Intrinsic subtypes, PIK3CA mutation, and the degree of benefit from adjuvant trastuzumab in NSABP trial B-31

Pogue-Geile, et al

Intrinsic subtypes, PIK3CA mutation, and the degree of benefit from adjuvant trastuzumab in NSABP trial B-31

Pogue-Geile, et al

Supplementary Files

SECTION I

TEXT, FIGS A and B

Power calculation for PAM50 and PIK3CA

To test the predictive effect of PAM50 and PIK3CA, we have tested the null hypothesis that there is no interaction between treatment and group (benefit group or non-benefit group) using a two-arm survival interaction test with the method developed by Peterson and George [1] using the SWOG webtool (http://www.swogstat.org/stat/public/int_survival.htm).

With the assumption of exponential distribution, the formula for power calculation is as follows:

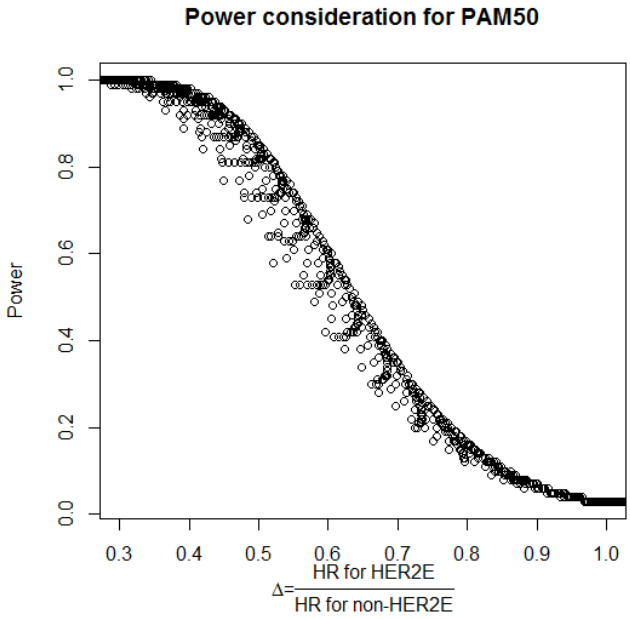
$$Z_{1-\beta} = \sqrt{N(log\Delta)^2 / \sum \frac{1}{(e_{ij}f_{ij})}} - Z_{1-\propto/2}$$

in which, N is the sample size; f_{ij} are the frequencies in treatment *i* (ACTH or ACT) in stratum *j* (benefit group or non-benefit group); e_{ij} are the event probabilities in treatment *i* and stratum *j*, with the exponential failure assumptions, e_{ij} is calculated using the accrual, minimum follow-up, competing risk, and the hazard rate λ_{ij}, where λ_{ij} represent hazard rate in treatment *i* and stratum *j*; and Δ is defined as $\Delta = \frac{\lambda_{11}/\lambda_{21}}{\lambda_{12}/\lambda_{22}}$. In this study, accrual is 5 years, minimum follow-up is 5 years, with DFS endpoint, competing risk is 0. For power consideration, variables α, N, f_{ij}, Δ, and λ_{ij} need to be specified.

For power calculations for the PAM50 analyses, the null hypothesis is that there is no interaction between treatment and intrinsic groups (HER2 enriched versus other subtypes). In this study, there are 1,578 patients and about 50% of these patients are HER2 enriched (HER2E). For patients with a PAM50 assignment, the observed ten-year DFS was 66.4% for the ACT arm and 78.9% for patients in the ACTH arm. Assuming exponential distribution, λ_1 (the hazard rate for ACT) is 0.041 and λ_2 (hazard rate for ACTH arm) is 0.024. To estimate the power and Δ ($\frac{HR\ for\ HER2E}{HR\ for\ other\ subtypes}$), we assumed λ_{11} (the hazard rate for patients who are **not HER2E** treated by **ACT**) to be in the range of (0.01, 0.07) and λ_{21} (the hazard rate for patients who are **not HER2E** and treated by **ACTH**) to be in the range of (0.01, 0.04). Because it is hypothesized that HER2E patients will receive more benefit from trastuzumab than non-HER2E patients, we focus on the results with Δ less than 1 and HR for patients with either HER2E or other subtypes is less than 2. As shown in the figure A, to

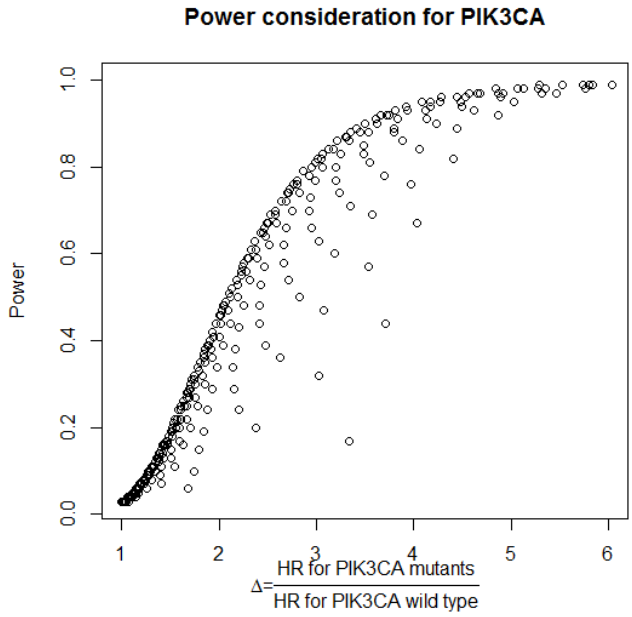
have a power larger than 0.6, the HR must be less than 0.6, to have power larger than 0.8, the HR needs to be less than 0.5.

Figure A



For power calculations for PIK3CA mutational status analyses, the null hypothesis is that there is no interaction between trastuzumab and PIK3CA mutational status. In this study, there were 671 patients for the PIK3CA study, among whom about 20% of the tumors had PIK3CA mutations. Similarly to the power consideration of PAM50, we assumed λ_{11} (the hazard rate for patients who are PIK3CA wild type [WT] treated by ACT) to be in the range of (0.01, 0.07) and λ_{21} (the hazard rate for patients who are PIK3CA WT treated by ACTH) to be in the range of (0.01, 0.04). Because it is hypothesized that patients with mutant PIK3CA will receive less benefit from trastuzumab than will those with WT PIK3CA, we focused on the results with a Δ larger than 1 and HR for patients with either WT or mutant is less than 2. As shown in the figure B, to have a power larger than 0.6, the HR must be larger than 2.3, to have power larger than 0.8, the HR needs to be greater than 2.9.

Figure B



Details of normalization of NSABP B-31 nCounter gene expression

To identify intrinsic subtype, we first normalized B-31 log- transformed nCounter gene expression data to remove technical platform differences and patient population differences between B-31 (who were mainly HER2 positive) and the original PAM50 development cohort published by Parker et al (which included all breast cancer).[2] We randomly sampled 90 patients from B-31 without replacement, so that it had a similar subtype distribution of ER and HER2 as the original PAM50 training dataset by Parker et al. We then calculated mean expression values for each gene within the constructed reference cohort. To stabilize the estimation random resampling was repeated 50 times and the mean for each gene was calculated, then we subtract the reference mean expression values for each gene were subtracted.

Detailed Methods for Mutation Profiling with TyplePlex chemistry and utilizing the Sequenom MassArray

Primer sequences for amplification and extention are available upon request. PCR and extension reaction conditions have been described previously.[3] Reactions with >25% unextended primer were excluded. Three personnel independently reviewed mass spectra of any allele with a frequency of 5% or more. Any discordances among the three reviewers were decided by joint re-inspection. This manual review was necessary to identify "real mutant" peaks from background peaks or shoulders. Detection of a mutation at any base was

considered evidence of mutation, whereas missing data for any reaction resulted in missing PIK3CA status only if all other positions were WT.

REFERENCES

1. Peterson B, George SL. Sample size requirements and length of study for testing interaction in a 2 x k factorial design when time-to-failure is the outcome [corrected]. *Controlled clinical trials* 1993; 14(6): 511-22.

2. Parker JS, Mullins M, Cheang MC, et al. Supervised risk predictor of breast cancer based on intrinsic subtypes. *J Clin Oncol* 2009; 27(8): 1160-7.

3. Fumagalli D, Gavin PG, Taniyama Y, et al. A rapid, sensitive, reproducible and cost-effective method for mutation profiling of colon cancer and metastatic lymph nodes. *BMC cancer* 2010; 10: 101.

SECTION II

ALL SUPPLEMENTARY TABLES (1, 2, and 3)

Supplementary Table 1. Demographics of Patients, NSABP B-31

Variable	Category	Eligible patients		Patients profiled by PIK3CA mutation	Patients profiled by nCounter
Total		2068		671	1578
Age	<60	1730 (83.7%)		564 (84.1%)	1324 (83.9%)
	≥60	338 (16.3%)		107 (15.9%)	254 (16.1%)
Tumor size	≤2 cm	823 (39.8%)		246 (36.7%)	584 (37%)
	2.1−5 cm	1047 (50.6%)		362 (53.9%)	841 (53.3%)
	>5 cm	191 (9.2%)		62 (9.2%)	148 (9.4%)
	Unknown	7 (0.3%)		1 (0.1%)	5 (0.3%)
Node	≥1, ≤3	1189 (57.5%)		385 (57.4%)	895 (56.7%)
	≥4, ≤9	598 (28.9%)		188 (28.0%)	471 (29.9%)
	≥10	281 (13.6%)		98 (14.6%)	212 (13.4%)
ER	Negative	978 (47.3%)		318 (47.4%)	737 (46.7%)
	Positive	1090 (52.7%)		353 (52.6%)	841 (53.3%)
HER2	Negative			60 (8.9%)	146 (9.3%)
	Positive			609 (90.8%)	1417 (89.7%)
	NA			2 (0.3%)	15 (1%)
DFS	Censored	1511 (73.1%)		481 (71.7%)	1146 (72.6%)
	Event	557 (26.9%)		190 (28.3%)	432 (27.4%)
Median follow up, years		6.84		6.93	7.03

Note: Eligible patients refers to patients who have follow up information, with known ER and nodal status, and have consented for either future cancer research or HER2 research. The p-value was calculated using chi-square test.

Supplementary Table 2. Consistency of Identifying Intrinsic Subtypes Using 50 Genes versus 49 Genes. Results are based on the Parker dataset (n=200).

		Intrinsic subtype identification based on 49 genes				
		Basal-like	Her2 Enriched	Luminal A	Luminal B	Normal-like
Intrinsic subtype identification based on 50 genes	Basal-like	49	0	0	0	0
	Her2 enriched	0	30	0	0	0
	Luminal A	0	0	51	0	0
	Luminal B	0	0	1	38	0
	Normal-like	0	0	0	0	31

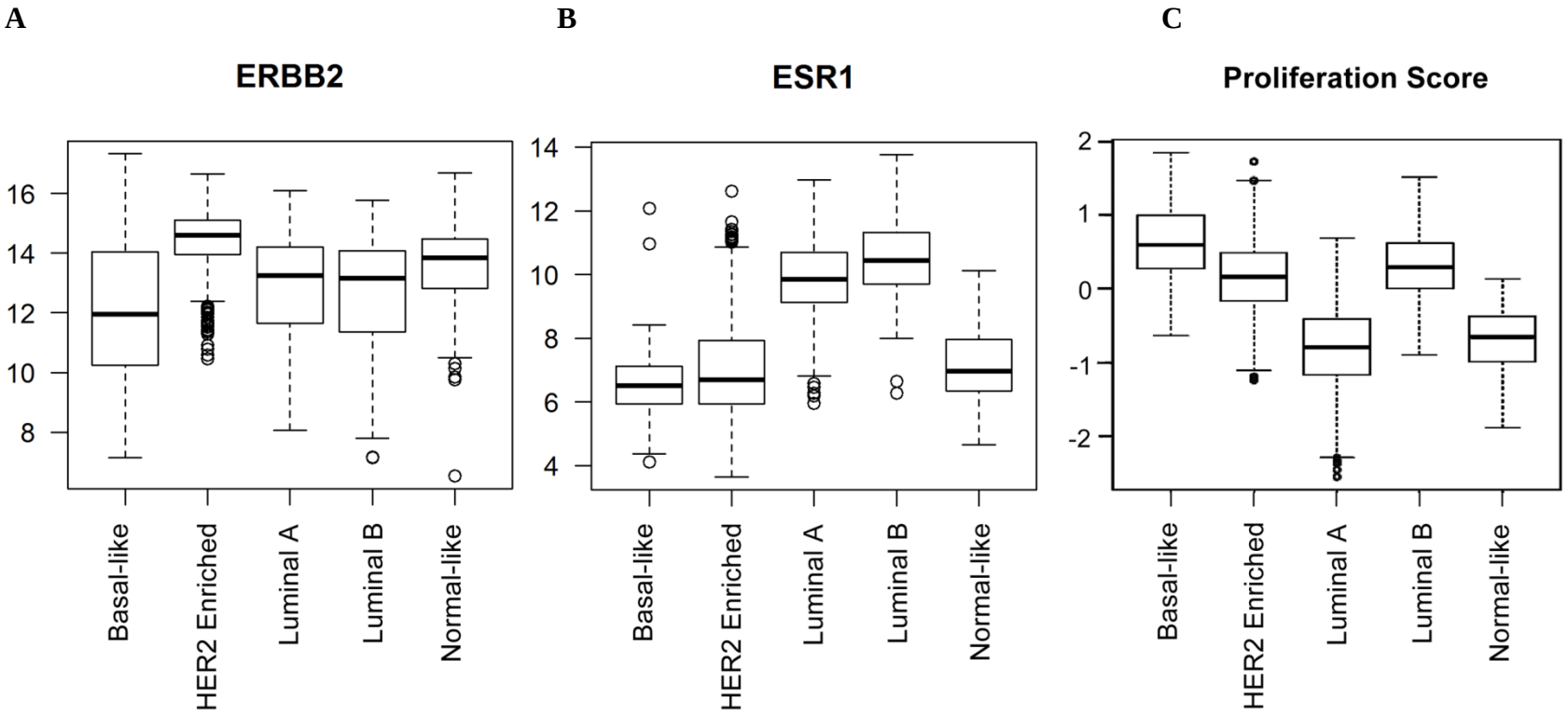
Supplementary Table 3. Frequency of PIK3CA Mutations in NSABP B-31 Cohort

Mutations	**N**	**Frequency**
PIK3CA_C420R	4	0.60%
PIK3CA_E542K	21	3.13%
PIK3CA_E545K	28	4.17%
PIK3CA_H1047L	10	1.49%
PIK3CA_H1047R	94	13.99%
PIK3CA_N345K	9	1.34%
PIKexon9	49	7.29%
PIKexon20	104	15.48%
PIK3CA_MUTATED	166	24.70%

SECTION III

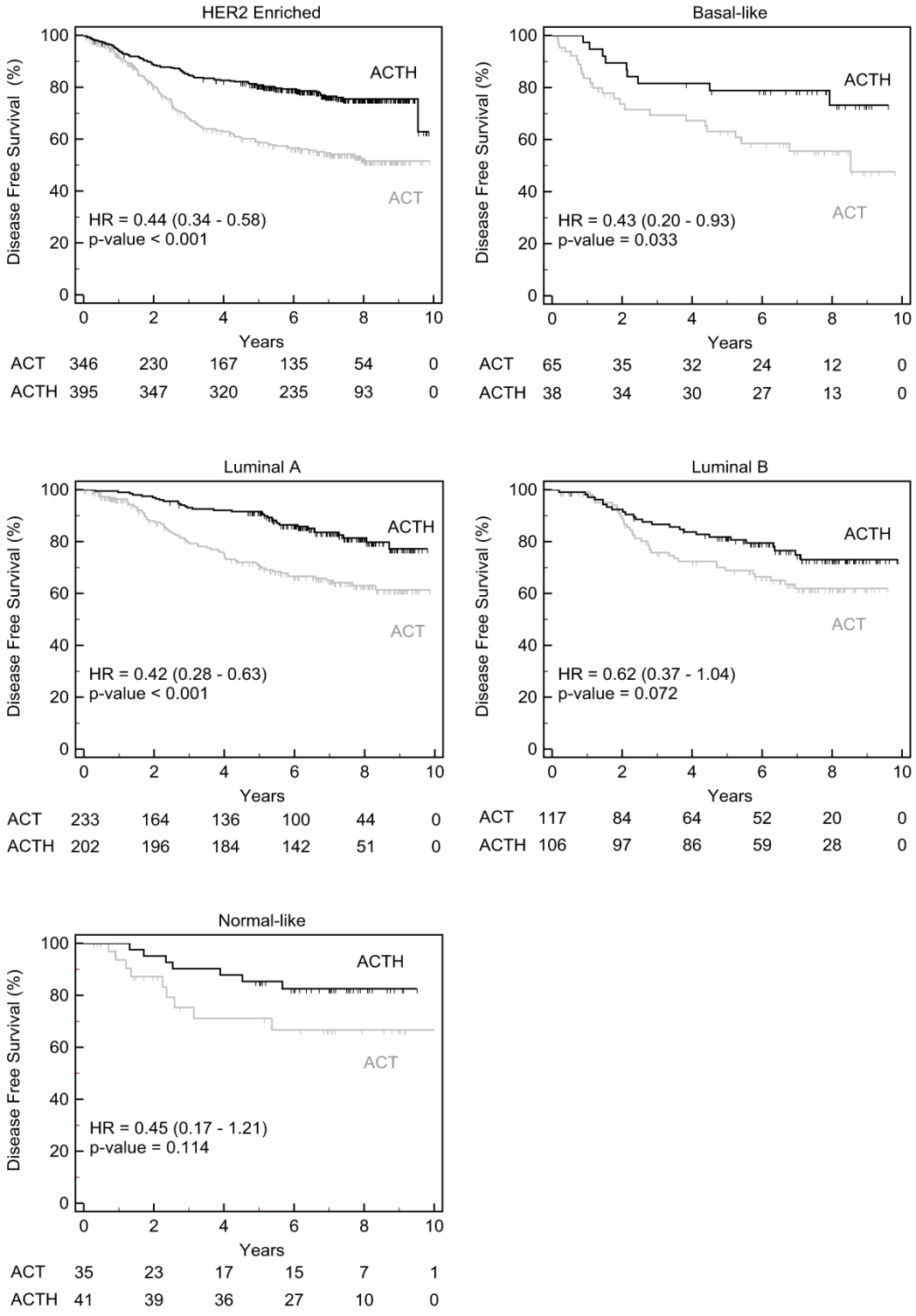
SUPPLEMENTARY FIGURES 1 – 4

Supplementary Figure 1. Boxplot of the association among intrinsic subtypes and expression levels of ERBB2, ESR1, and proliferation genes.



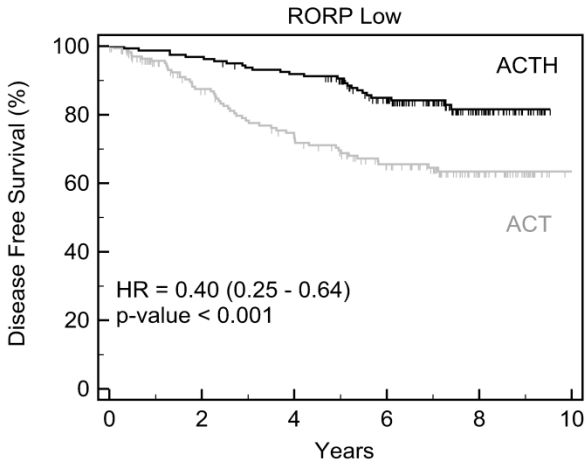
Supplementary Figure 2

Disease-free survival (DFS) for HER2-enriched, Basal-like, Luminal A, and Luminal B patients treated with adriamycin and cyclophosphamide followed by taxol (ACT) versus those treated with adriamycin and cyclophosphamide followed by taxol plus trastuzumab (ACTH).

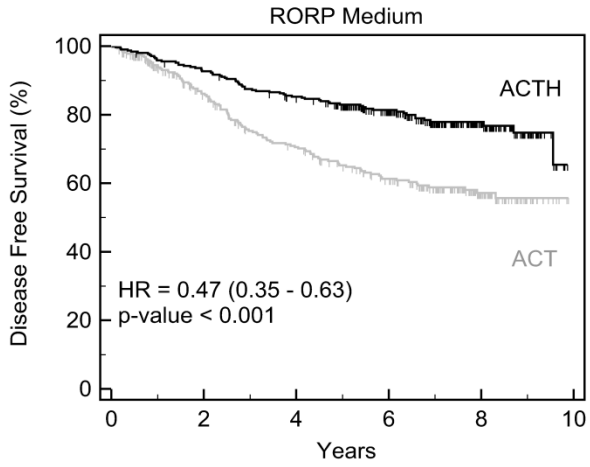


Supplementary Figure 3

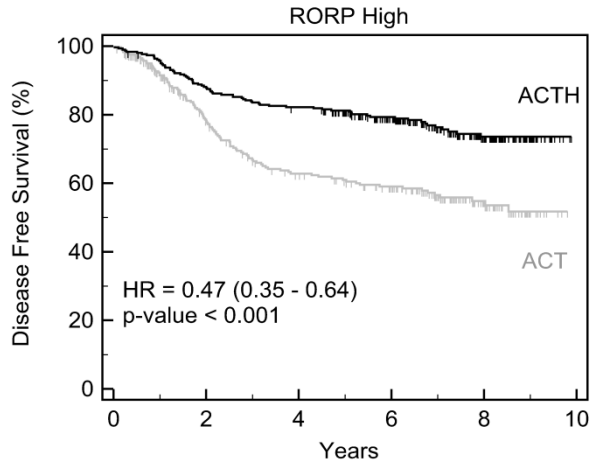
Disease-free survival (DFS) for RORP low, medium, and high patients, and RORS low, medium, and high patients treated with adriamycin and cyclophosphamide followed by taxol (ACT) versus those treated with adriamycin and cyclophosphamide followed by taxol plus trastuzumab (ACTH).



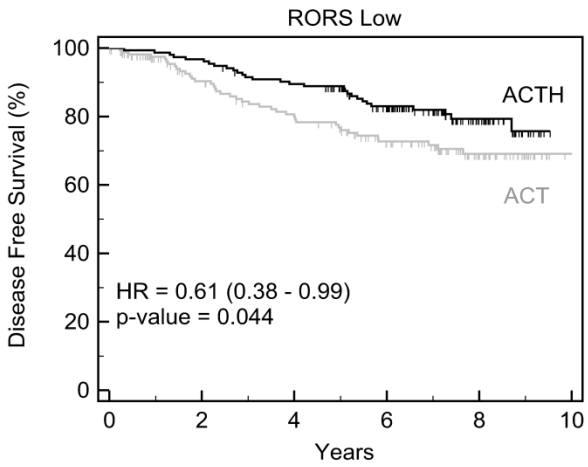
ACT	174	125	101	75	38	1
ACTH	161	156	146	112	40	0



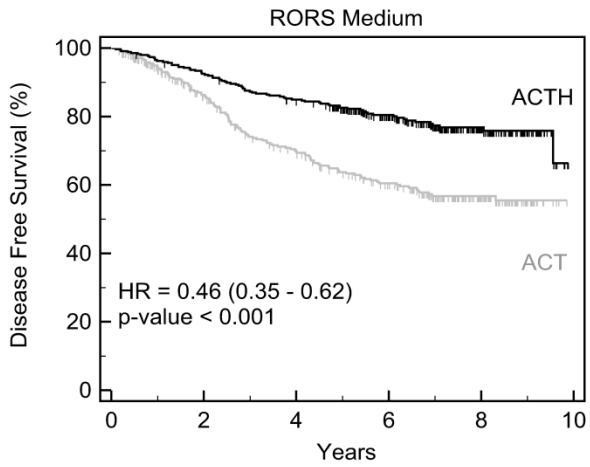
ACT	329	229	177	136	55	0
ACTH	315	290	261	190	72	0



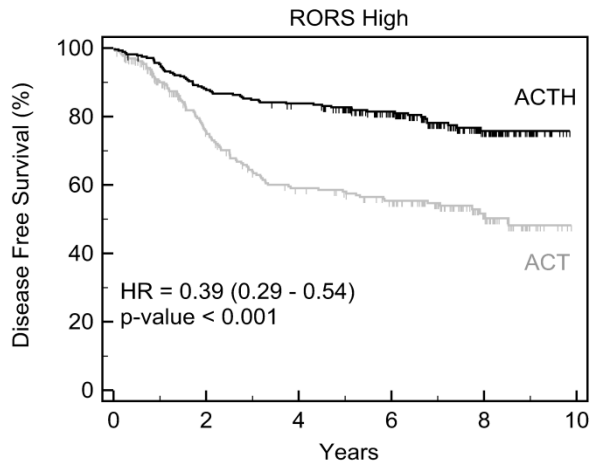
ACT	293	182	138	115	44	0
ACTH	306	267	249	188	83	0



ACT	171	124	106	81	37	1
ACTH	154	149	136	106	40	0

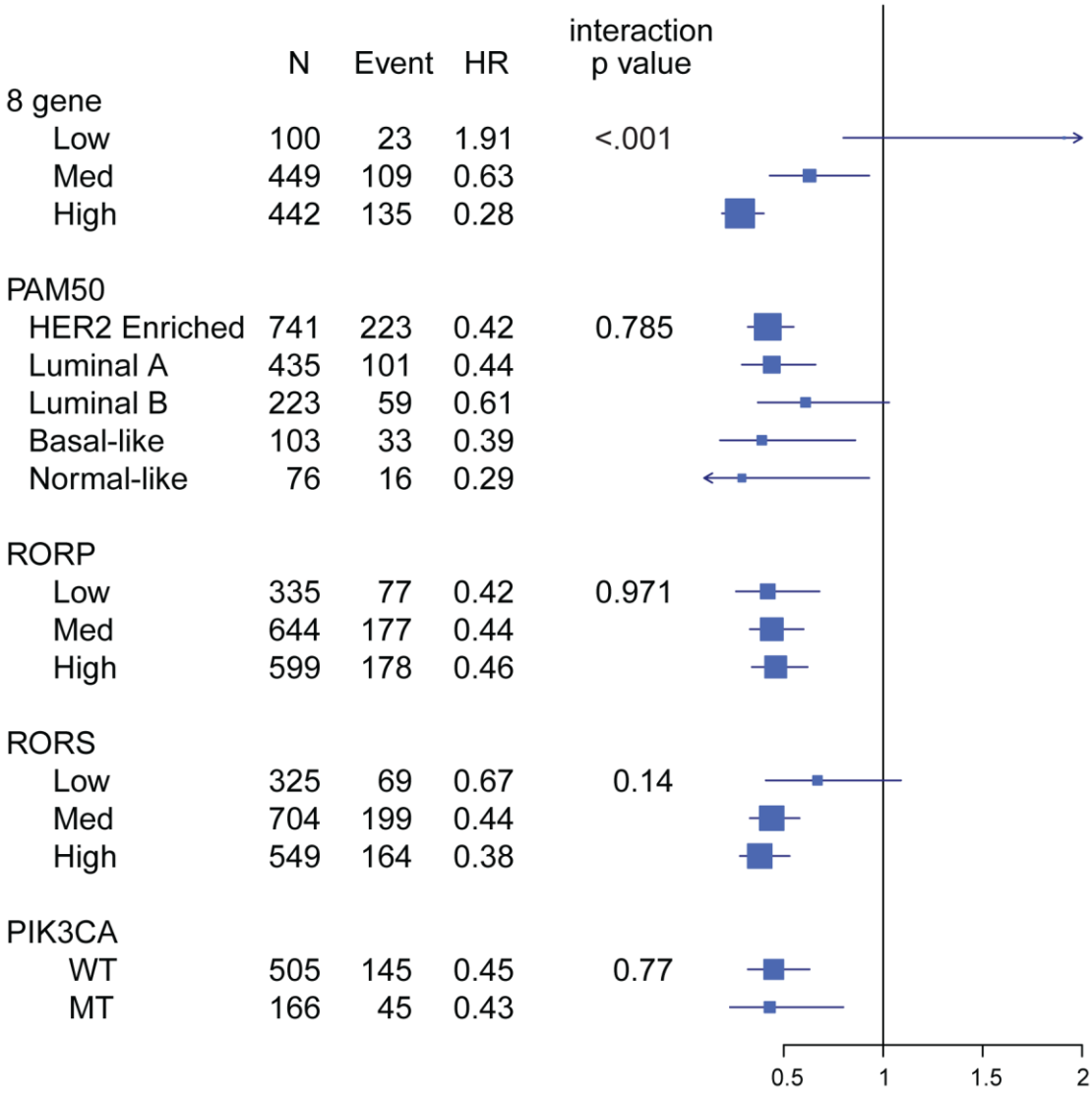


ACT	356	251	191	144	61	0
ACTH	348	320	290	208	79	0



ACT	269	161	119	101	39	0
ACTH	280	244	230	176	76	0

Supplementary Figure 4. Forest plot of hazard ratios (HR) for subtypes identified by the 8-gene predictive model, PAM50, RORP, RORS, and PIK3CA mutational status. HR and interaction p-values were determined by multivariate analysis by controlling for nodal status, ER status, age, and tumor size. With multiple hypothesis adjustment, only the 8-gene model, (which was analyzed on confirmation cohort only), has significant treatment-group interaction.



SECTION IV

R Code for PAM 50 analysis

```
############################
# Code to identify PAM50 subtype based on B31 nanostring data
##############################
rm(list=ls())
library(ctc)  # required by PAM50
library(heatmap.plus)  # required by PAM50
setwd("C:\\Users\\nsong\\Desktop\\Research\\B31-subtypes\\")
library(survival)



############################
# PAM50 for Nanostring
########################
# load data
# nano_clinical_info: include clinical information
# nano_gse: gene expression data
load("Final_results_2014\\nano_gse.rdata")
nano_gse = nano_gse[rownames(nano_clinical_info),]

# First, we need to normalize the data because
#   - the patient distribution in B-31 is not similar to the training dataset of PAM50
#   - different technical platform

# To normalize data,
# 1. re-sample a subset of patients who have similar distribution to that of the PAM50 training dataset
```

```
# 2. subtract the mean value of each gene which was estimated from the re-sampling

# Note: In the PAM50 paper, it is specifically mentioned that ER positive patients is about 50%
#    but there is no data about HER2 distribution, because in other datasets, the HER2 neg vs pos is about 3:1
#    when we resampl, we aim to have the following distribution as specified below:
#      - Her2 pos, ER pos: 12%
#      - Her2 neg, ER pos: 38%
#      - Her2 pos, ER neg: 12%
#      - Her2 neg, ER neg: 12%
#    we also experimented with other distributions of patients by varying HER2 distribution, although there are
#    some slight changes for the identification, the conclusion about PAM50 vs 2 benefit is very similar

set.seed(123)
num_samples = nrow(nano_clinical_info)
index = (1:num_samples)[!is.na(nano_clinical_info[,'ER'])&!is.na(nano_clinical_info[,'her2_info'])]
her2_pos_er_pos_index = index[nano_clinical_info[index,'ER']=='2' &
nano_clinical_info[index,'her2_info']=='1']
her2_neg_er_pos_index = index[nano_clinical_info[index,'ER']=='2' &
nano_clinical_info[index,'her2_info']=='0']
her2_pos_er_neg_index = index[nano_clinical_info[index,'ER']=='1' &
nano_clinical_info[index,'her2_info']=='1']
her2_neg_er_neg_index = index[nano_clinical_info[index,'ER']=='1' &
nano_clinical_info[index,'her2_info']=='0']

MAdata = t(nano_gse)
f_mean = matrix(NA, dim(MAdata)[1], 50)

her2_pos_er_pos = 0.12 # her2 neg, er pos
her2_neg_er_pos = 0.38#, her2 pos, er pos
her2_pos_er_neg = 0.12 #, her2 neg Er neg
```

```
her2_neg_er_neg = 0.38 # her2 pos, er pos


total = 90
for (i in 1:50){
        select_index = c(sample(her2_neg_er_neg_index, total*her2_neg_er_neg),
                                        sample(her2_neg_er_pos_index, total*her2_neg_er_pos),
                                        sample(her2_pos_er_pos_index, total*her2_pos_er_pos),
                                        sample(her2_pos_er_neg_index, total*her2_pos_er_neg))
        f = MAdata[, select_index]
        f_mean[,i]= apply(f,1, mean)
}



MA.mean_val = apply(f_mean, 1, mean)
MA.mean = t(matrix(rep(MA.mean_val, each=dim(MAdata)[2]),nrow=dim(MAdata)[2]))
MA.stand = MAdata-MA.mean
MAdata[1, 1:5] - MA.stand[1, 1:5]
foo = cbind(rownames(MA.stand), MA.stand)
fn = "B31_PAM50_2014\\nano_pam50.txt"
write.table(foo, fn, row.names=F, sep="\t")



inputFile <- "nano_pam50.txt" # the input data matrix as a tab delimited text file
short<-"nano_pam50" # short name that will be used for output files

calibrationParameters<- -1

hasClinical<-FALSE
```

```
####
# run the assignment algorithm
####

source(paste(paramDir,"subtypePrediction_functions.R",sep="/"))
source(paste(paramDir,"subtypePrediction_distributed.R",sep="/"))
```