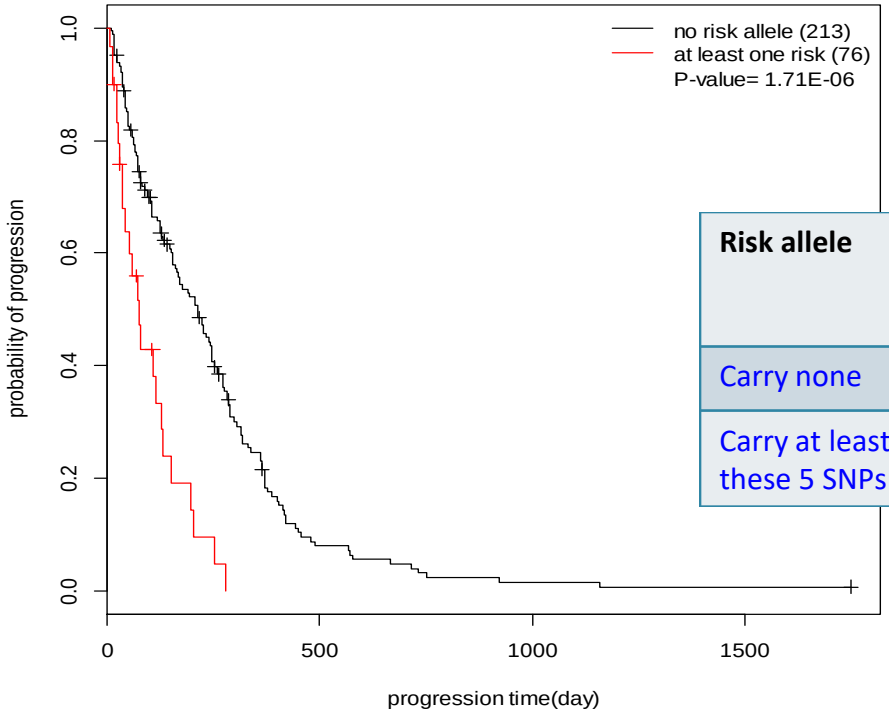


Figure 1f, SNP1+ SNP2+ SNP3+ SNP4_202



No risk allele(212)

At least one risk(76)

P-value=1.71x 10^{-6}

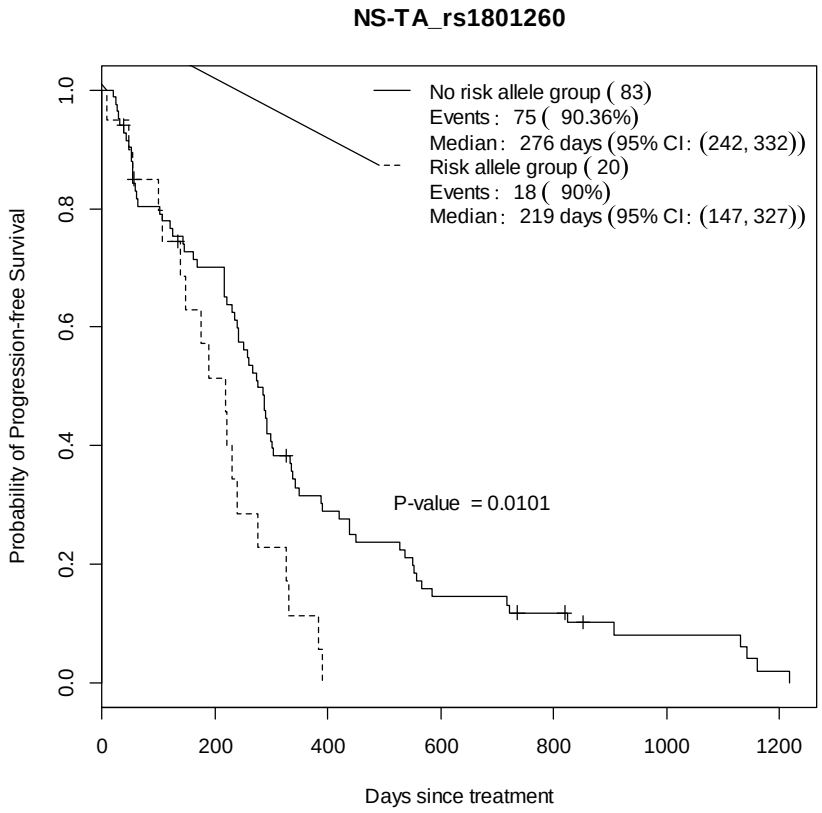
Risk allele	Median PFS (days)	95%CI
Carry none	212	(159, 247)
Carry at least one at these 5 SNPs	76	(59, 151)

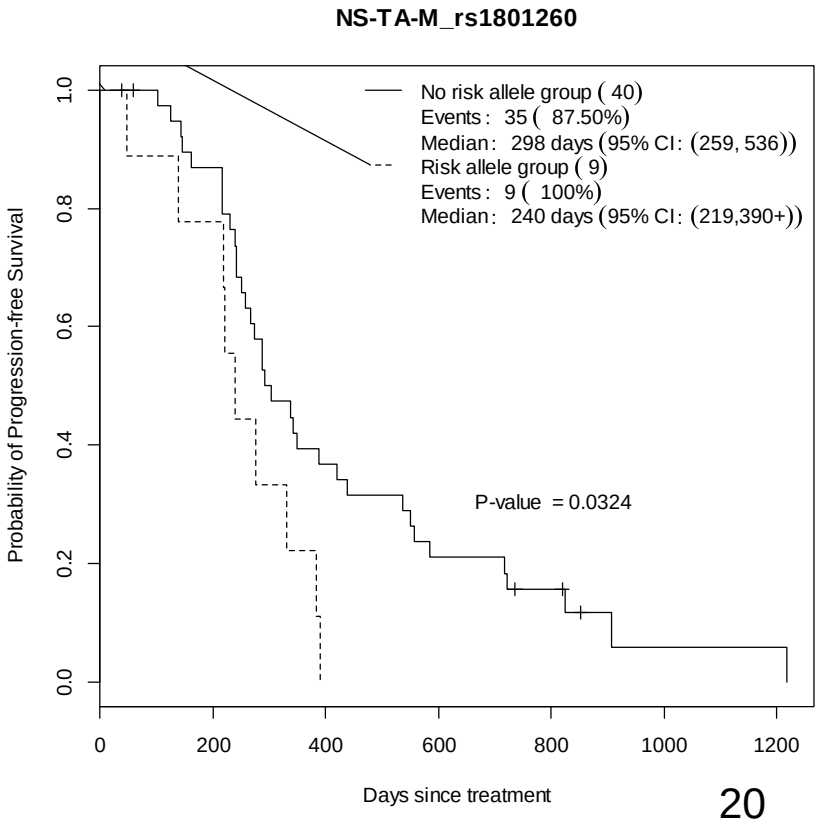
Successful Replication of the Association based on Validation Cohort NTU ADC

Log-rank test for the NTU NSCLC cohort. Kaplan-Meier curves for progression-free survival

(c), for the entire never-smoking patients with lung ADC;

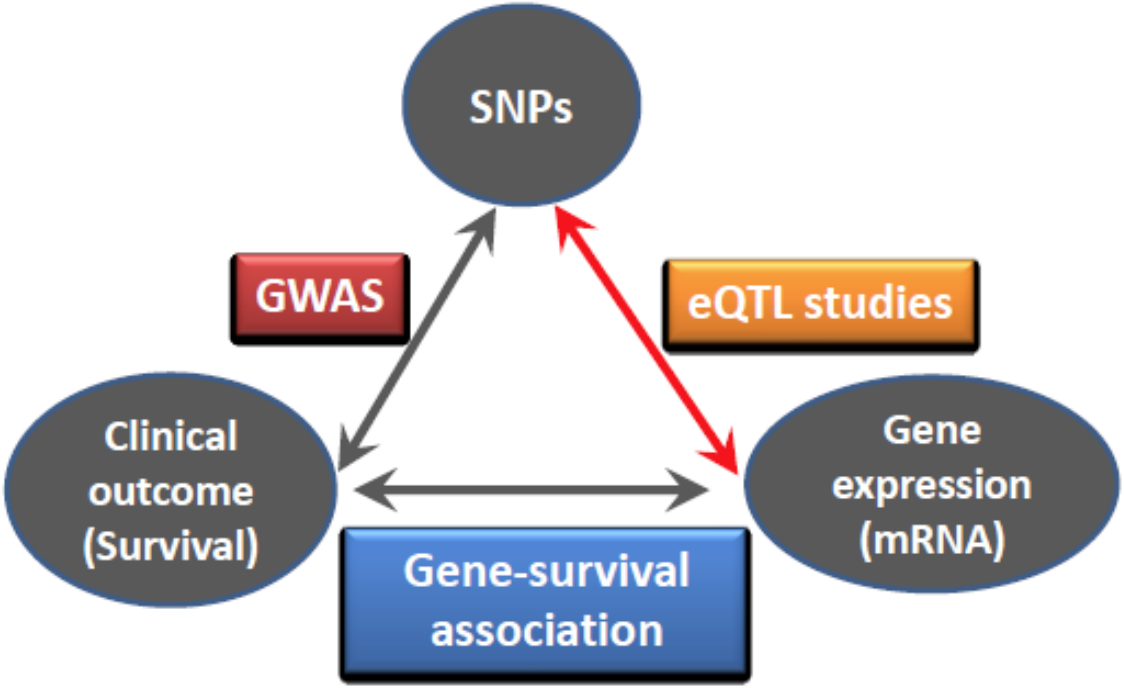
(d), for all the never-smoking lung ADC testing positive for *EGFR* mutation.





Functional Study －Enhancing the biological plausibility of the association by eQTL (expression quantitative trait loci)

- 115 normal/tumor **tissue samples**
- genes located **within 500kb of these SNPs**
- Correlation between SNPs and gene expression levels
- Query public microarray gene expression data



eQTL Study + Other Functional Studies → Identify Drug Target

eQTL studies

1. EGFR signaling pathway candidate genes: *BRAF, EGFR, ERBB2, GRB2.*
2. *cis*-eQTL genes: *CEP135, CLOCK, NMU, PDCL2, SRD5A3, TMEM165.*
3. Allelic expression imbalance (AEI) assay of *NMU.*

Association between the SNPs & expression probes within 500kb

eQTL analysis for rs3805383

Sample	Tissue	*BRAF*		*CLOCK*		*EGFR*				*ERBB2*					
		ILMN_1652472		ILMN_1682399		ILMN_1755535		ILMN_1696521		ILMN_1717902		ILMN_1728761		ILMN_2352131	
		β	P-value	β	P-value	β	P-value	β	P-value	β	P-value	B	P-value	β	P-value
ALL (n=115)	T	0.118	0.038	0.000	0.998	0.279	0.040	0.081	0.461	-0.087	0.087	0.158	0.020	0.066	0.123
	N	0.031	0.438	-0.078	0.237	0.049	0.609	0.038	0.542	-0.072	0.152	-0.099	0.135	0.055	0.018
EGFR mut+ (n=72)	T	0.198	0.008	-0.024	0.815	0.398	0.016	0.123	0.304	-0.112	0.098	0.103	0.236	0.091	0.070
	N	-0.007	0.883	-0.188	0.015	0.037	0.777	0.052	0.507	-0.038	0.559	-0.038	0.634	0.045	0.151
L858R (n=46)	T	0.339	0.000	-0.004	0.980	0.755	0.001	0.333	0.029	-0.177	0.046	0.141	0.206	0.116	0.087
	N	-0.030	0.643	-0.216	0.020	0.022	0.886	0.061	0.560	-0.016	0.862	0.049	0.599	0.047	0.240
GTEx		ENSG00000157764.7		ENSG00000134852.10		ENSG00000146648.11				ENSG00000141736.9					
		-0.03	0.38	-0.12	0.036	0.07	0.1			-0.042	0.3				

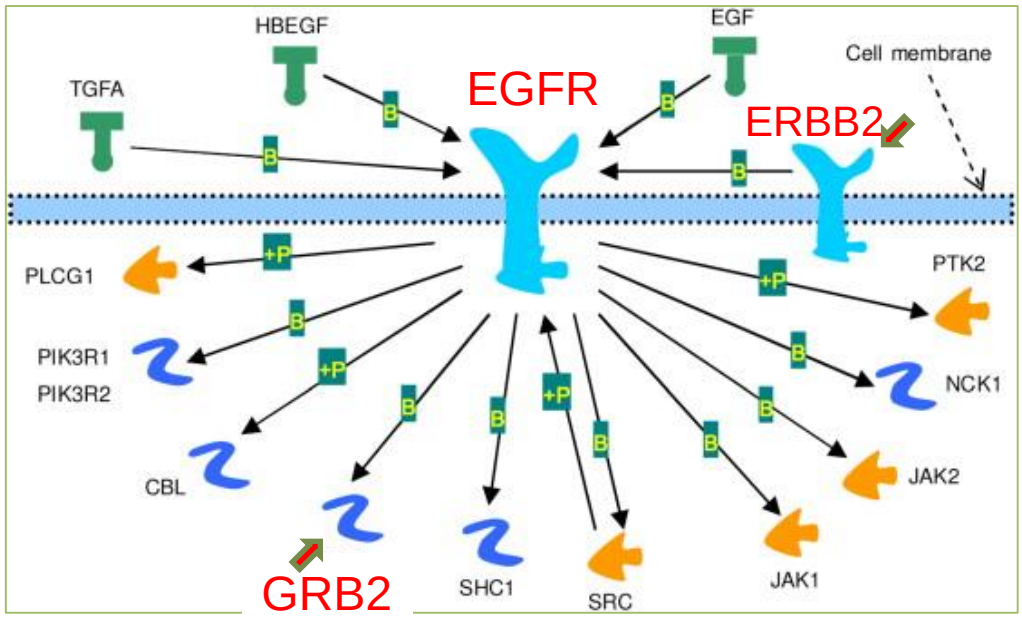
These SNPs are associated with the expression of **EGFR signaling pathway genes**, including *EGFR*, which encodes the target of TKI.

eQTL analysis for rs3805383(cont.)

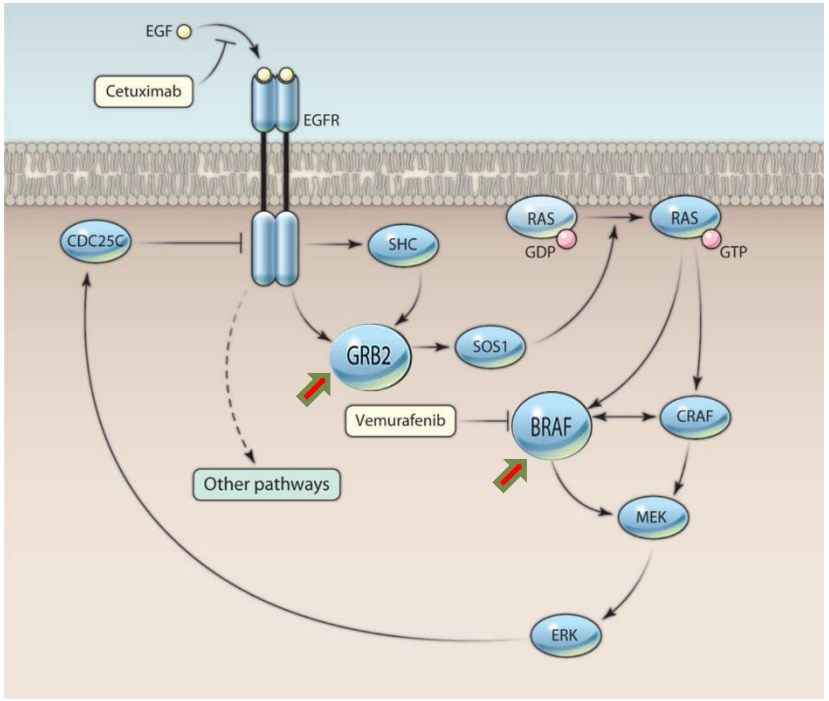
Sample	Tissue	*EXOC1*		*GRB2*		*NMU*		*PDCL2*		*SRD5A3*	
		ILMN_1745583		ILMN_1748797		ILMN_2162253		ILMN_1762409		ILMN_1678435	
		β	P-value	β	P-value	β	P-value	β	P-value	β	P-value
ALL (n=115)	T	0.108	0.100	0.060	0.094	-0.071	0.774	0.149	0.102	0.151	0.117
	N	0.142	0.013	-0.016	0.637	0.283	0.051	-0.090	0.106	0.046	0.427
EGFR mut+ (n=72)	T	0.094	0.287	0.065	0.166	0.300	0.349	0.269	0.027	0.306	0.015
	N	0.112	0.108	-0.010	0.821	0.297	0.102	-0.105	0.155	0.024	0.744
L858R (n=46)	T	0.199	0.077	0.036	0.517	0.725	0.058	0.208	0.143	0.453	0.002
	N	0.161	0.061	-0.046	0.437	0.515	0.044	-0.121	0.215	0.097	0.335
GTEx		ENSG00000090989.13		ENSG00000177885.9		ENSG00000109255.7				ENSG00000128039.6	
		0.01	0.79	-0.025	0.58	0.24	0.001			-0.16	0.006

These SNPs are associated with the expression of nearby **cis-eQTL genes**, including *NMU* (neuromedin U), which encodes a GPCR (G protein-coupled receptor) ligand involved in the progression of NSCLC.

eQTL gene products in EGFR signaling pathways



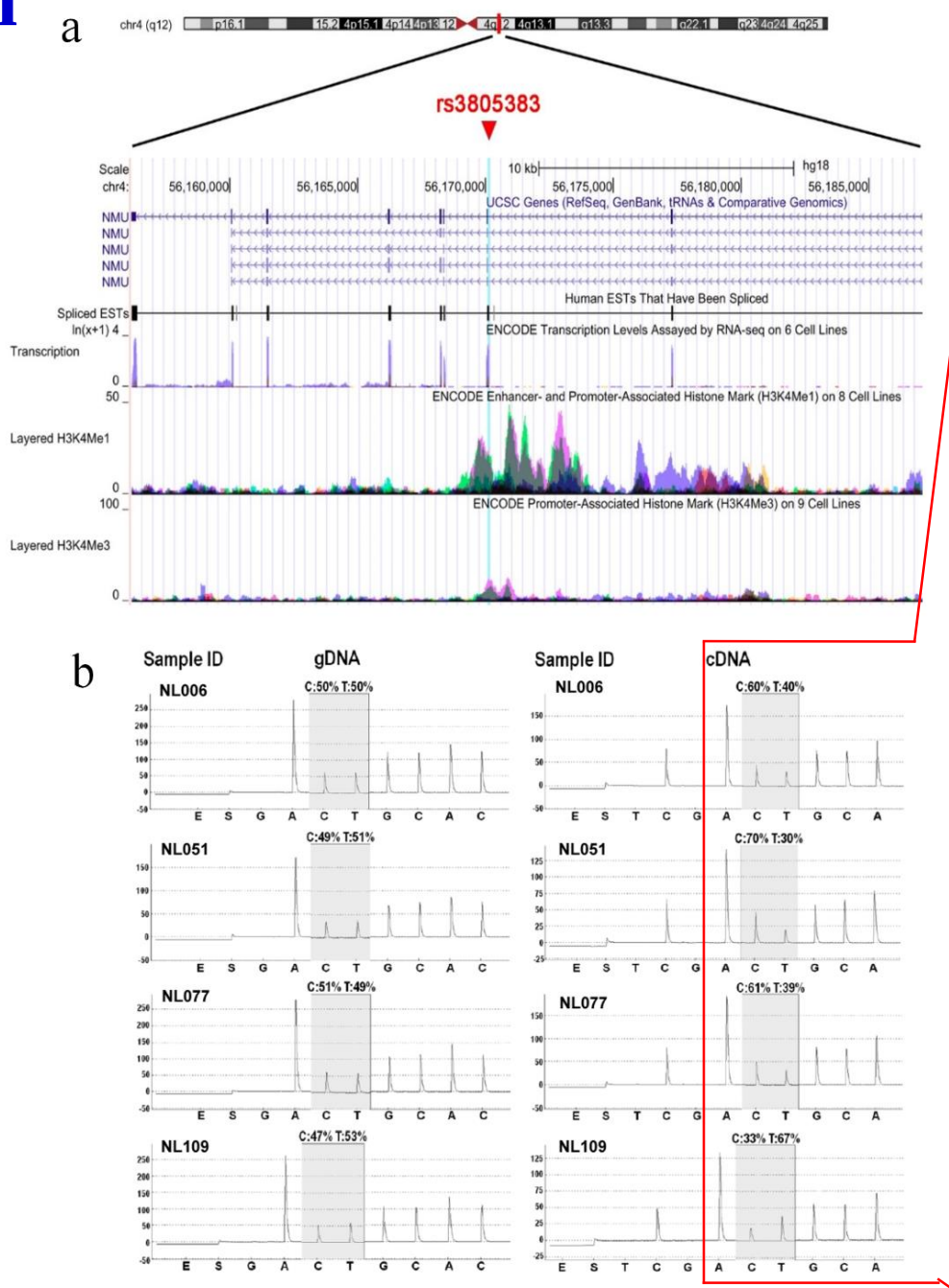
Genomics, 2014, 104 (6), 504-511



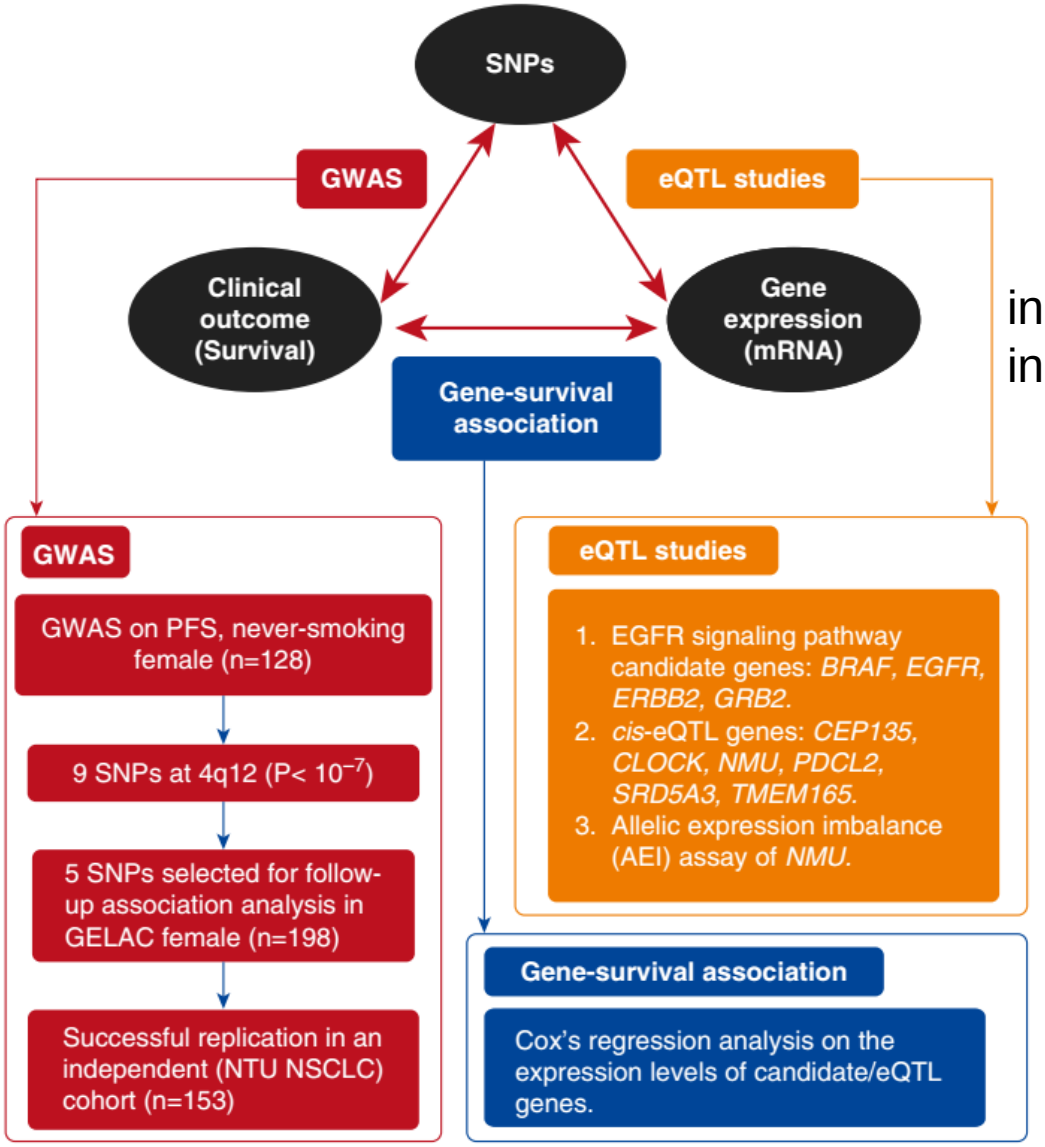
Sci. Signal, 2012 246(5), pe46

Allelic expression imbalance (AEI)

Allelic expression imbalance (AEI) and genomic location and of rs3805383. (b) Pyrograms showing pyrosequencing-based allele quantification at rs3805383 of genomic DNA (gDNA), from buffy coat, and cDNA, from adjacent normal lung tissues, of four patients who showed allele expression imbalance. The ratios in gDNA were always near 1:1, confirming their heterozygosity; and their ratios in cDNA were away from 1:1. The allele specific expression ratios (ASER) for these four patients were **1.48, 2.40, 1.52, and 0.57**, respectively.



Biological plausibility of the Association



in-house,
in public domain

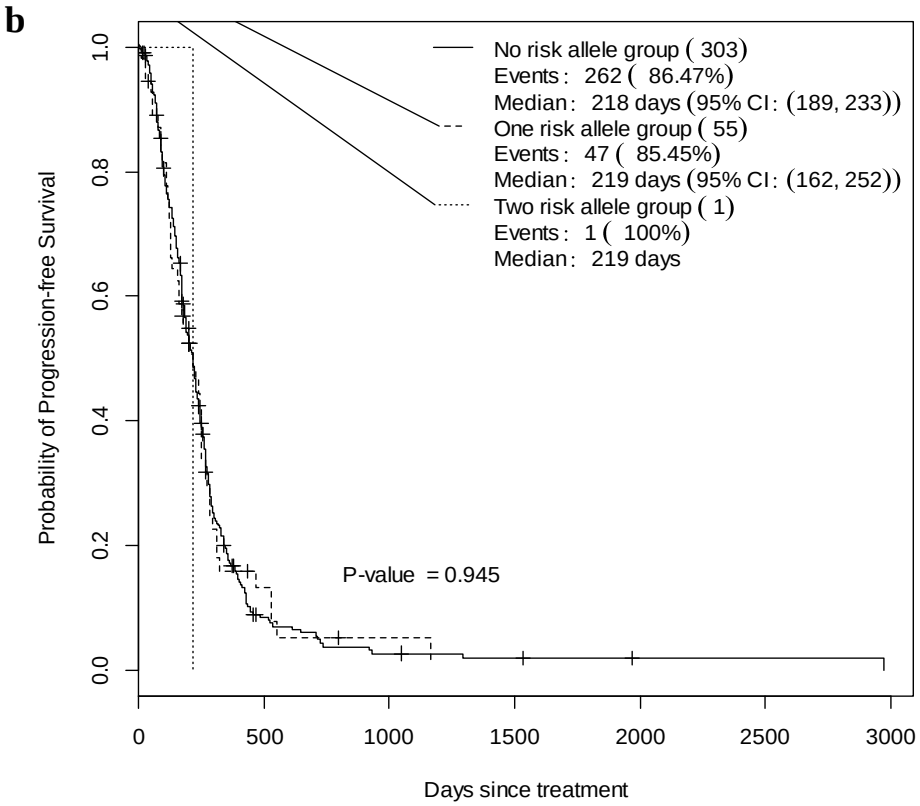
Build Confidence from Integrative Studies

- Association between SNPs and PFS.
- Association between SNPs and gene expression.
- These genes are differentially expressed between lung cancer tissues and adjacent normal tissues; both in-house and in public domain.
- Their expression levels are correlated with time to tumor recurrence.
- These four associations are compatible.

Novel Genetic Markers

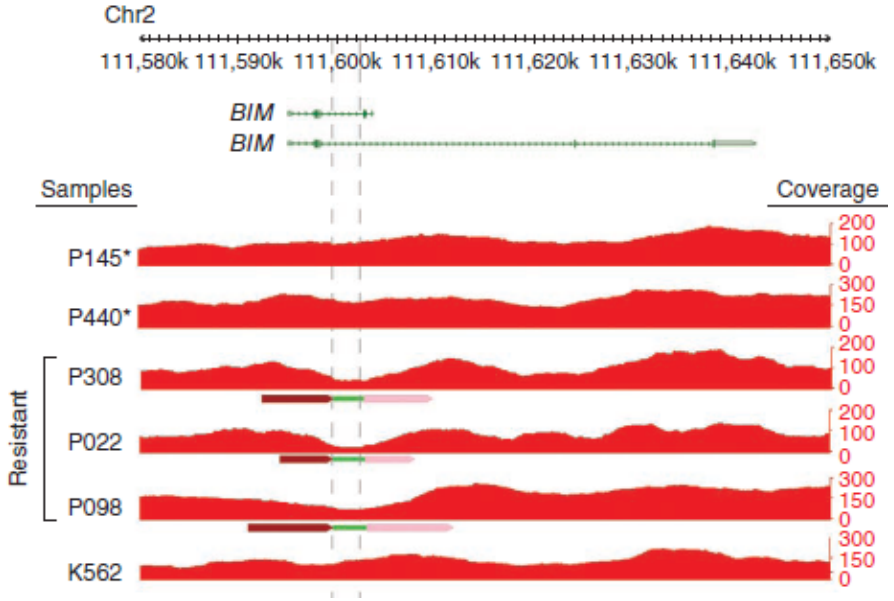
- Not associated with EGFR mutation status
- Nor with *BIM* polymorphism,
- Nor with PFS in never-smoking patients with late-stage lung ADC treated with first-line platinum-based doublets.

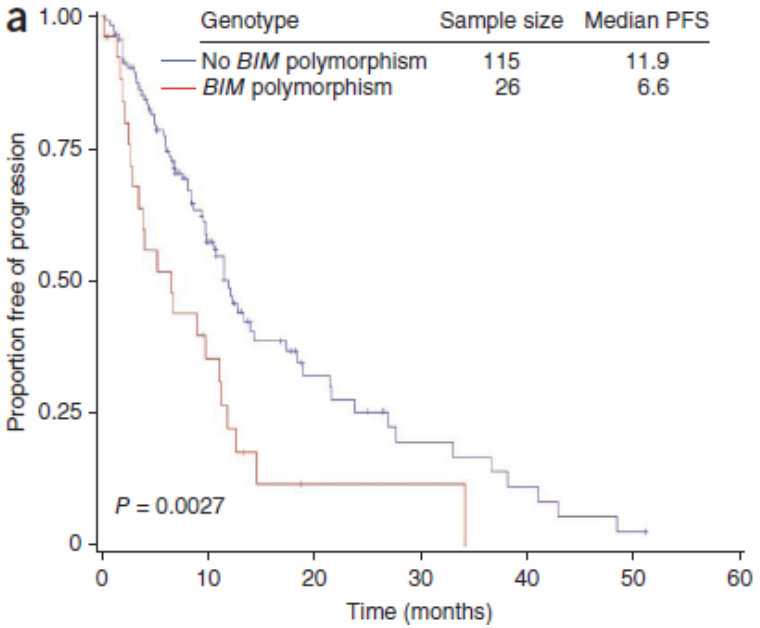
Progression-free survival (PFS) is not significantly associated with the genotype at rs476184 in those treated with first-line platinum-based doublets. (b) Based on PGFNS (all of the 360 female never-smoking patients with lung ADC treated with first-line platinum-based doublets and included in the GELAC).



BIM deletion polymorphism

- *BIM* polymorphisms affecting TKI sensitivity might account for the 20% of TKI-treated individuals with poor responses.
- The *BIM* deletion polymorphism predicts shorter PFS in individuals with *EGFR*-mutant NSCLC treated with EGFR TKI therapy.





Ng et al., *Nature Medicine* 2012

SNPs at 4q12 are not correlated with either *EGFR* mutation status or *BIM* polymorphism, based on LCTCSMB and NTU NSCLC.

	LCTCSMB				NTU NSCLC[a]	
	EGFR[a]		BIM		EGFR[a]	
SNP[c]	OR	P-value[b]	OR	P-value[b]	OR	P-value[b]
rs576732*	0.75	5.71E-01	0.62	5.37E-01	0.57	2.54E-01
rs476184	0.90	8.41E-01	0.64	5.59E-01	0.59	2.82E-01
rs1801260*	0.83	7.22E-01	0.68	6.14E-01	0.59	2.82E-01
rs17725110	0.74	5.69E-01	0.74	6.94E-01	0.64	3.55E-01
rs3805383*	0.74	5.69E-01	0.74	6.94E-01	0.64	3.55E-01

[a] Only wild type and common mutant in EGFR are considered. Frequencies and percentages of EGFR mutation status in LCTCSMB are in Table E9. For the NTU NSCLC cohort, only association with EGFR mutation was studied, using data in ALL-TN, described in Table E1a, which includes the frequencies of EGFR mutation status.

[b] Logistic regression model is used for all the analyses in this table.

[c] Genotype data in LCTCSMB were from Illumina 660W array except those marked with*, which were from imputation. All of the SNP data in NTU NSCLC subjects were obtained using the Taqman assay.

Minor allele frequencies comparison, based on HapMap3.r2

Chr	SNP	Minor allele	Major allele	MAF		
				JPT	CEU	CHB+CHD
4	rs576732	T	C	0.22	0.29	0.07
4	rs476184	G	A	0.2	0.29	0.07
4	rs1801260	G	A	0.21	0.27	0.08
4	rs17725110	G	A	0.22	0.28	0.08
4	rs3805383	A	G	0.22	0.27	0.08

Medical decision based on the genetic variation in this region might encourage **more** lung ADC patients in Japanese population to explore first line treatment other than TKIs.

Conclusions

- Genetic variants in 4q12 merit further investigation to assess their potential as pharmacogenomic predictors for and to understand the biology underlying its influence on PFS in patients treated with TKI therapy.

Clinical Implication

- Among advanced lung ADC patients, the median PFS is about 200 days for first-line gefitinib group, 180 days for first-line chemotherapy group.

  (Mok *et al.* 2009)

- For patients carrying any risk allele in this region, TKI might not be a good choice.

- Key consideration:

Availability of alternative therapy for individuals with high-risk genotype.

Reference

➢*Am J Respir Crit Care Med*. 2017 Mar; 195(5):663-673

(IF: 13.118, Journal ranking: RESPIRATORY SYSTEM : 2/58 3.4%)

ORIGINAL ARTICLE

Genetic Modifiers of Progression-Free Survival in Never-Smoking Lung Adenocarcinoma Patients Treated with First-Line Tyrosine Kinase Inhibitors

I-Shou Chang[1*], Shih Sheng Jiang[1*], James Chih-Hsin Yang[2,3*], Wu-Chou Su[4], Li-Hsin Chien[5], Chin-Fu Hsiao[5,6], Jih-Hsiang Lee[2], Chih-Yi Chen[7,8], Chung-Hsing Chen[1], Gee-Chen Chang[9,10], Zhaoming Wang[11], Fang-Yi Lo[5], Kuan-Yu Chen[12], Wen-Chang Wang[5,13], Yuh-Min Chen[14,15], Ming-Shyan Huang[16], Ying-Huang Tsai[17], Yu-Chun Su[5], Wan-Shan Hsieh[5], Wen-Chi Shih[5], Shwn-Huey Shieh[18,19], Tsung-Ying Yang[10], Qing Lan[20], Nathaniel Rothman[20], Chien-Jen Chen[21], Stephen J. Chanock[20], Pan-Chyr Yang[22], and Chao A. Hsiung[5]

Future work

- Our GWAS, validation studies, eQTL studies are initial steps in pharmacogenomics.
- More replications are needed to confirm these associations.
- Mechanism studies are desirable to understand how genetic variant alters PFS and to provide clue for drug target.
- From pharmacogenomics to implementation.
- Actionable inherited pharmacogenes.

Future work

- Analytic validity: how well the test predicts the presence or absence of a specific genetic variant?
- Clinical validity: how well the genetic variant being analyzed is related to the presence, absence or risk of a specific disease. Sensitivity, specificity, positive predictive value, negative predictive value.
- Clinical utility: whether the test can provide information about diagnosis, treatment, management, or prevention of a specific disease that will be helpful for a consumer.

Other stories

- Early detection, diagnosis, treatment
- EGFR T790M resistance mutation (EGFR T790M) ultimately emerged in most of the patients treated by first generation of TKIs.
- The second generation EGFR-TKIs, afatinib and dacomitinib, were designed to have more potent inhibition of EGFR and to overcome EGFR T790M. Third generation.
- How are variants in 4q12 behave when the second or third generation TKIs are used?

Future work on Pharmacogenomics Studies

- TKI
  - ⇨Gefitinib (Iresssa)
  - ⇨Erlotinib (Tarceva)
  - ⇨Afatinib (Giotrif)
  - ⇨Crizotinib (Xalkori)
  - ⇨……
- Chemotherapy

Lung Cancer in Precision Medicine

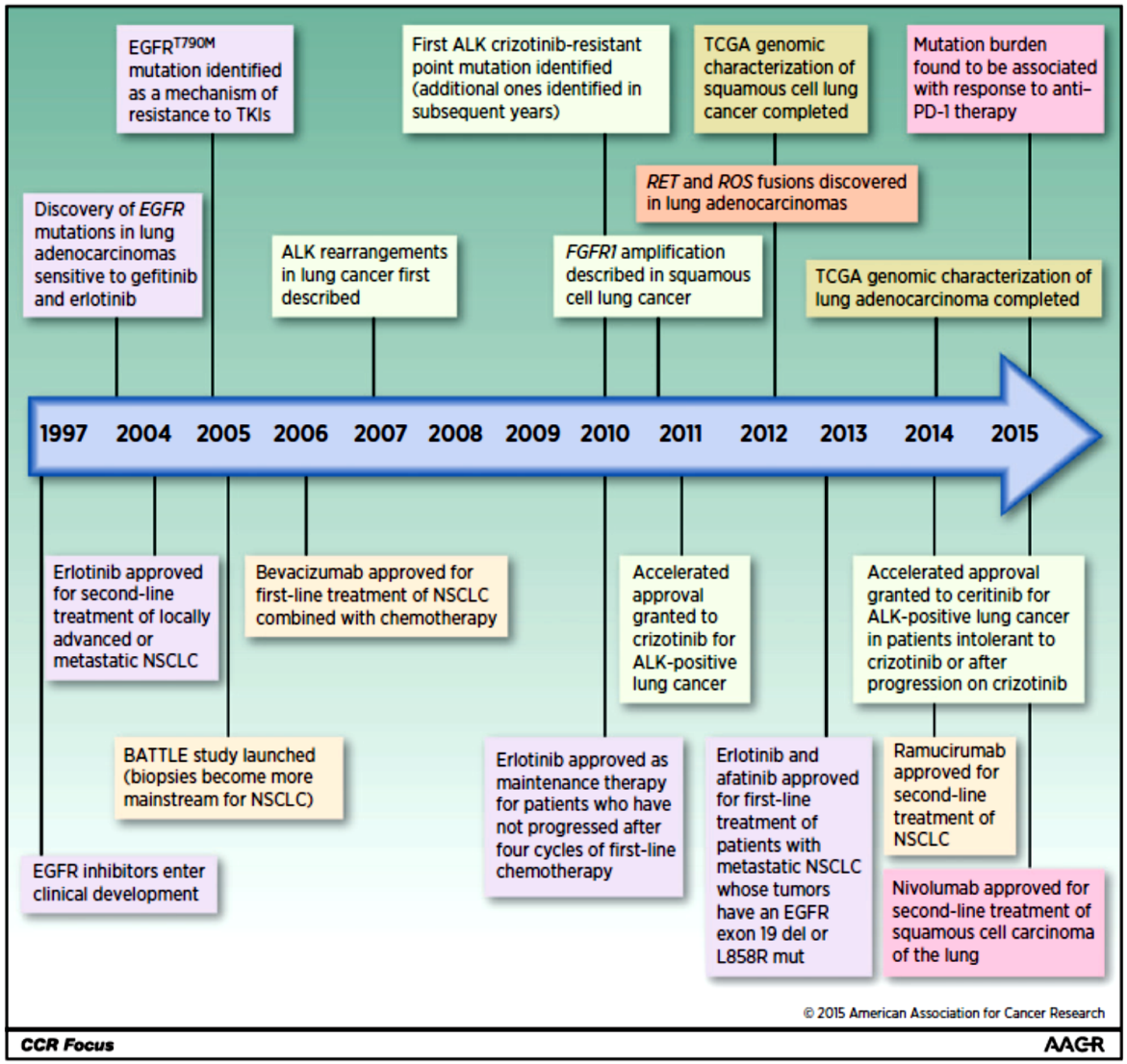


Figure 1.
Timeline of selected major discoveries in lung cancer in recent years (above the arrow) and related clinical trials (below the arrow).

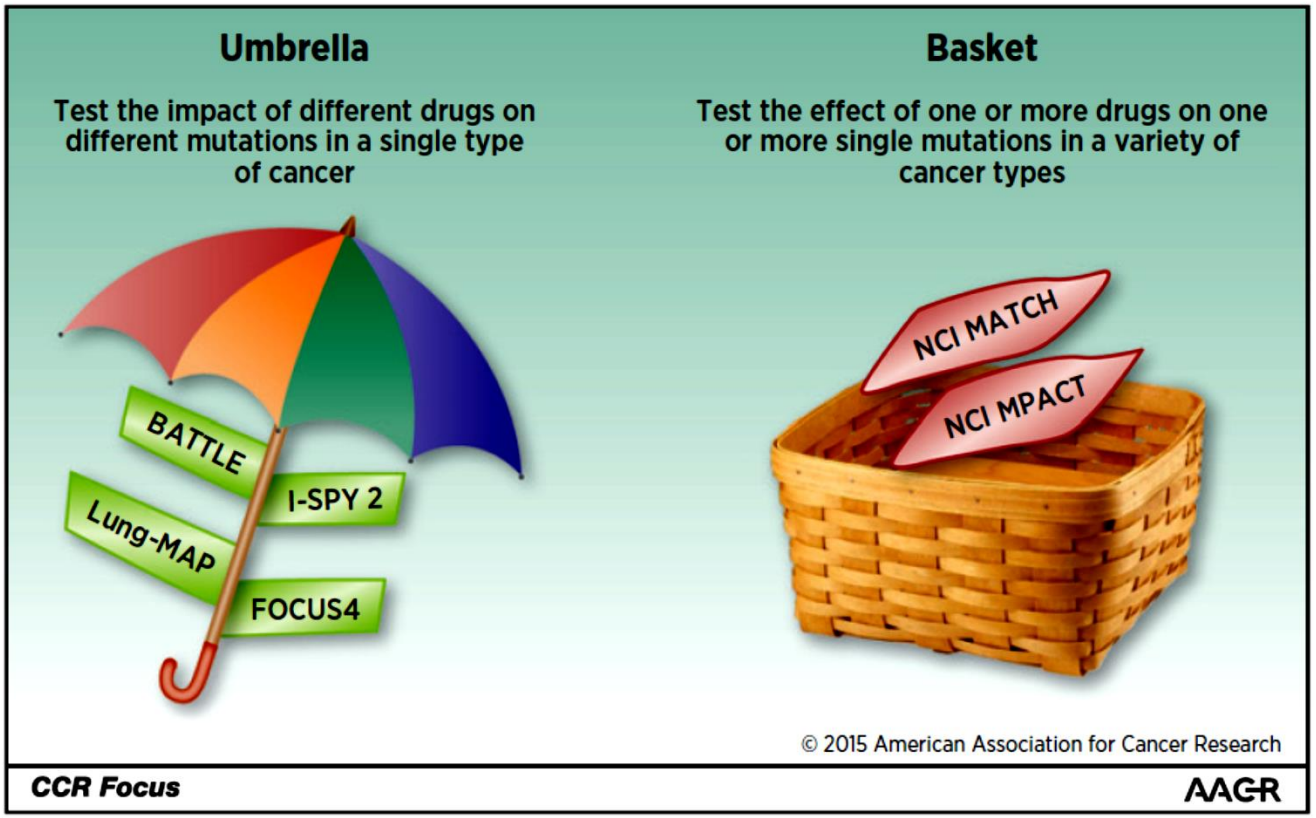
Lung Cancer in the Era of Precision Medicine. Politic & Herbs, *CCR Focus*, 2015

The BATTLE Trial: Personalizing Therapy for Lung Cancer

- Biomarker-integrated Approaches of Targeted Therapy for Lung Cancer Elimination. (Cancer Dis 2011)
- Evaluated utility of targeted therapies in refractory lung cancer, by biopsy-mandated prospective adaptively randomized therapy, based on tissue biomarkers.
- 8 week disease control rate (DCR).
- Feasibility of performing re-biopsies on patients in real time, assigning patients to treatment accordingly, utility of DCR as surrogate for OS.

Clinical Trial Design in Precision Medicine

- New trial designs have been used to match the right drug to the right patient at the right time.



Lung Cancer in the Era of Precision Medicine. Politic & Herbs, *CCR Focus*, 2015

BATTLE-2 (JCO 2016)

- Umbrella studies: to test the impact of different drugs on different mutations in a single type of cancer.
- To facilitate screening and accrual in view of low prevalence biomarkers.
- Basket studies: to test a single drug in patients with a single gene alteration regardless of the primry tissue, based on the idea that the presence of a molecular marker predicts response to therapy independent of tumor histology.

Acknowledgements

NHRI

- Dr. I.S. Chang
  Dr. S.S. Jiang
  PostDoctorals & RAs

GELAC

- Dr. C.J. Chen
- Dr. P.C. Yang 、 Dr. K.Y. Chen
  Dr. W.C. Su 、 Dr. J.T. Chang
  Dr. Y.H. Tsai 、 Dr. Y.M. Chen
  Dr. M.S. Huang 、Dr. C.Y. Chen

NCI

- NCI ：Dr. Nat Rothman、 Dr. Qing Lan
  Dr. Stephen Chanock
- FLCCA Consortium

NTUH

- Dr. C.H. Yang