

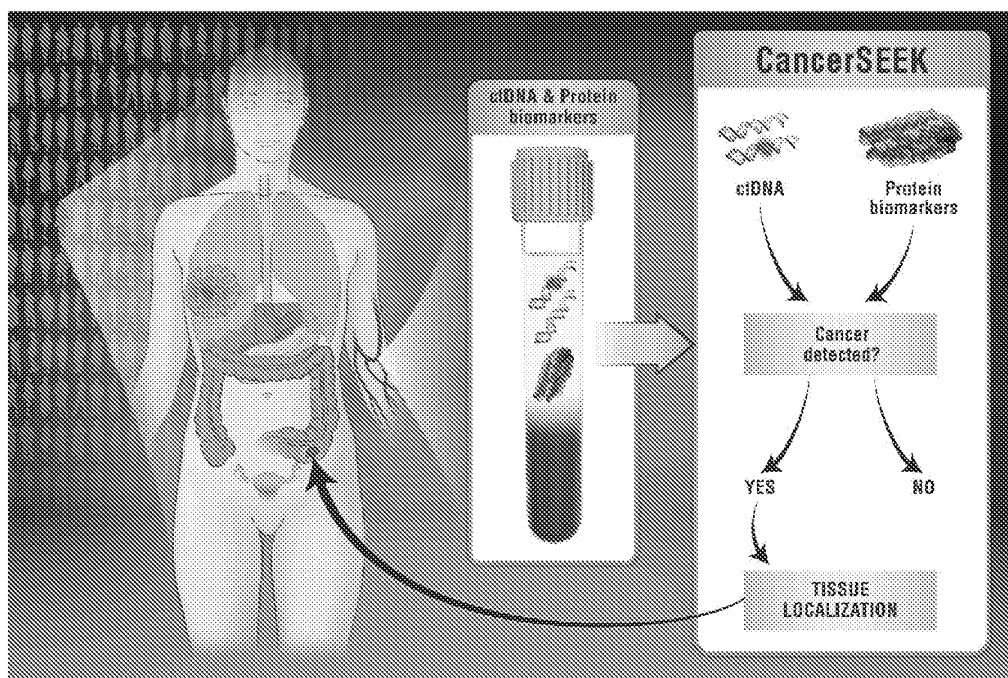


FIG. 1

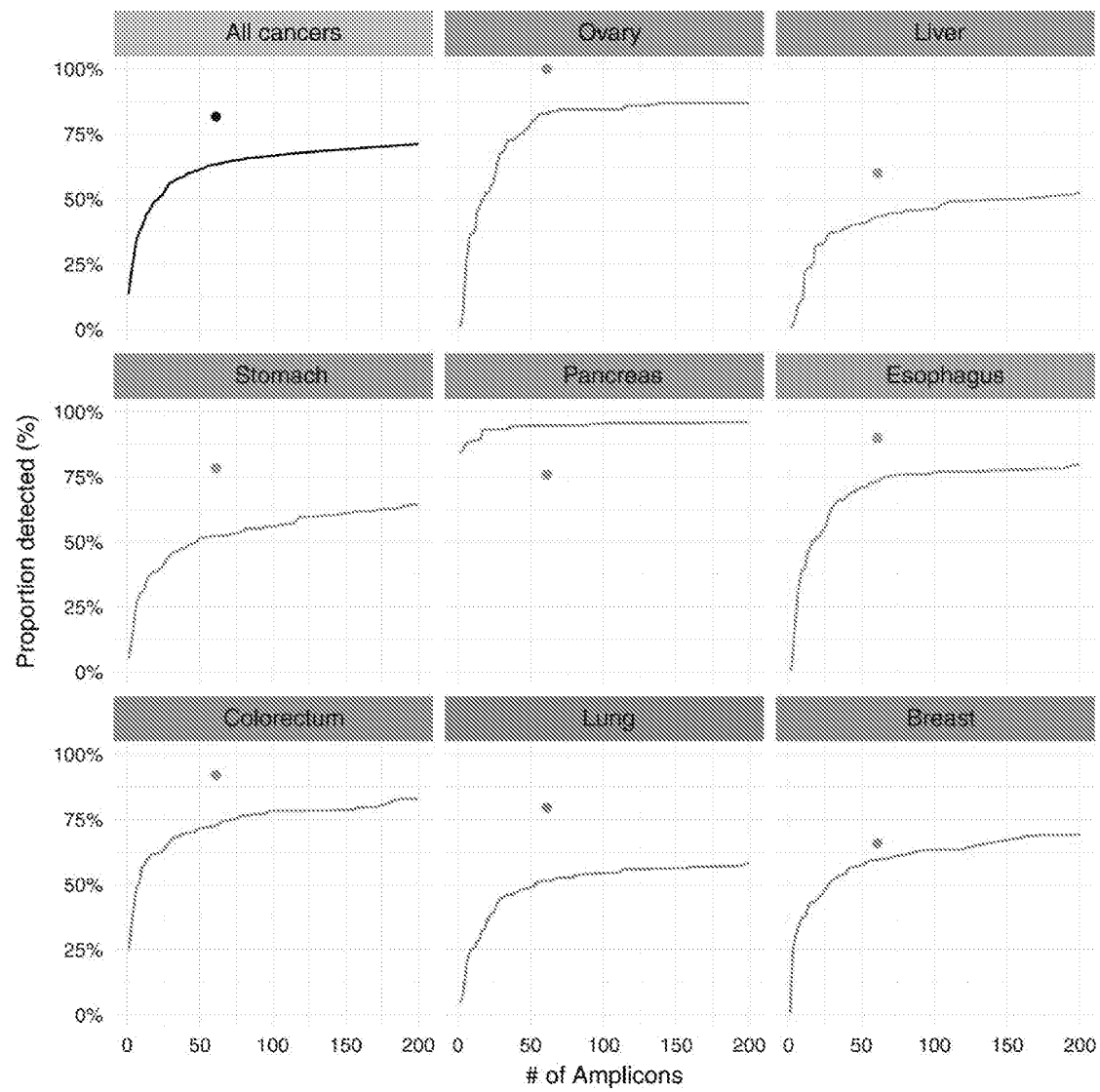


FIG. 2

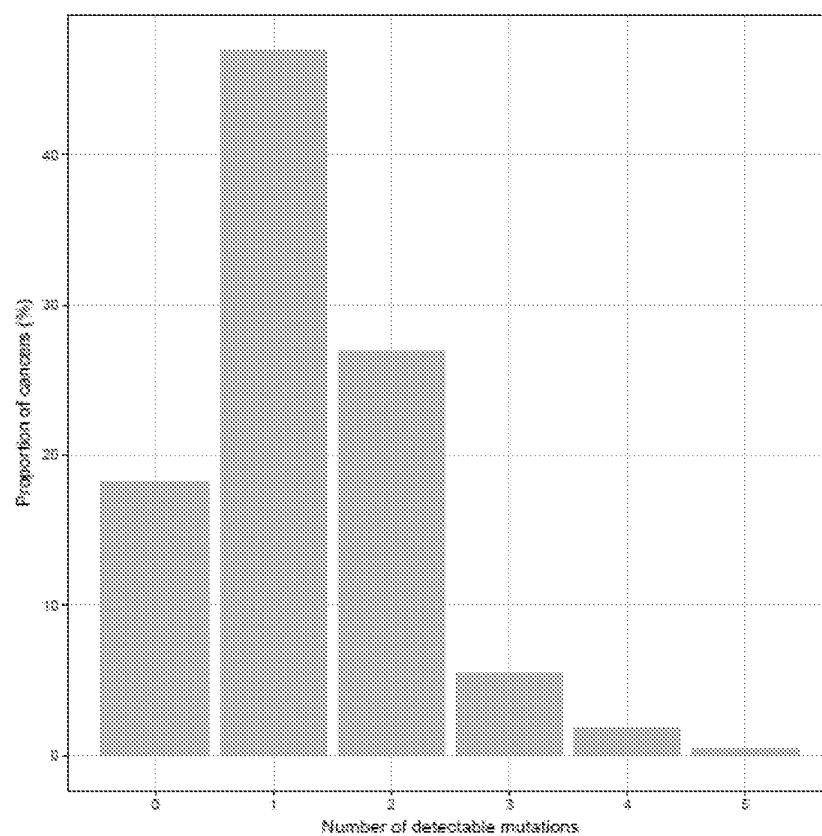


FIG. 3

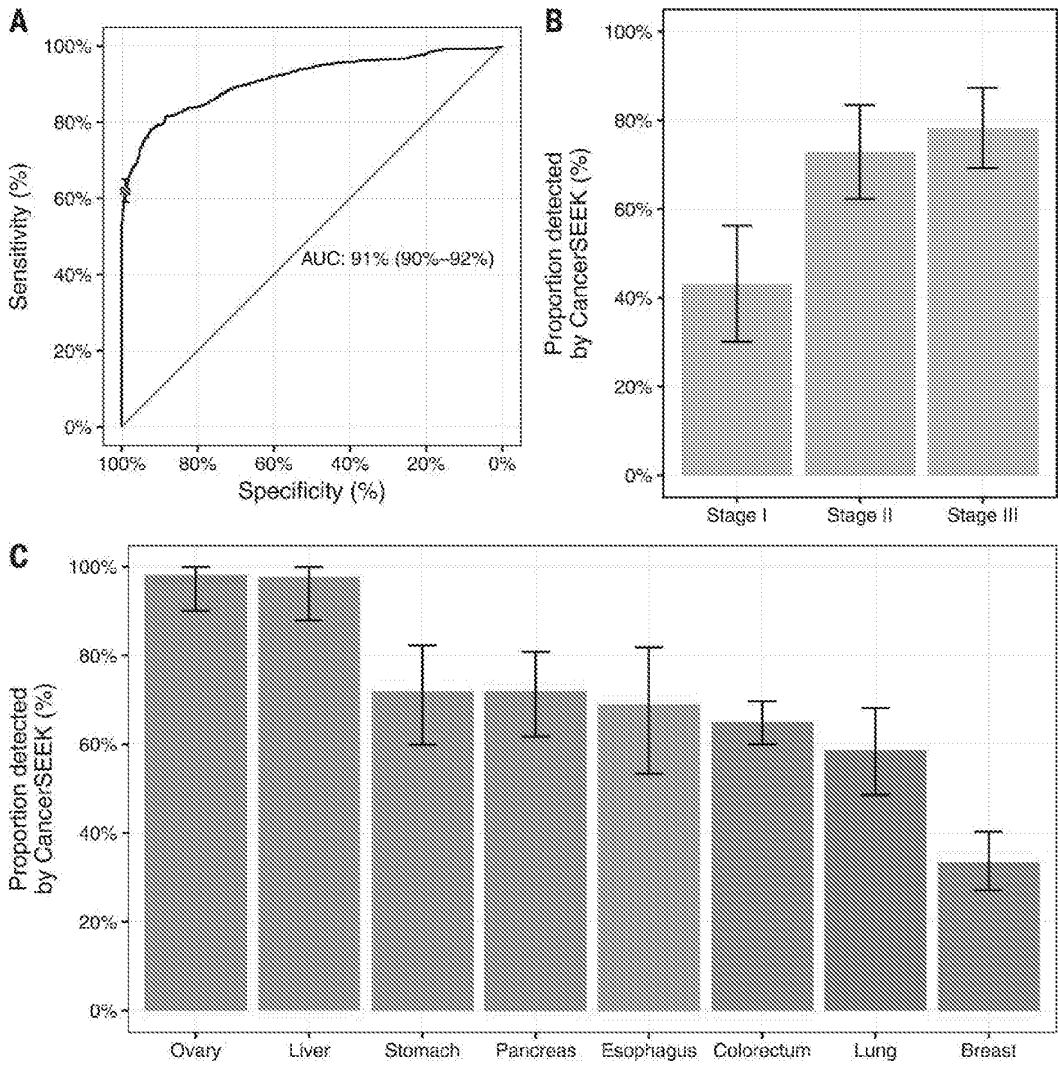


FIG. 4

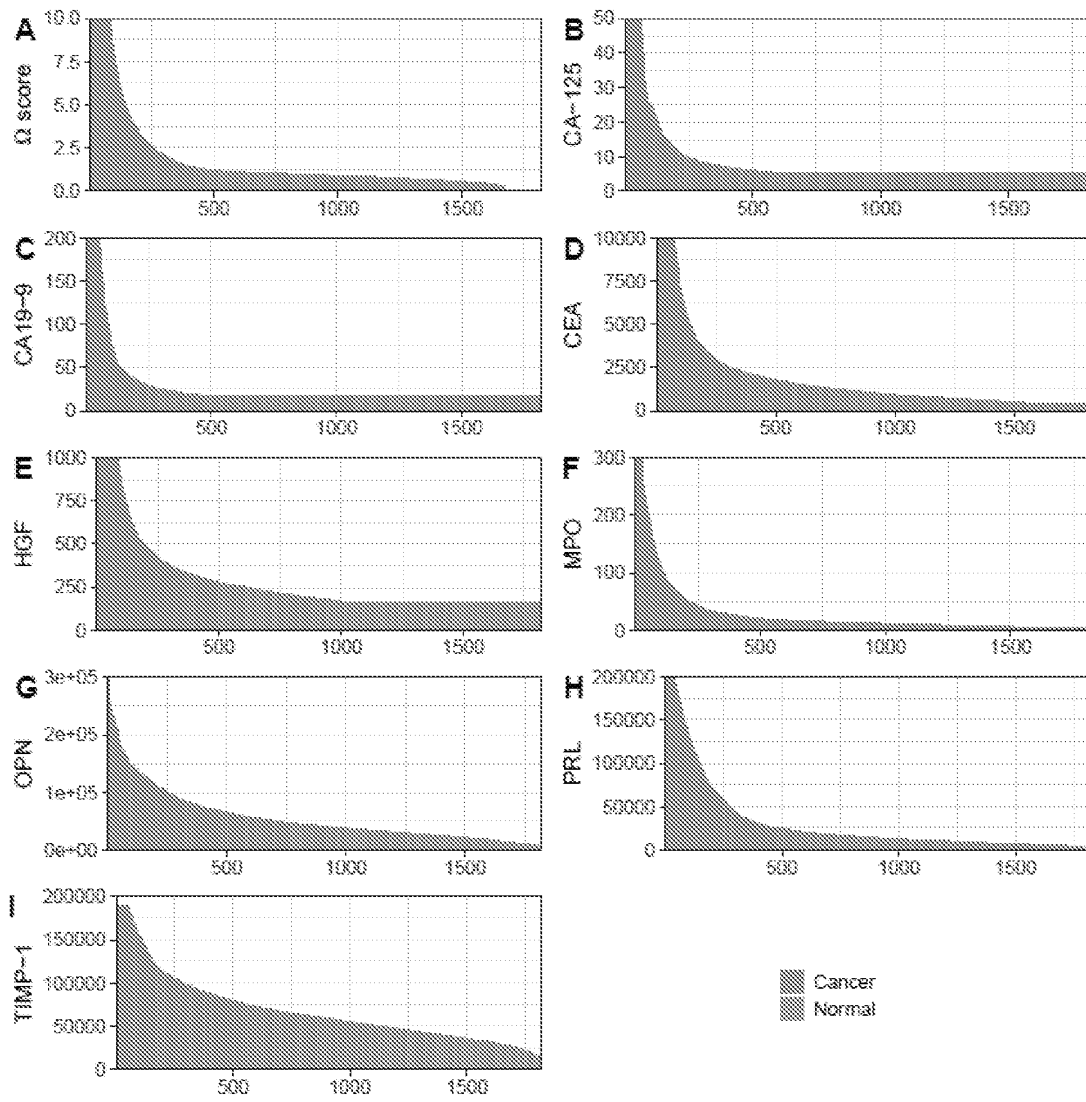


FIG. 5

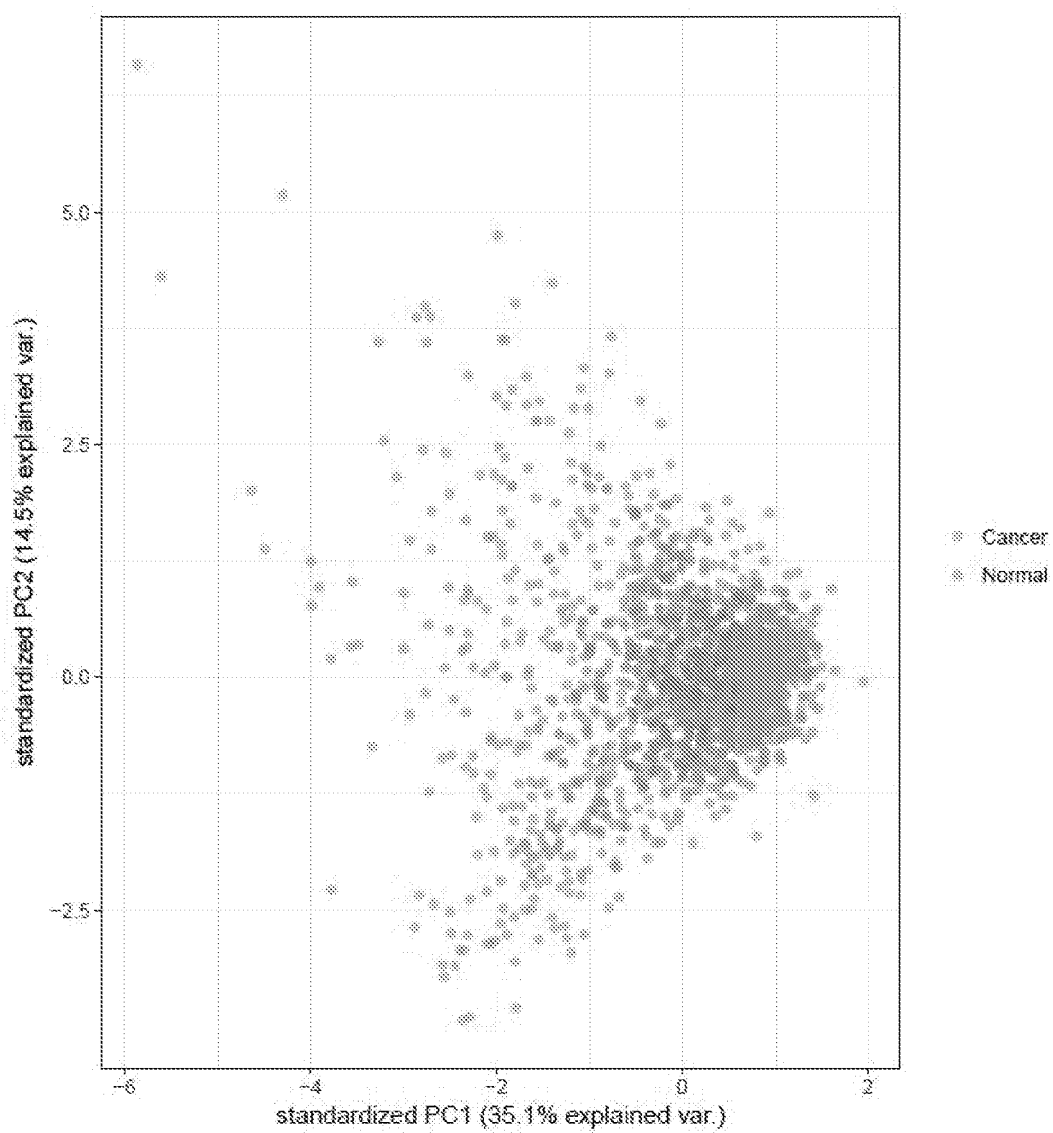


FIG. 6

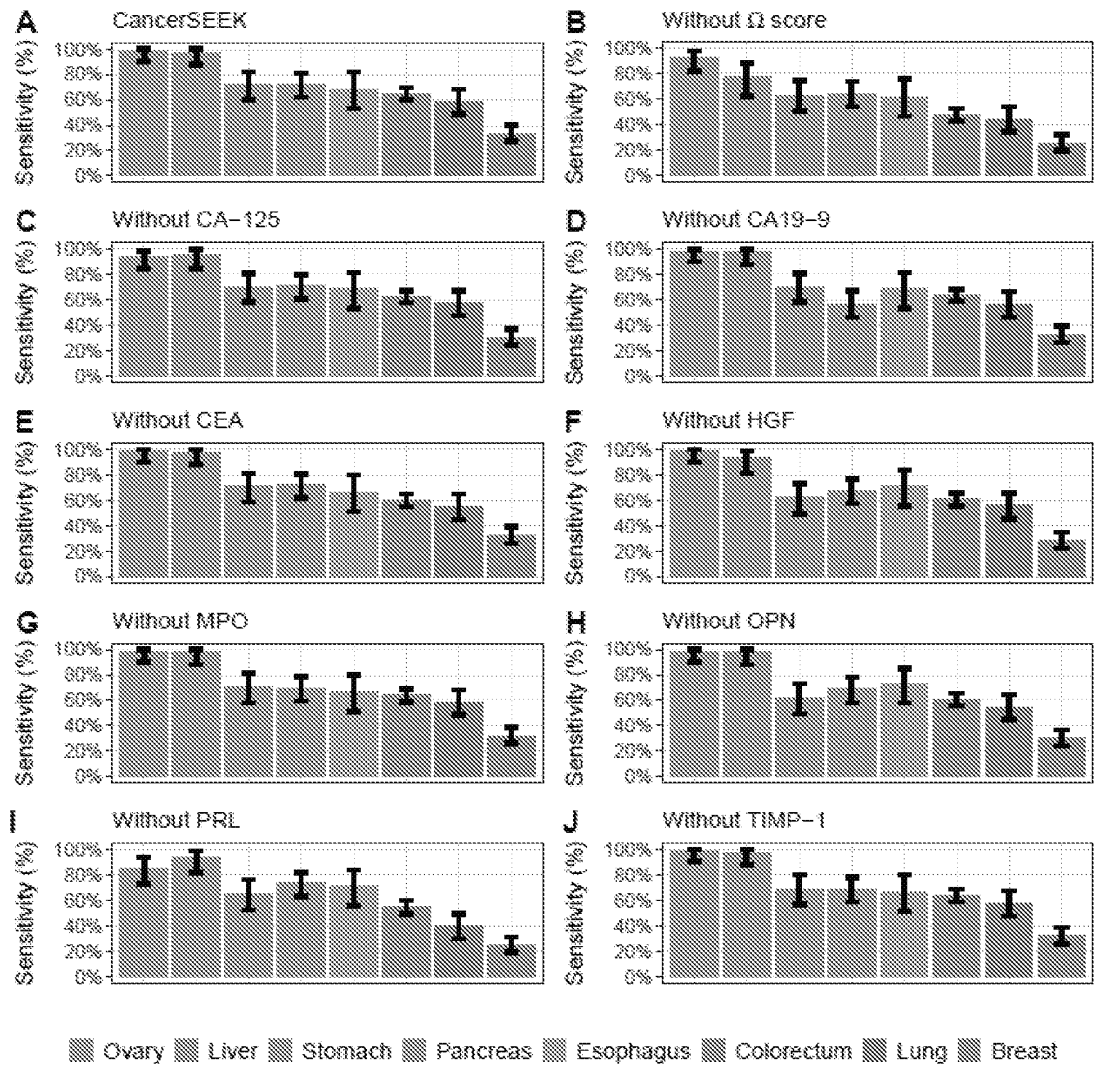


FIG. 7

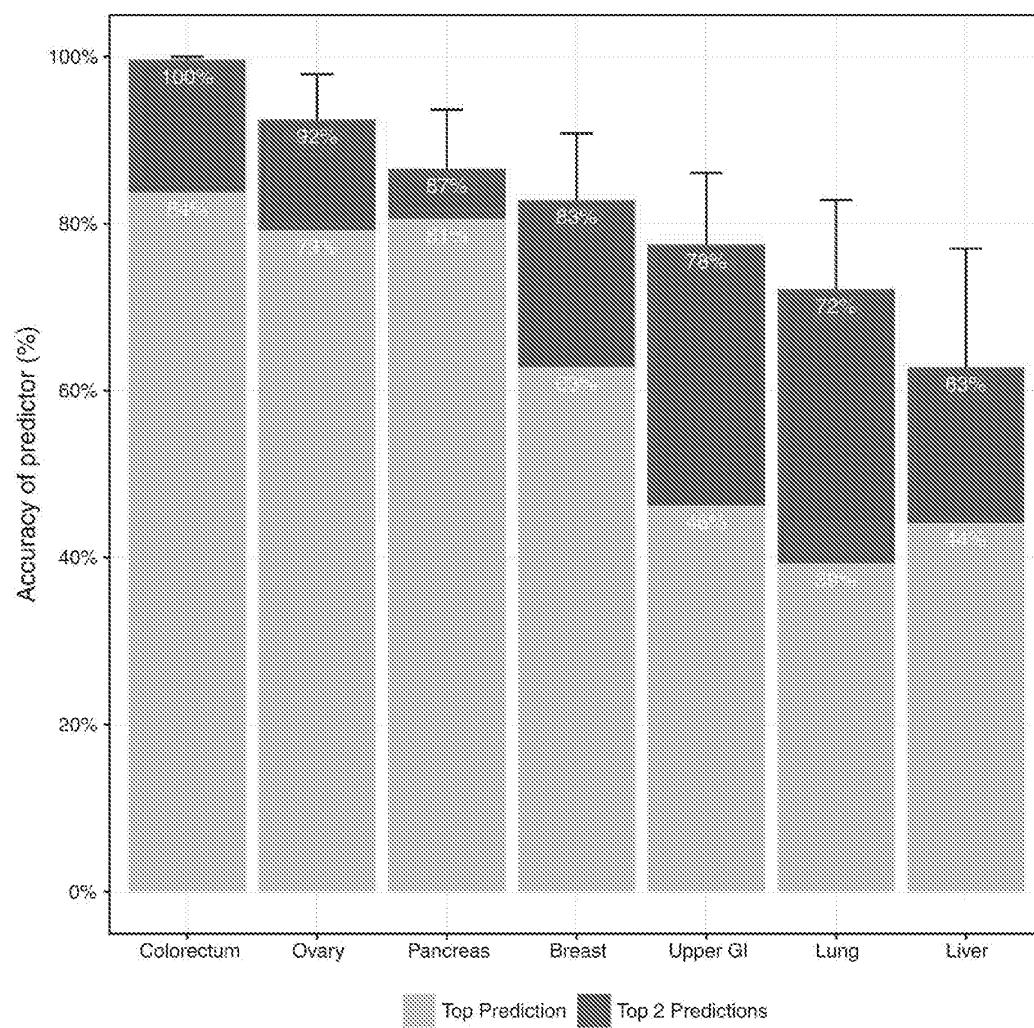


FIG. 8

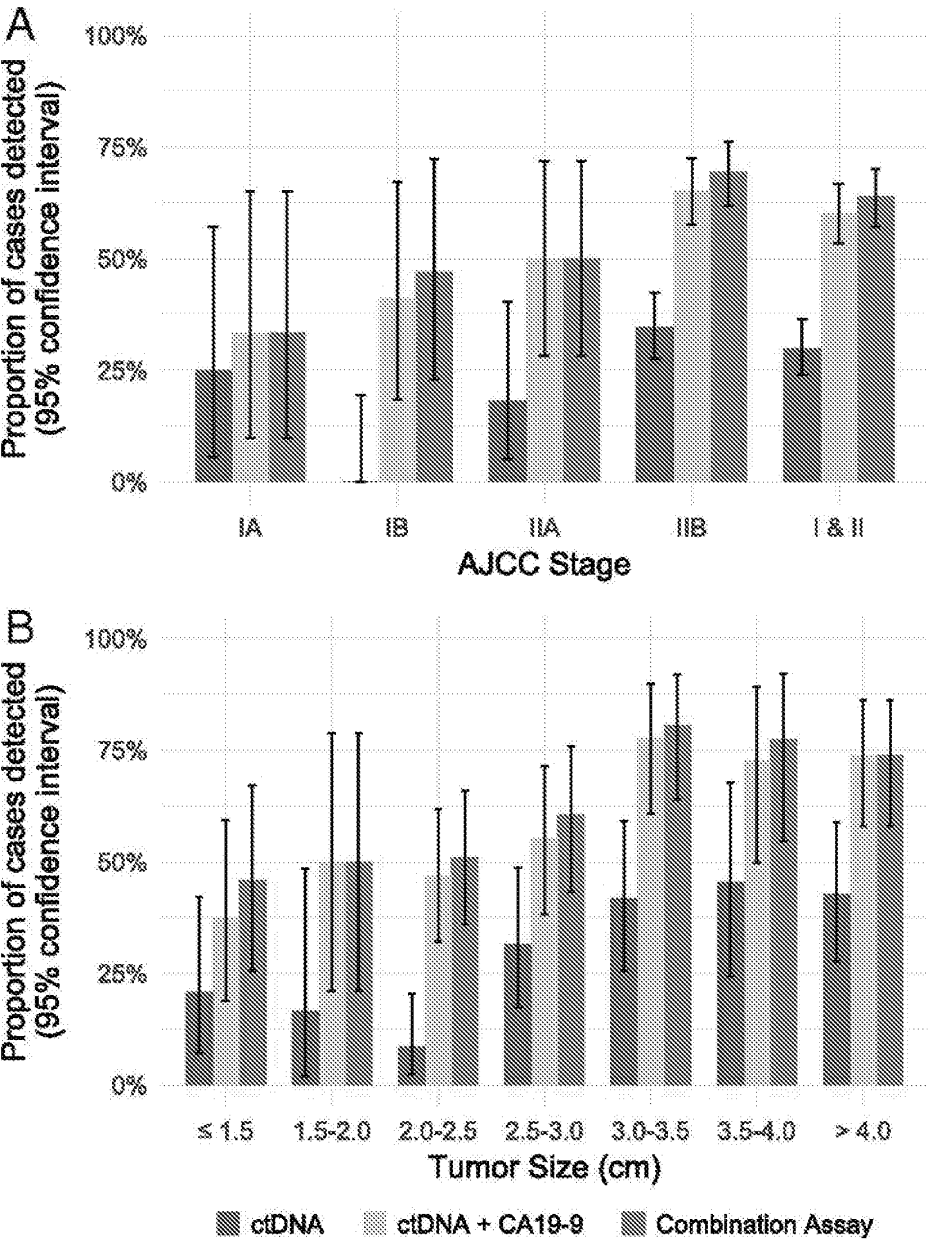


FIG. 9

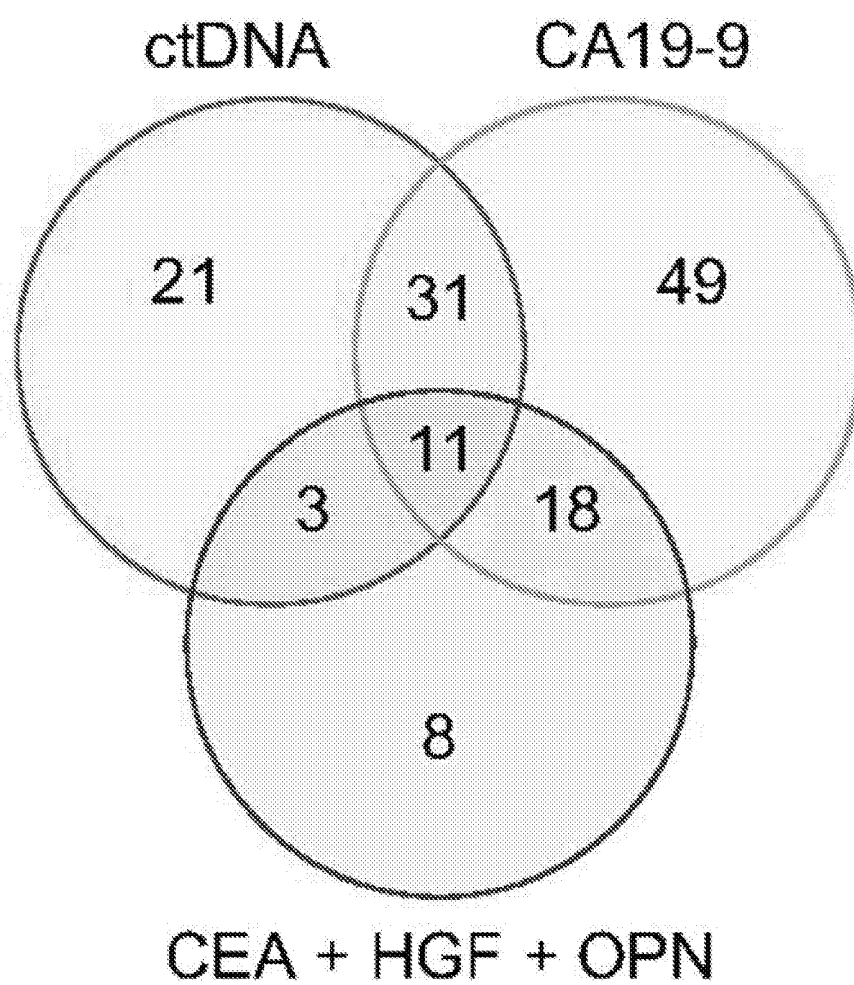


FIG. 10

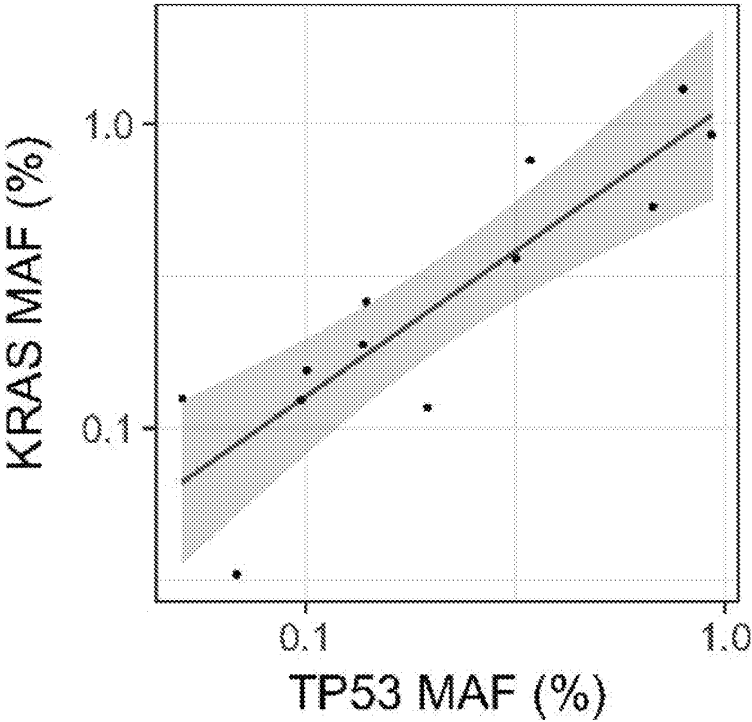


FIG. 11

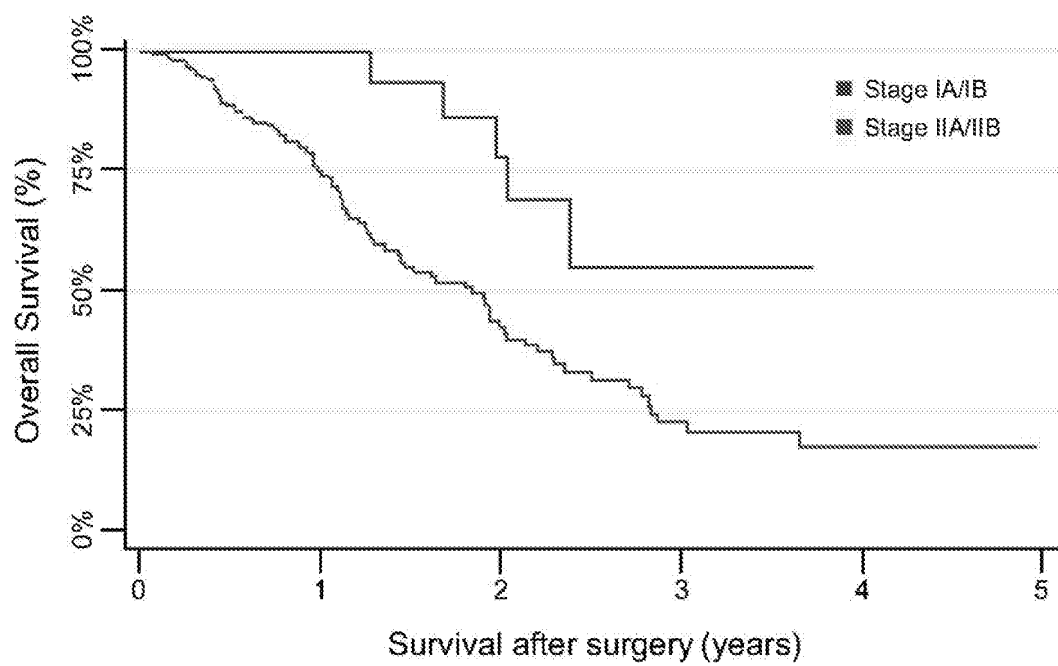


FIG. 12

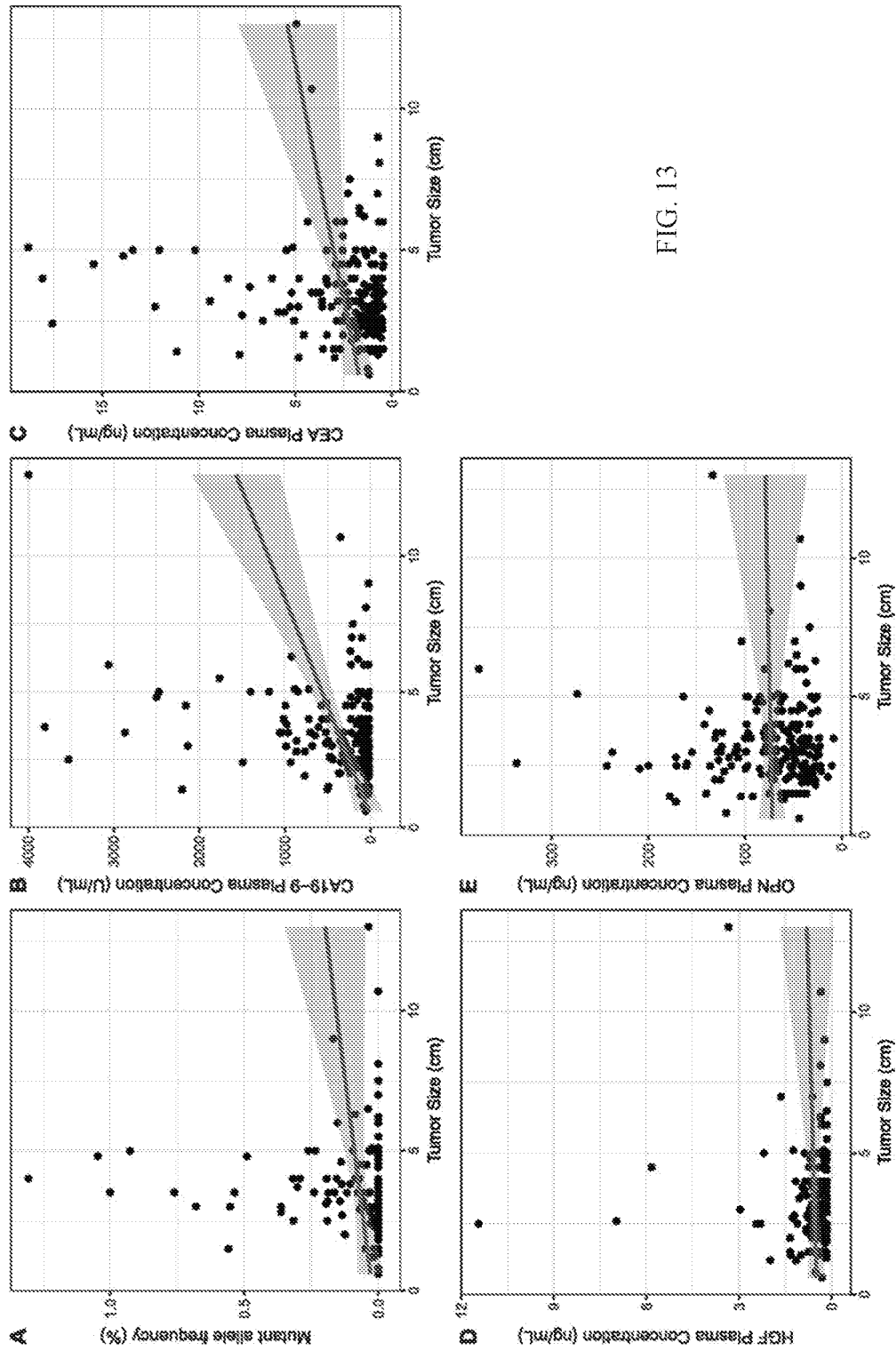


FIG. 13

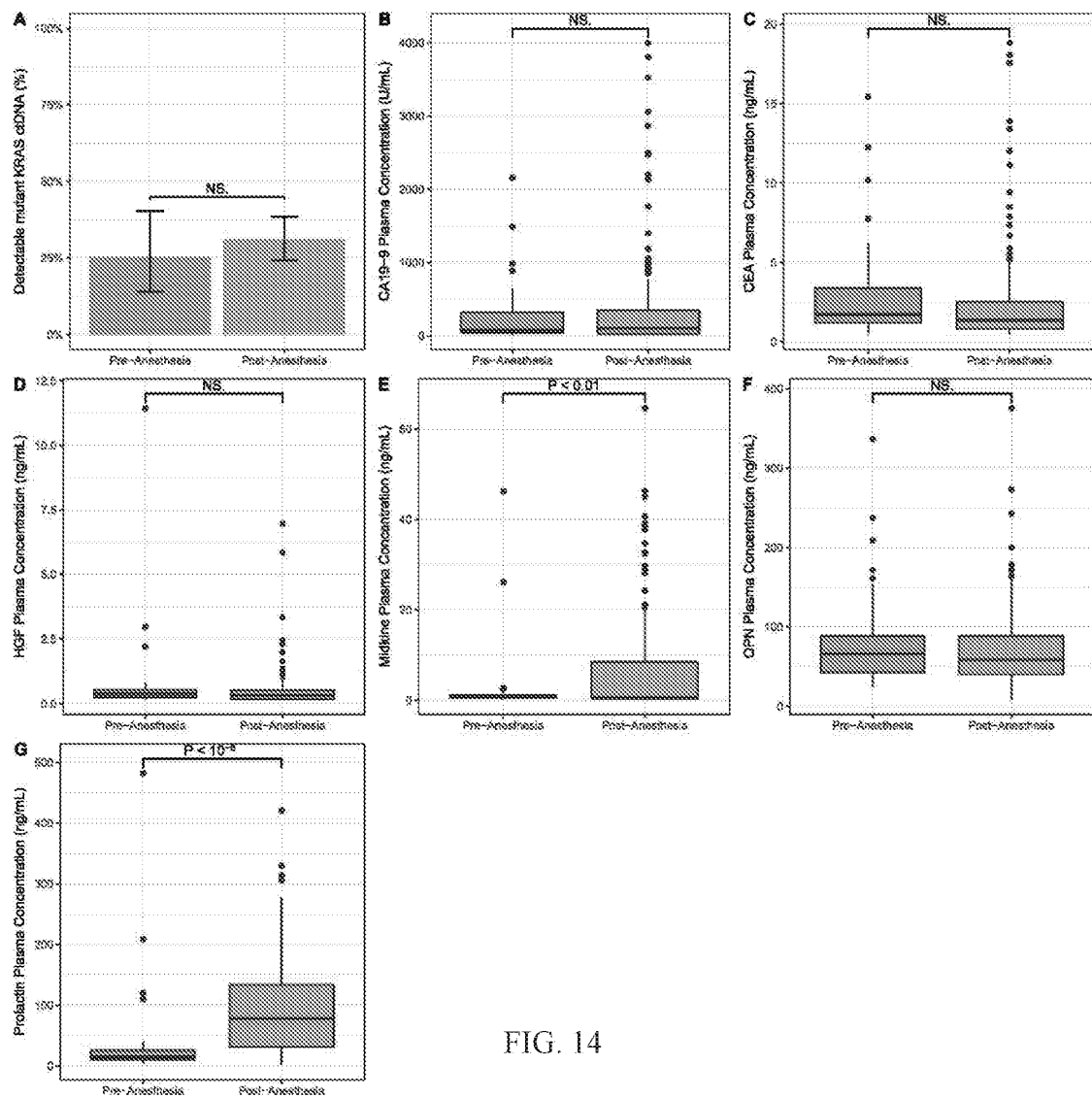


FIG. 14

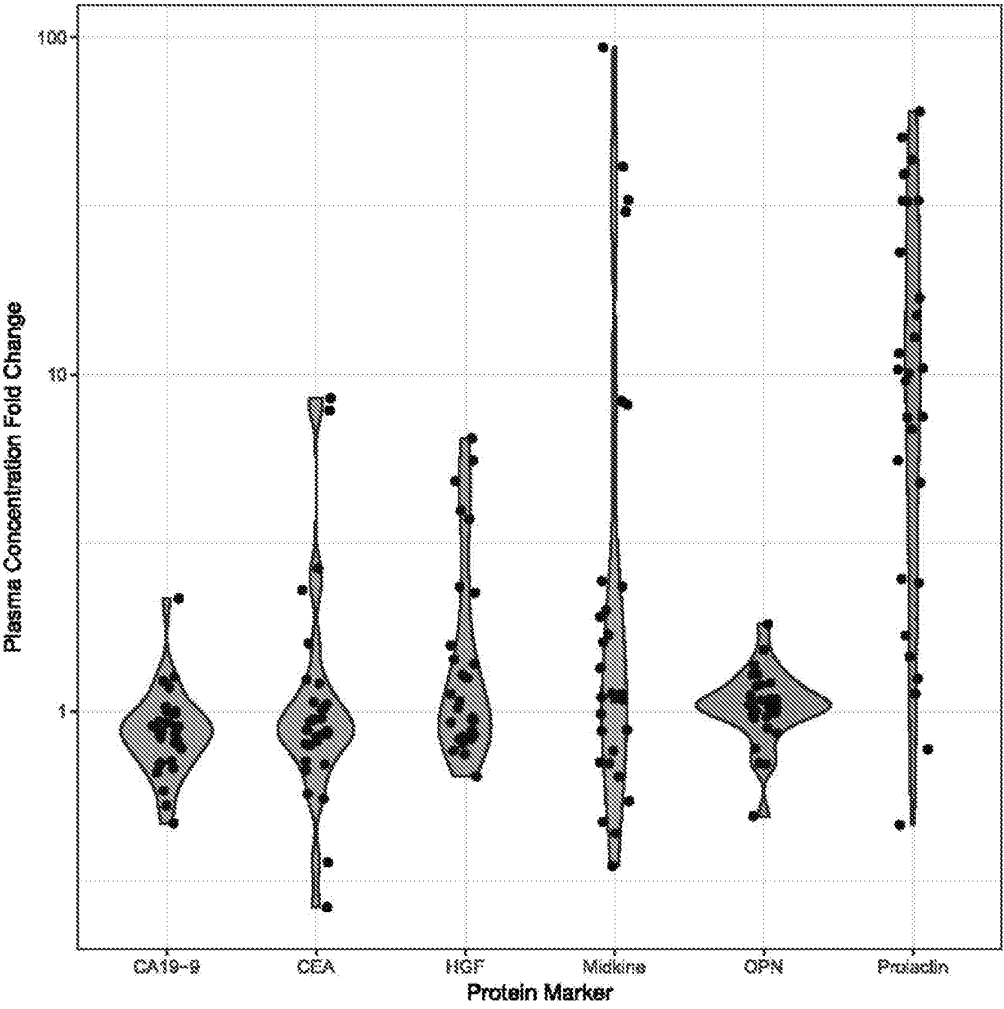


FIG. 15

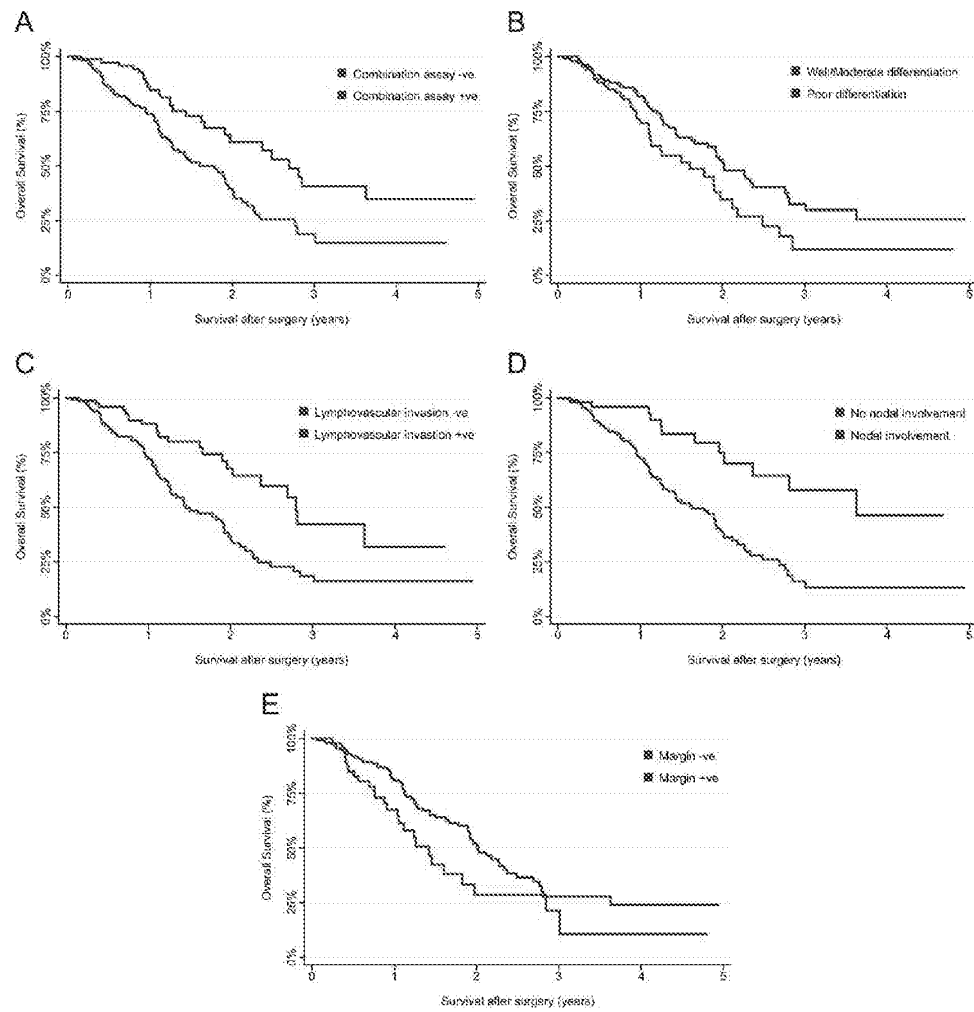


FIG. 16

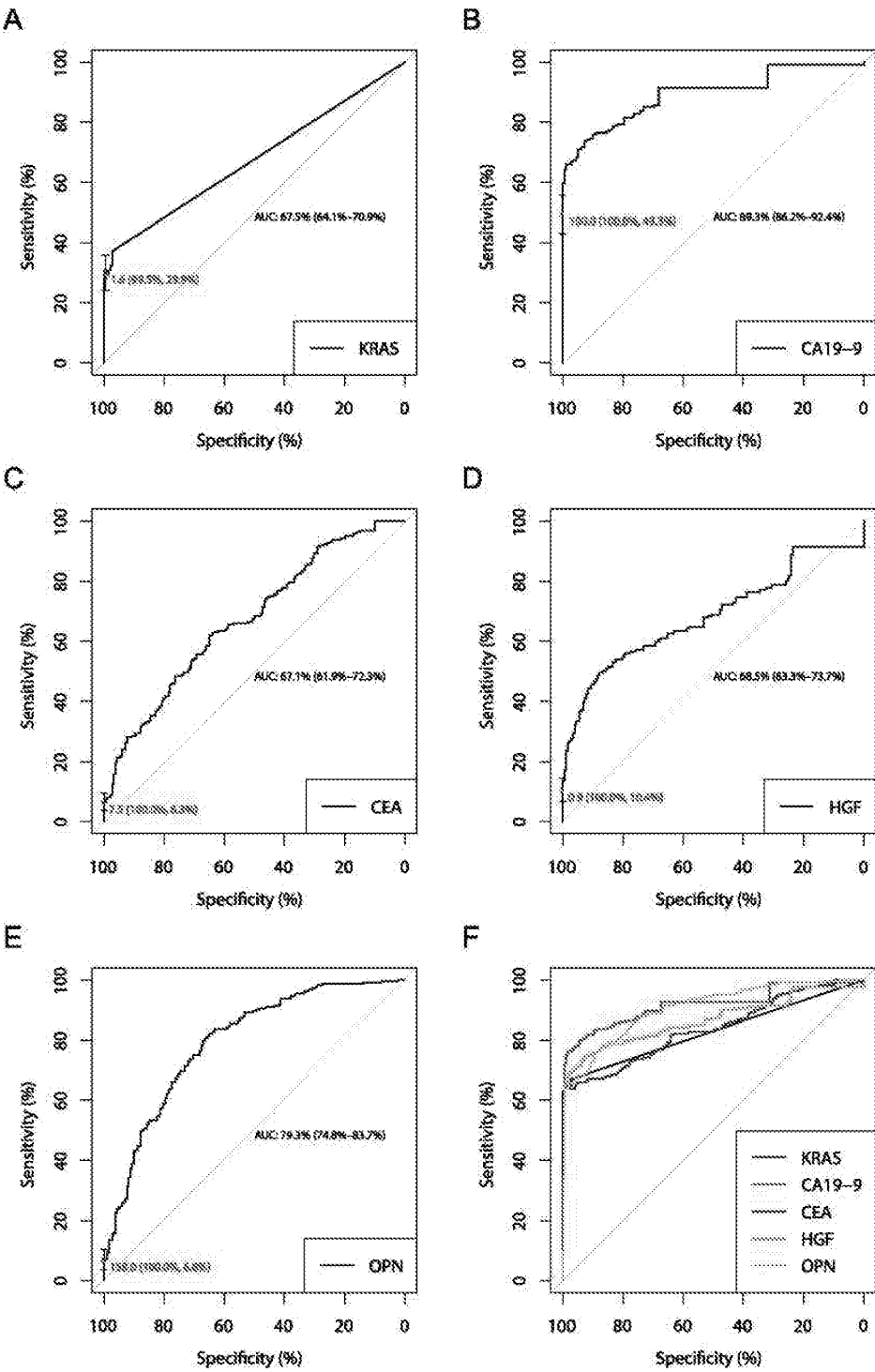


FIG. 17

A

Tumor Type	SEER Incidence	Total #	# ctDNA Positive	# Protein Positive	# ctDNA + Protein Positive	% ctDNA Positive	% Protein Positive	% ctDNA + Protein Positive
Breast	43.8	150	24	30	47	16%	20%	31%
CRC	10	322	130	168	231	40%	52%	72%
Esophagus	0.63	43	21	28	34	49%	65%	79%
Gastric	1.34	65	20	38	39	31%	58%	60%
Liver	1.5	53	39	43	50	74%	81%	94%
Lung	3.31	109	43	30	61	39%	28%	56%
Ovarian	3.83	86	43	45	59	50%	52%	69%
Pancreas	3.8	412	91	248	271	22%	60%	66%
All Cancers	68.21	1240	411	630	792	33%	51%	64%
All Cancers (Incidence Weighted)	68.21	12239	3102.55	4341.11	5574.96	25%	35%	49%

B

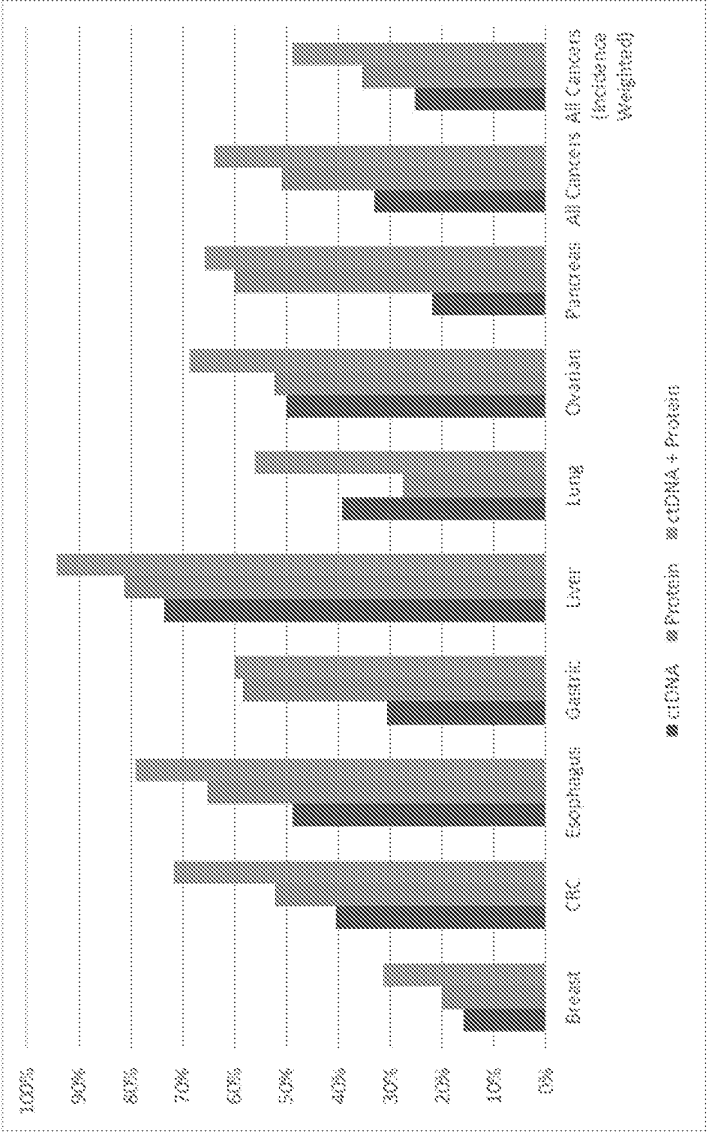


FIG. 18

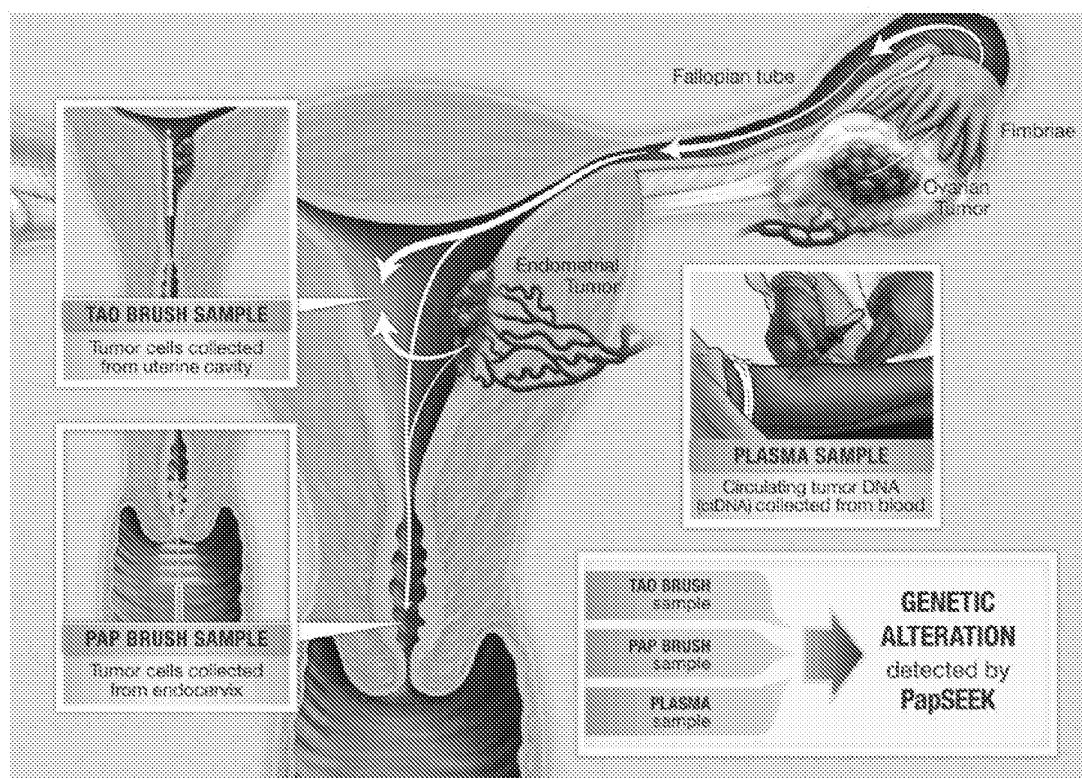


FIG. 19

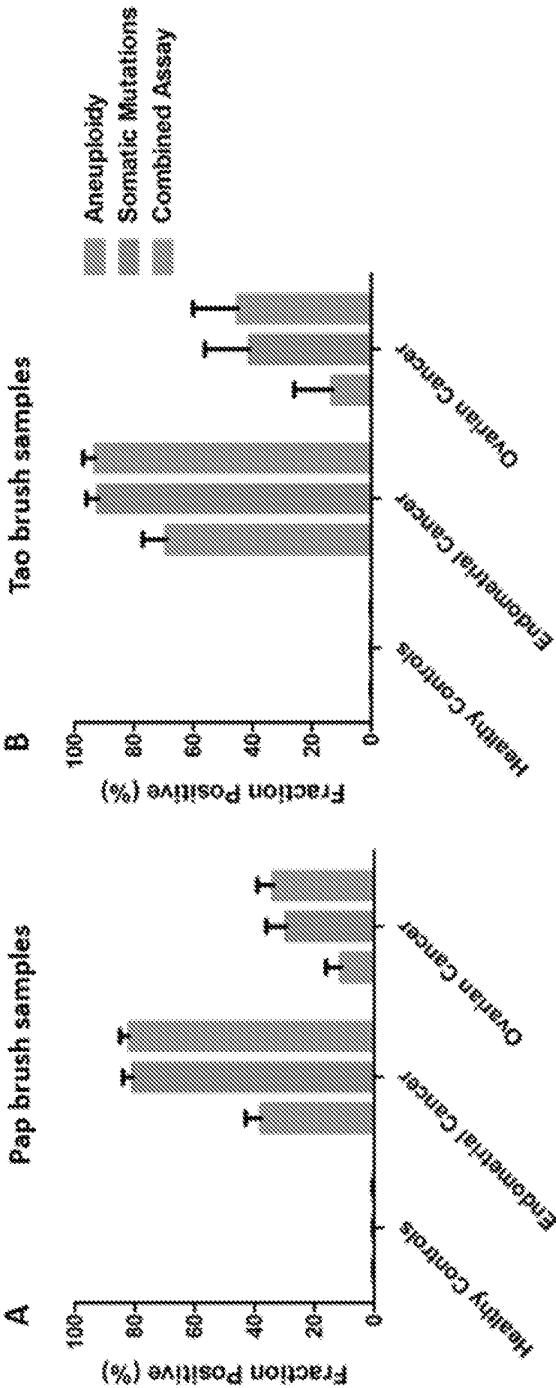


FIG. 20

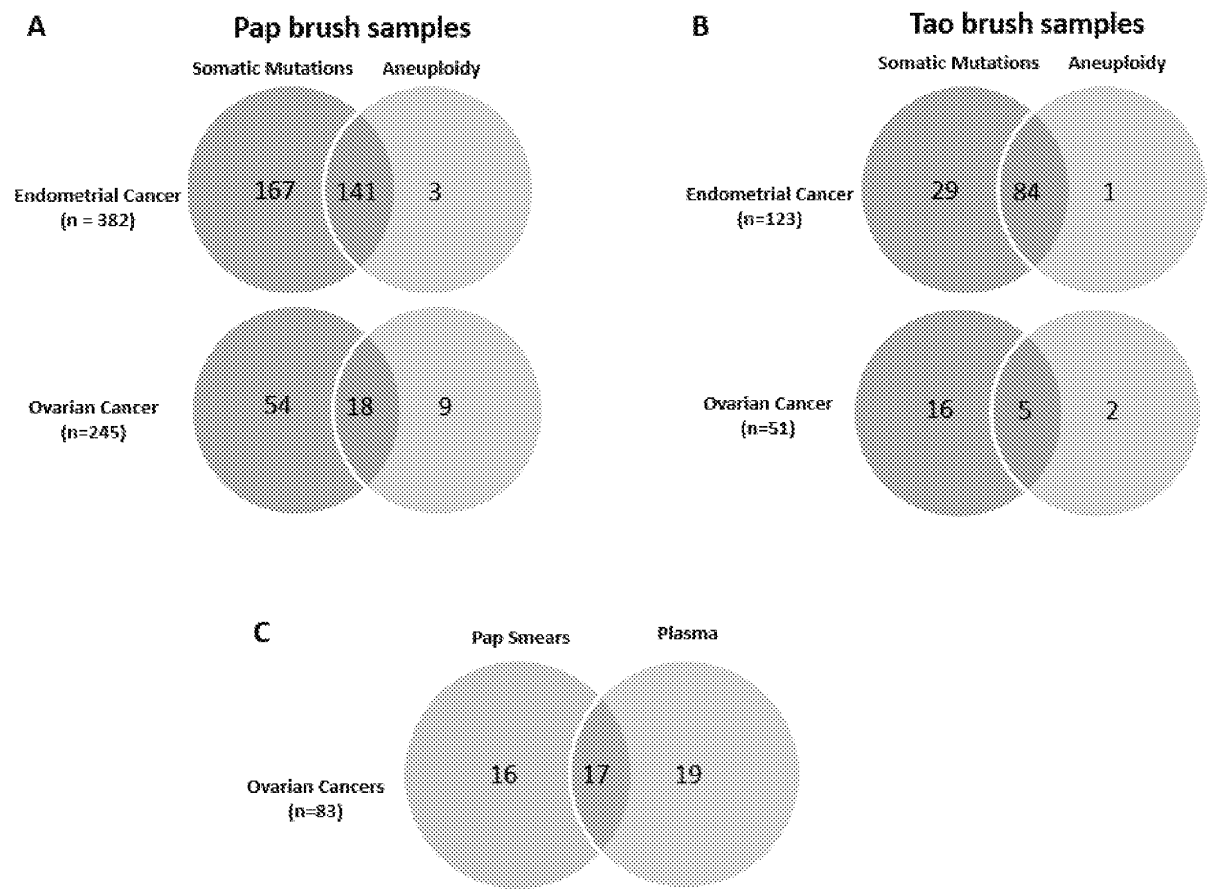


FIG. 21

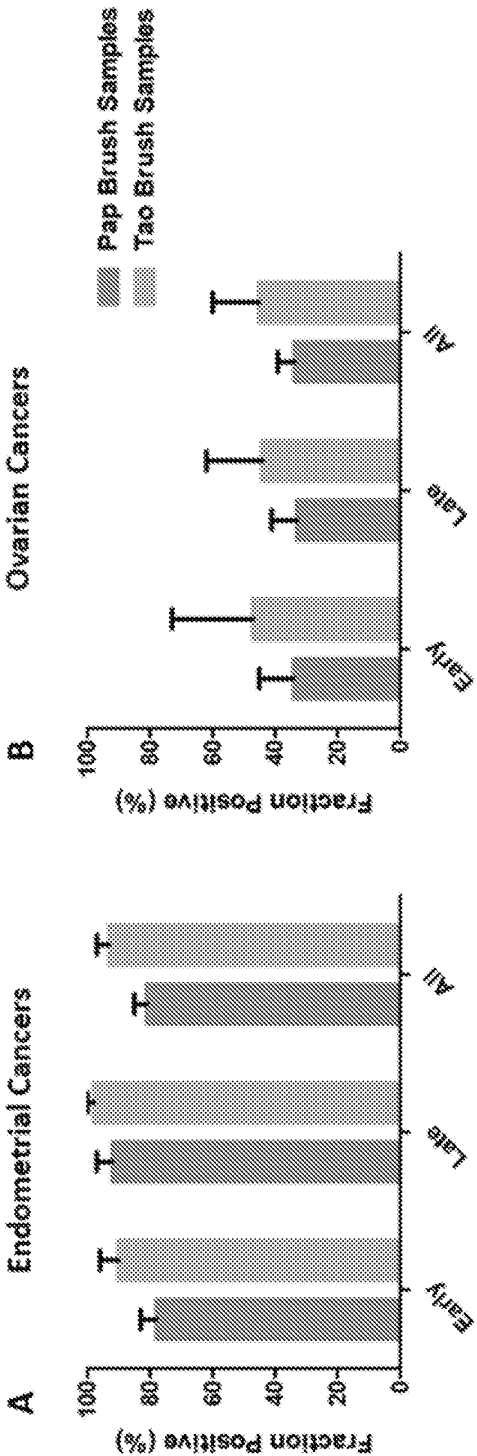


FIG. 22

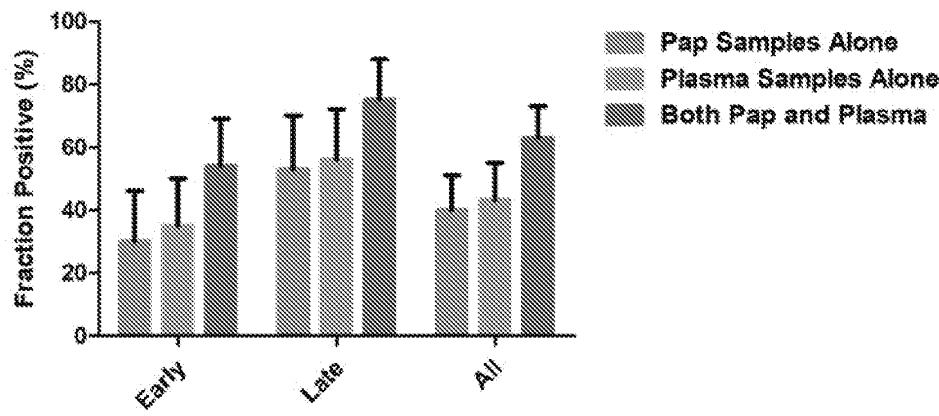


FIG. 23

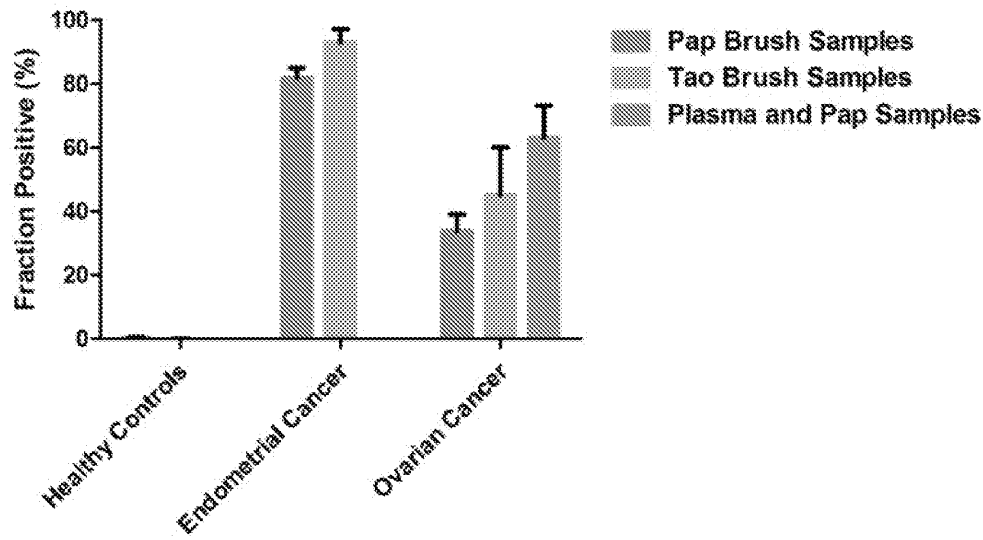


FIG. 24

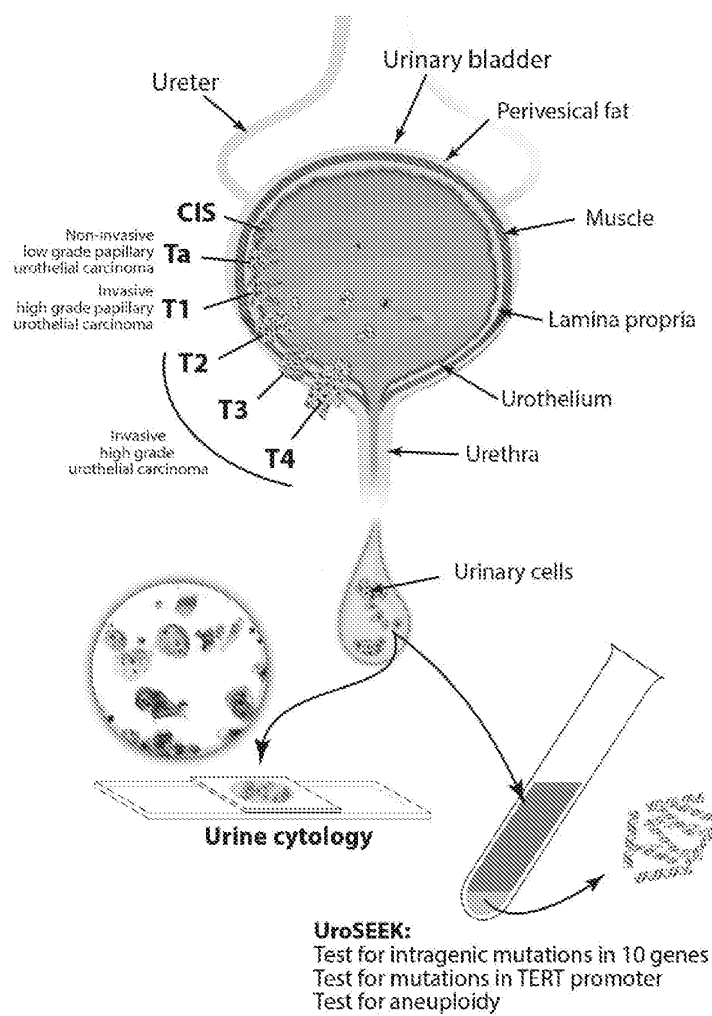


FIG. 25

Early Detection Cohort				Surveillance Cohort			
Urine collected at patient presentation (n= 670)				Urine collected following surgery (n= 322)			
Developed Bladder Cancer 31%	Did Not Develop Bladder Cancer 69%			Developed Bladder Cancer 58%	Did not Develop Bladder Cancer 42%		
UroSEEK positive 83%	Cytology positive 43%	UroSEEK positive 7%	Cytology positive 0%	UroSEEK positive 68%	Cytology positive 25%	UroSEEK positive 20%	Cytology positive 0%
Sensitivity (UroSEEK/Cytology Combined) 95%		Specificity (UroSEEK/Cytology Combined) 93%		Sensitivity (UroSEEK/Cytology Combined) 71%		Specificity (UroSEEK/Cytology Combined) 82%	

FIG. 26

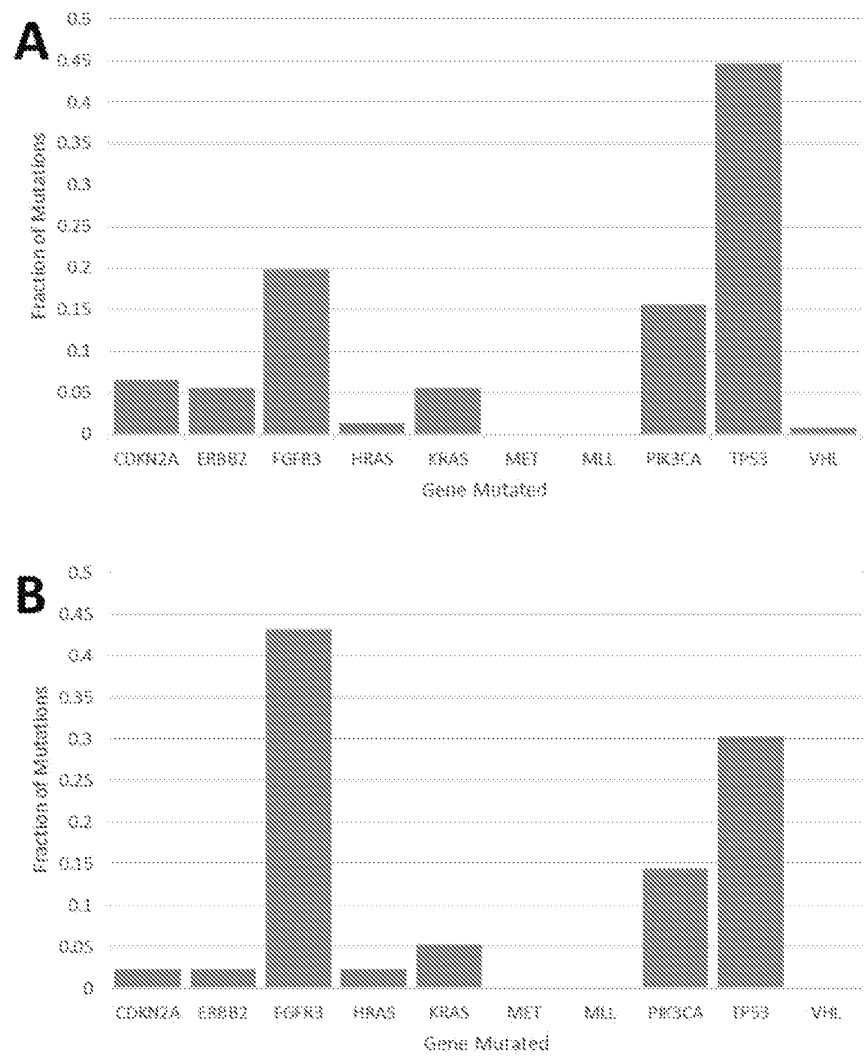


FIG. 27

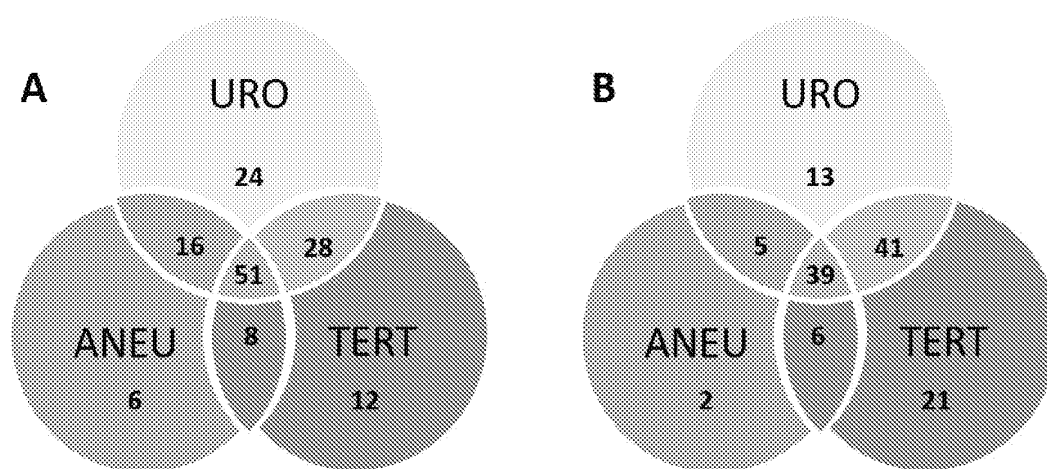


FIG. 28

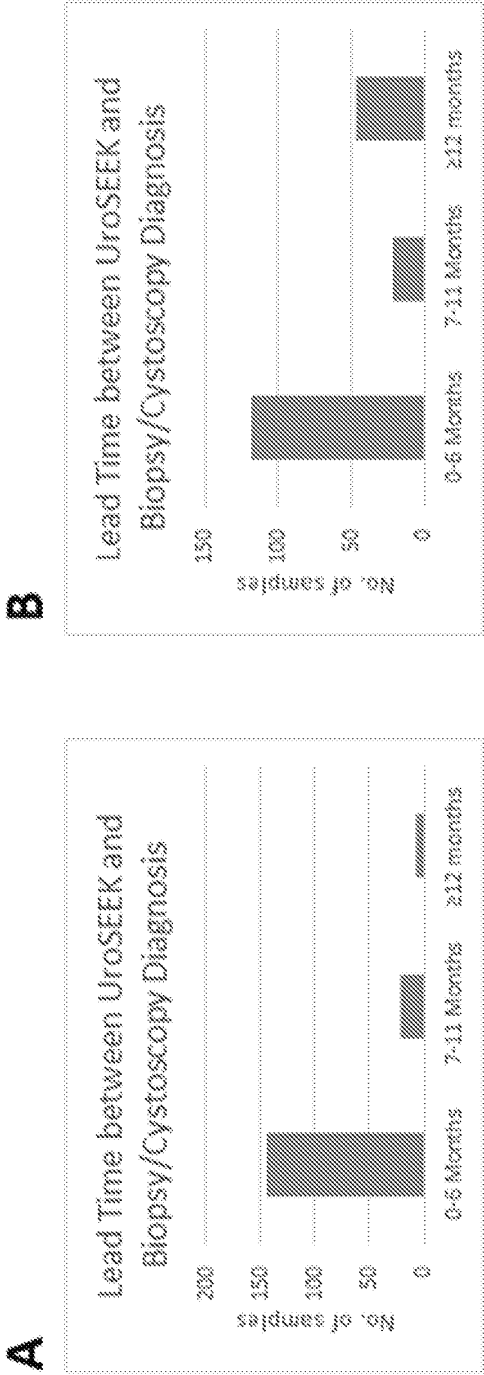


FIG. 29

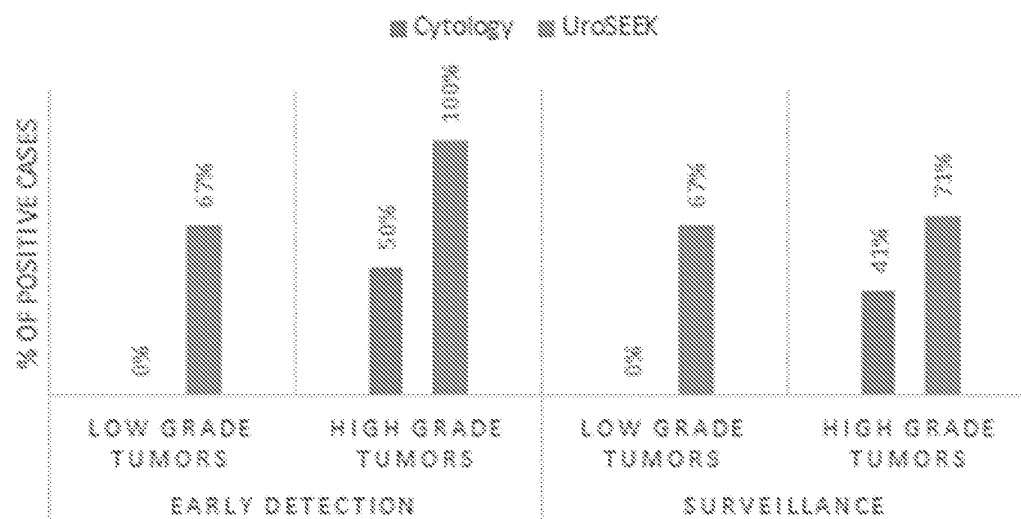


FIG. 30

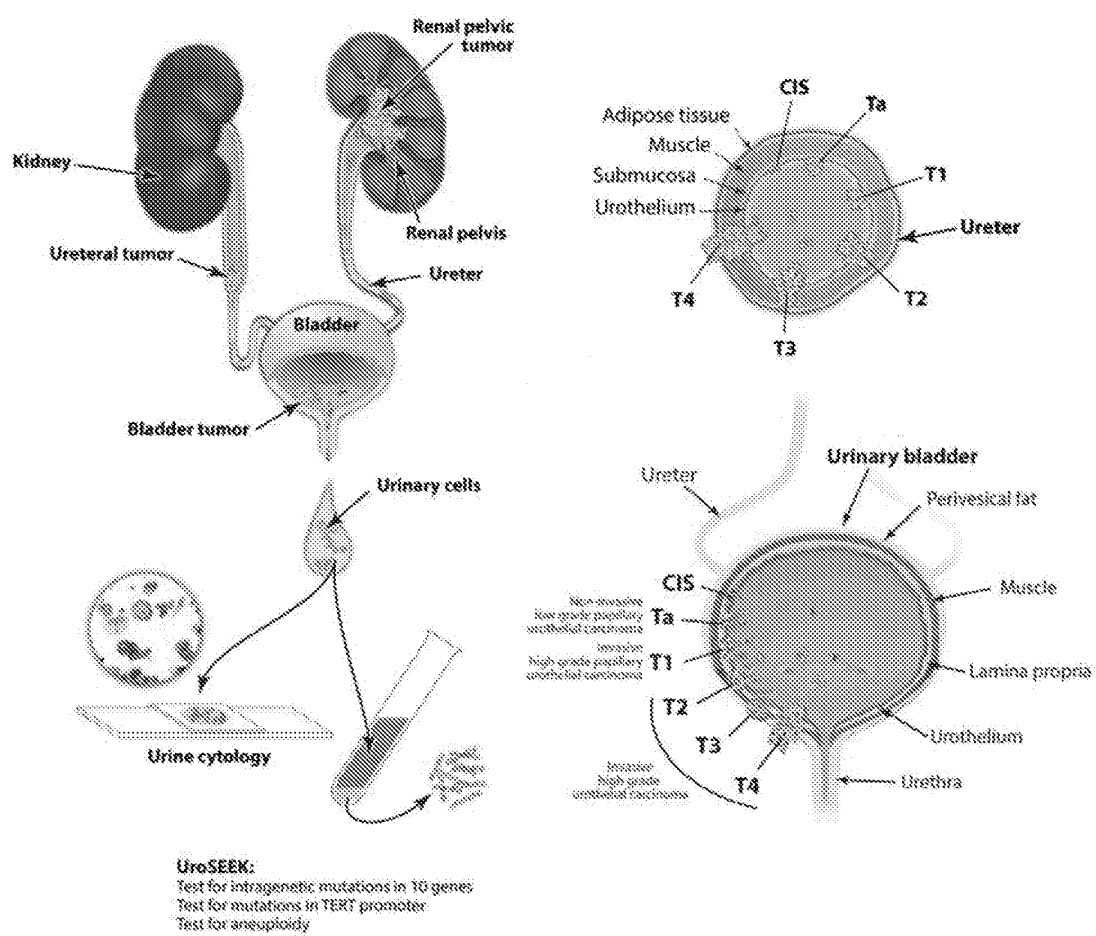


FIG. 31

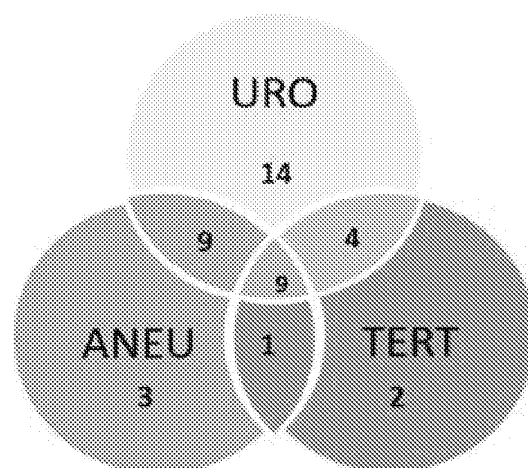


FIG. 32

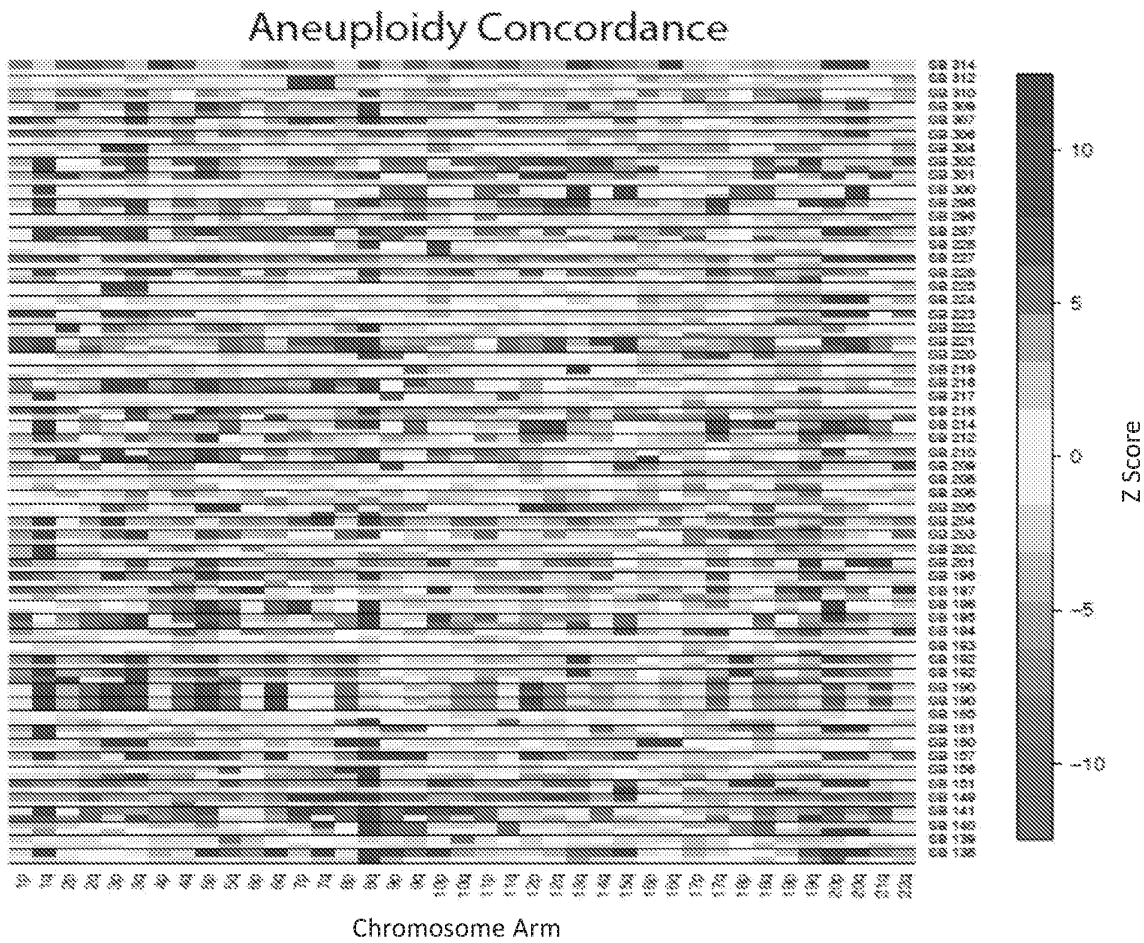


FIG. 33

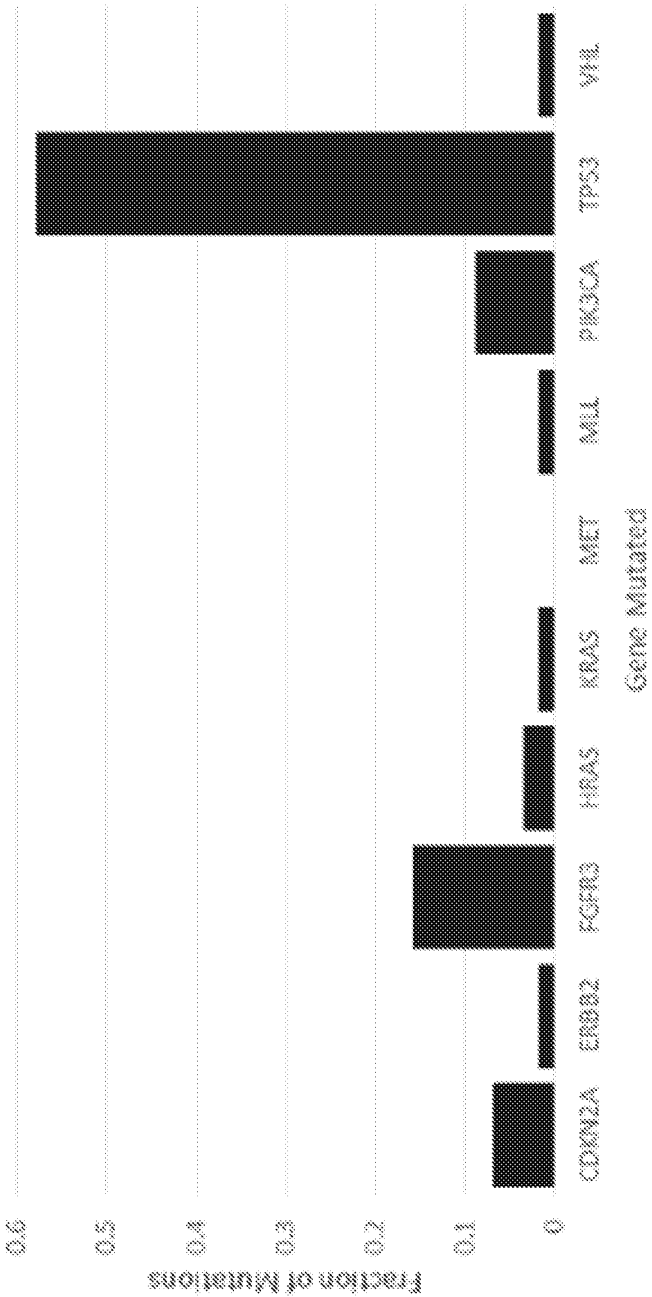


FIG. 34

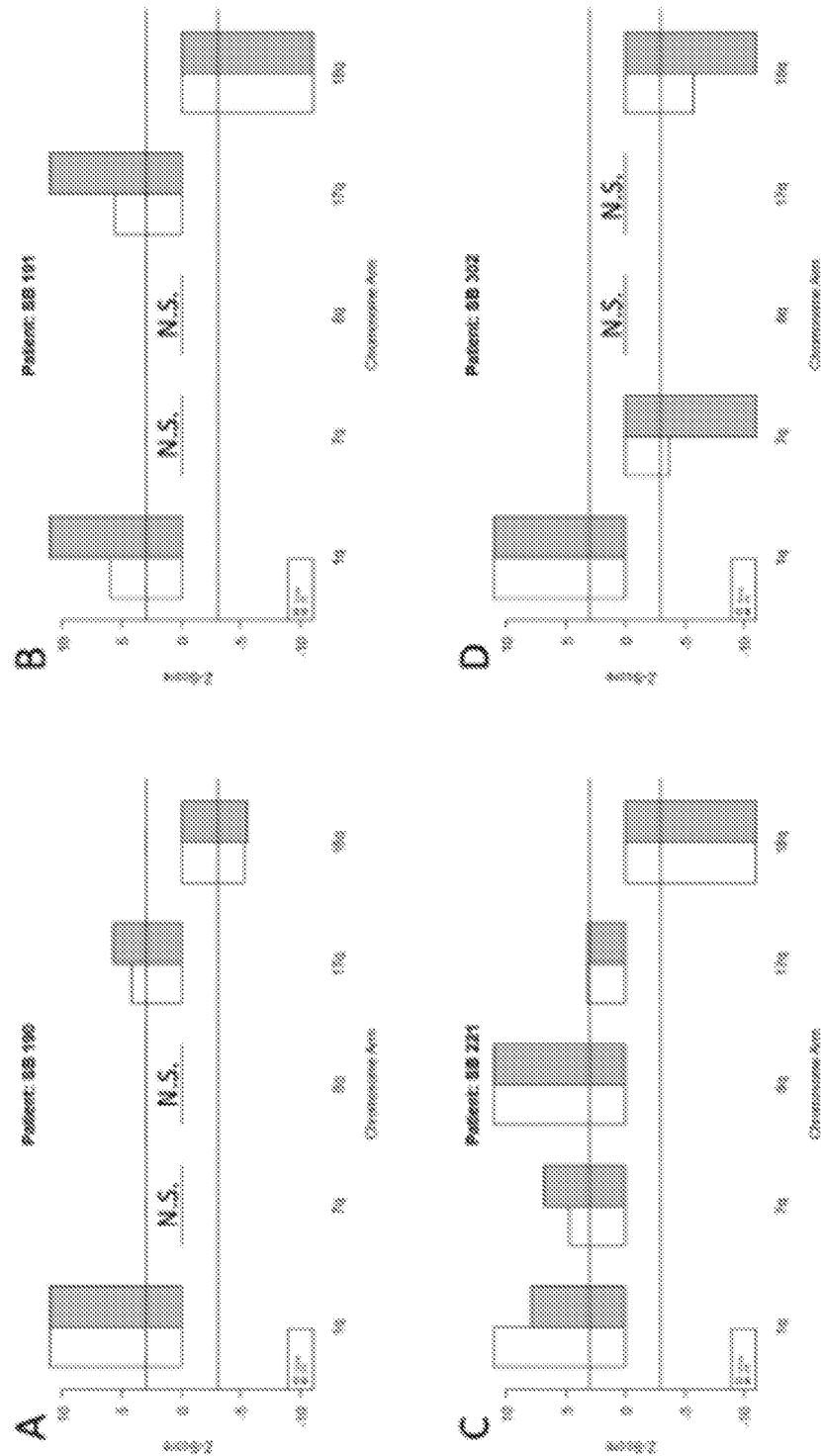


FIG. 35

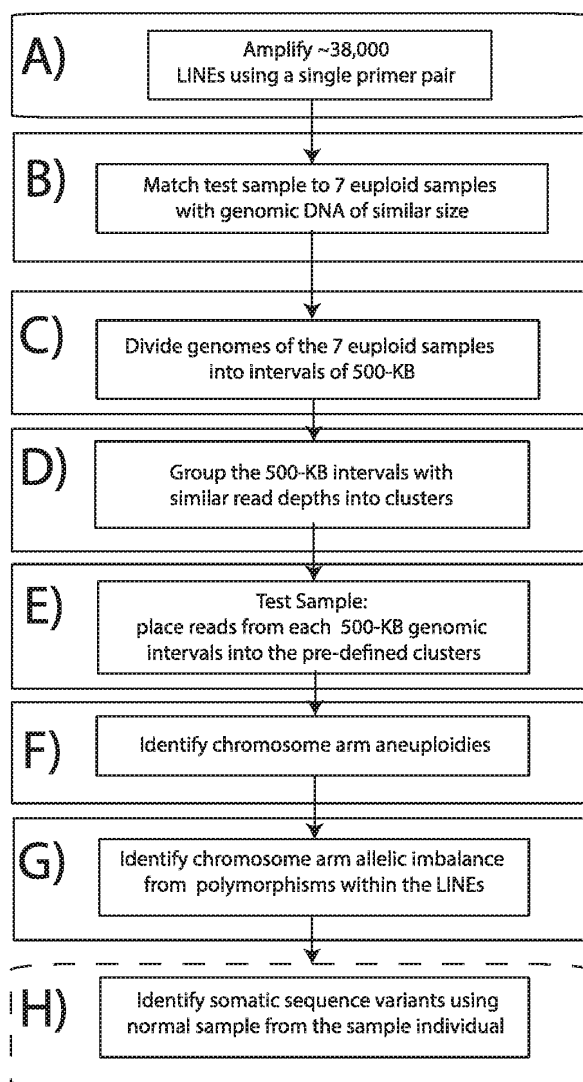


FIG. 36

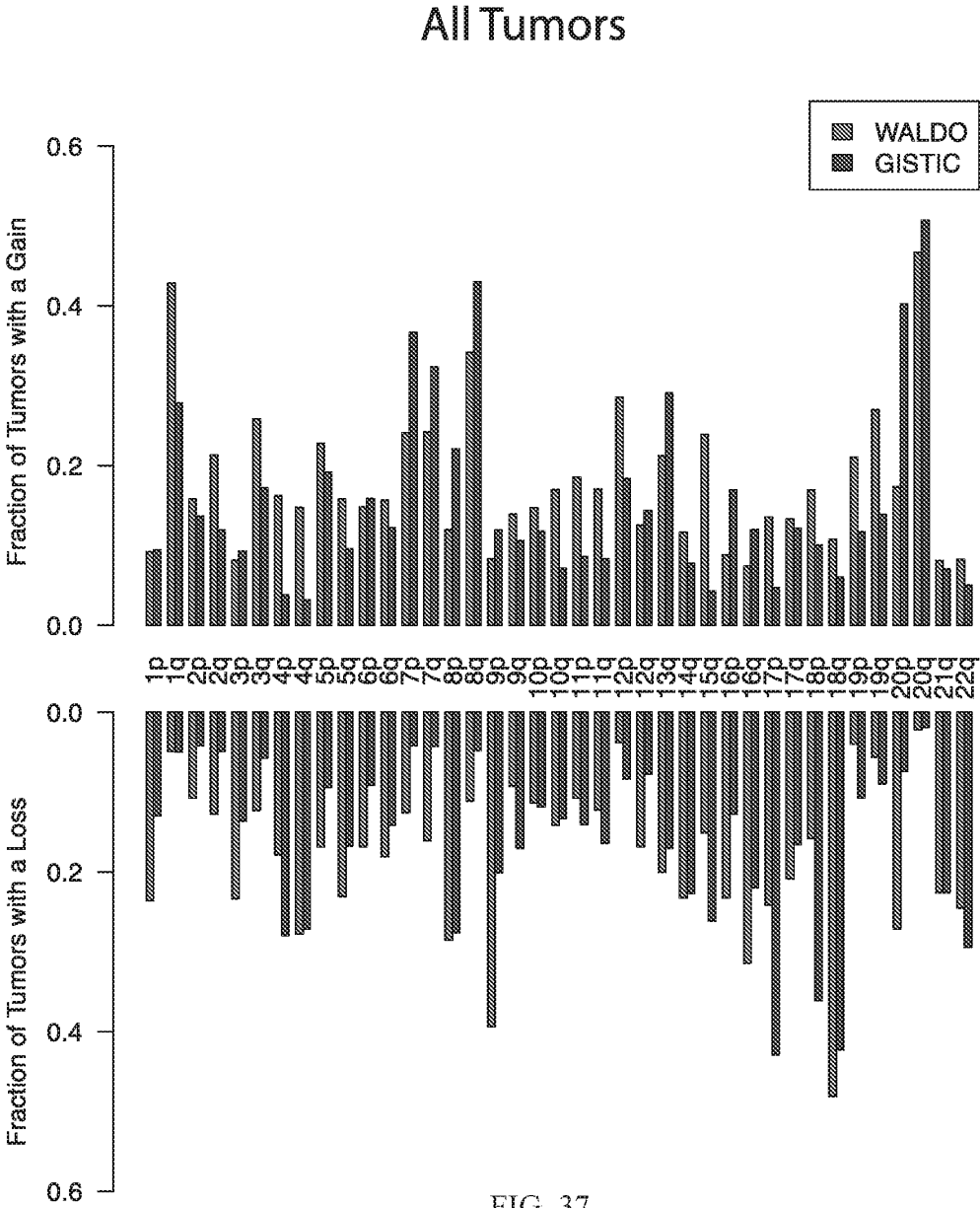


FIG. 37

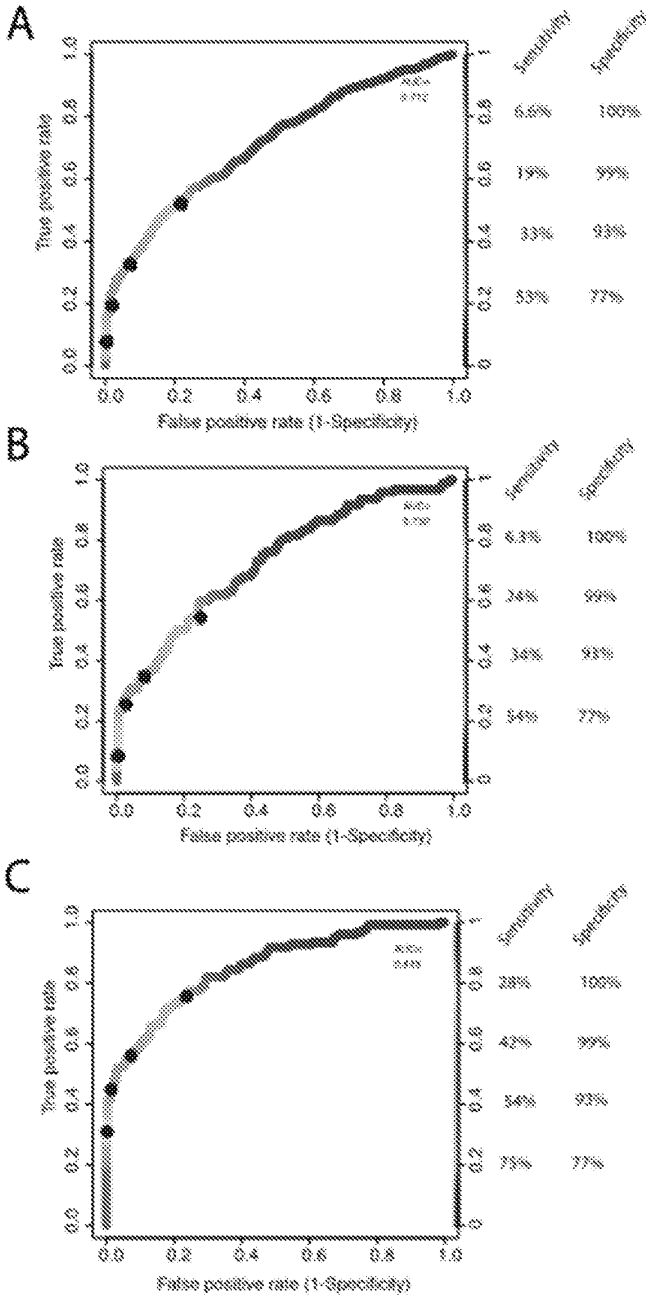


FIG. 38

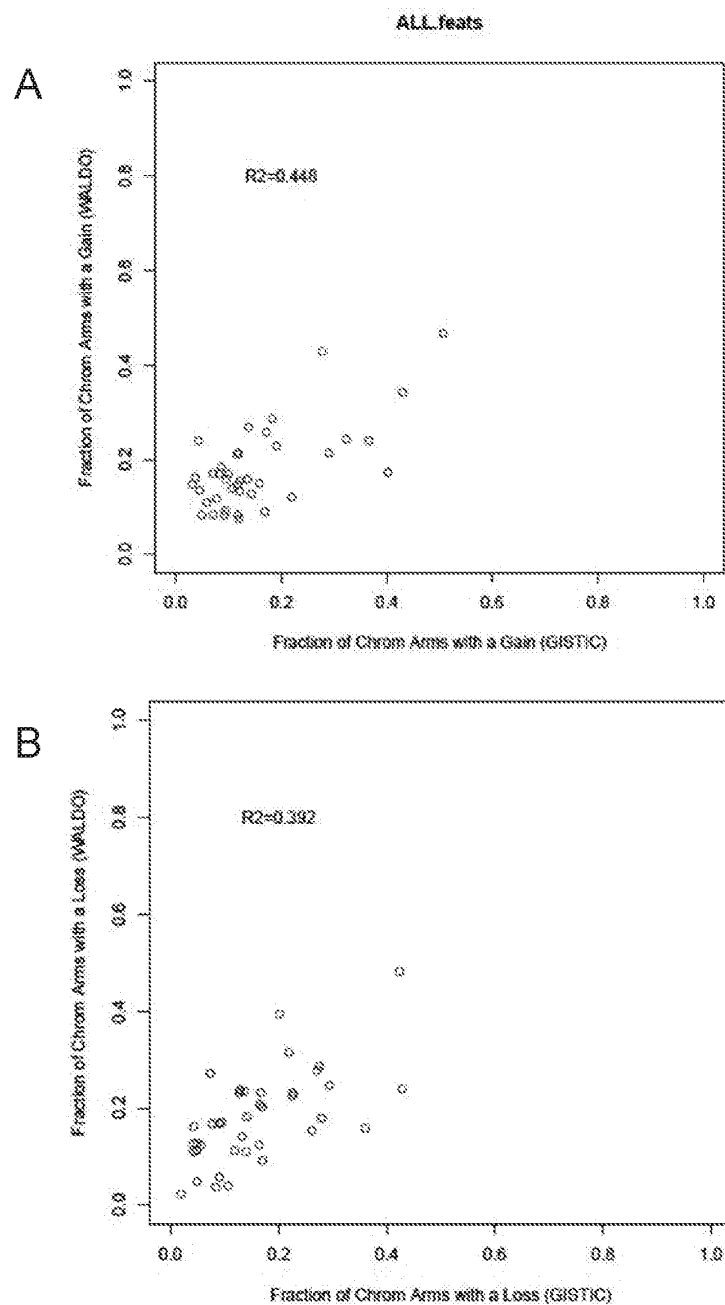


FIG. 39

A

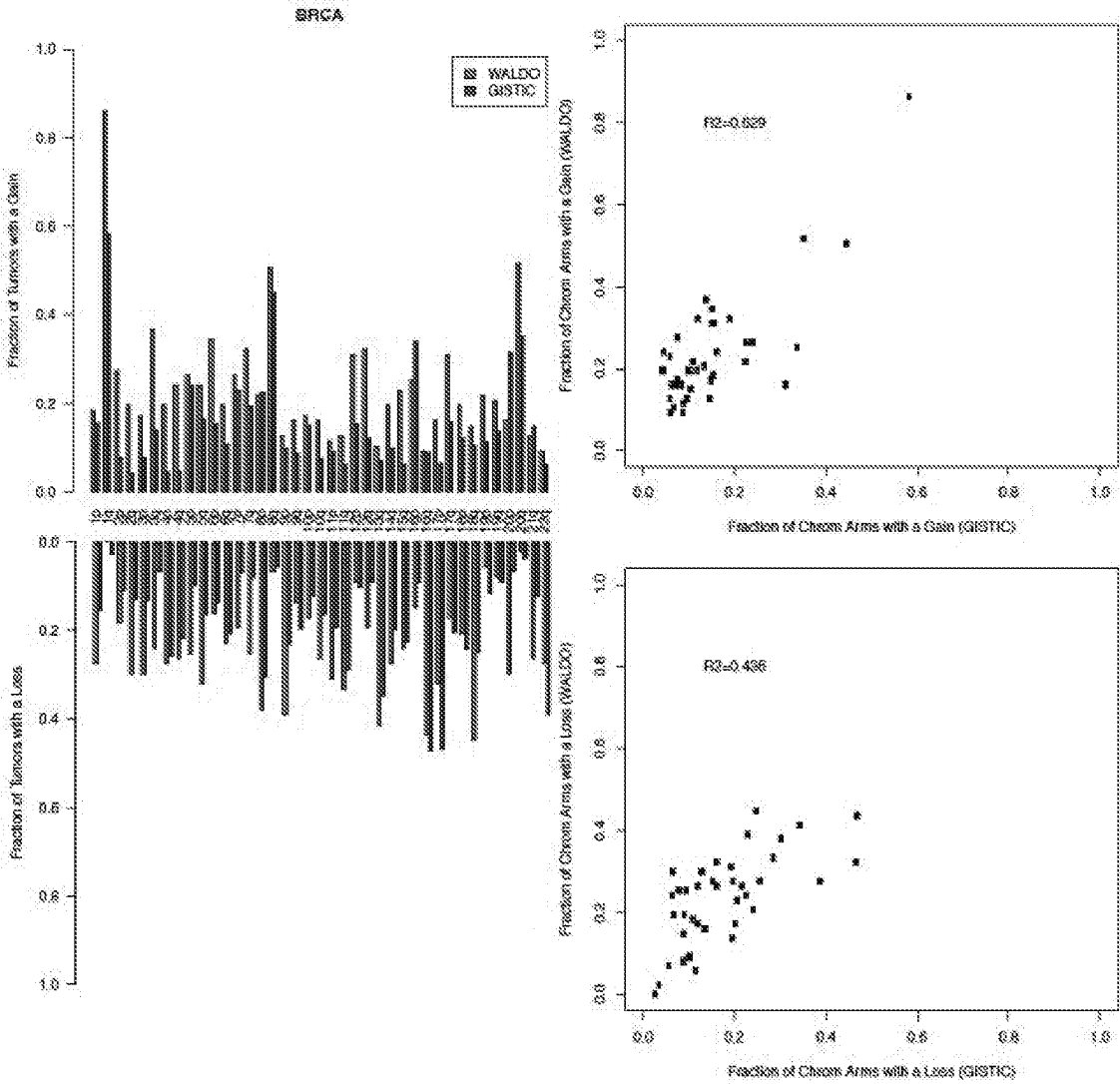


FIG. 40

B

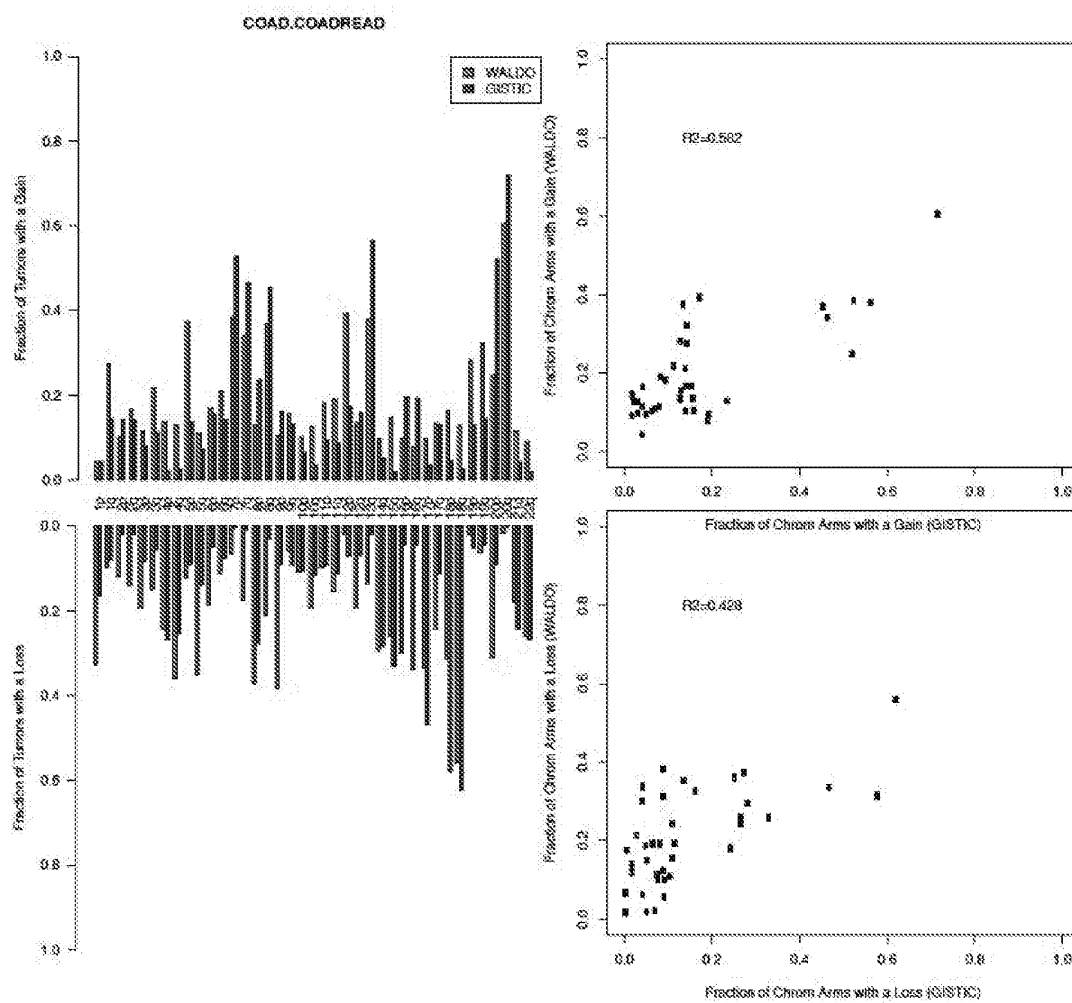


FIG. 40 (continued)

C

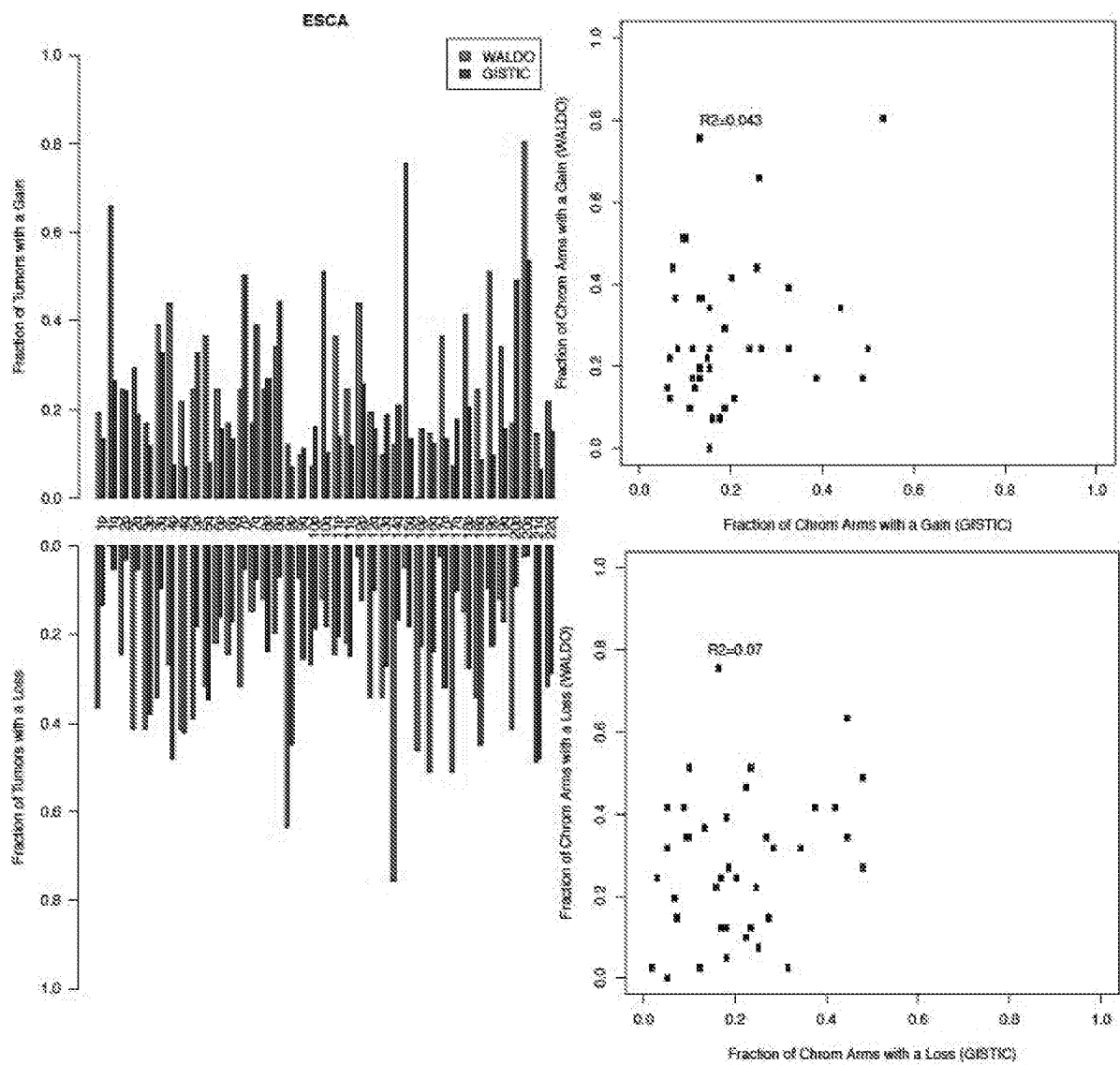


FIG. 40 (continued)

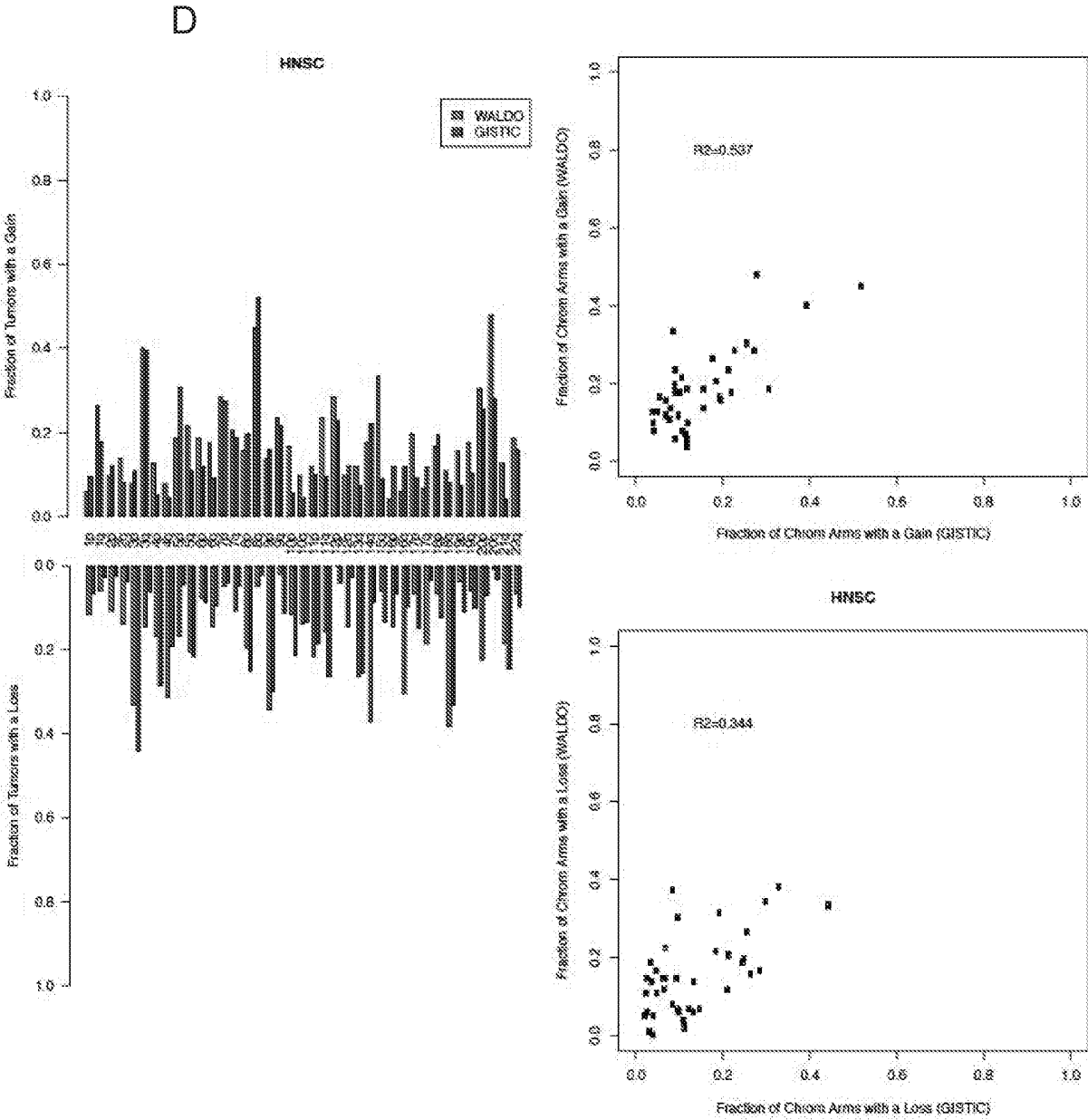


FIG. 40 (continued)

E

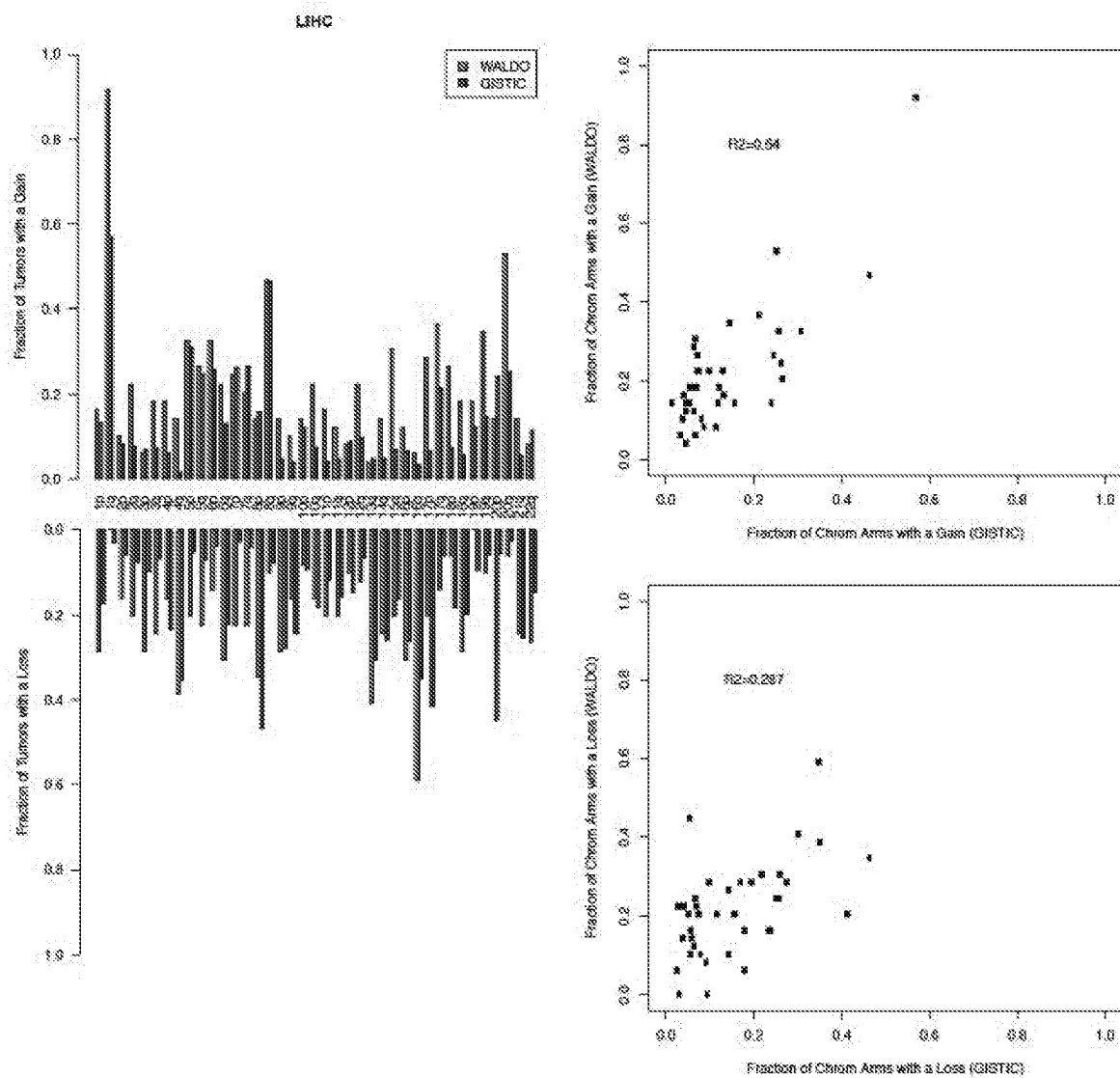


FIG. 40 (continued)

F

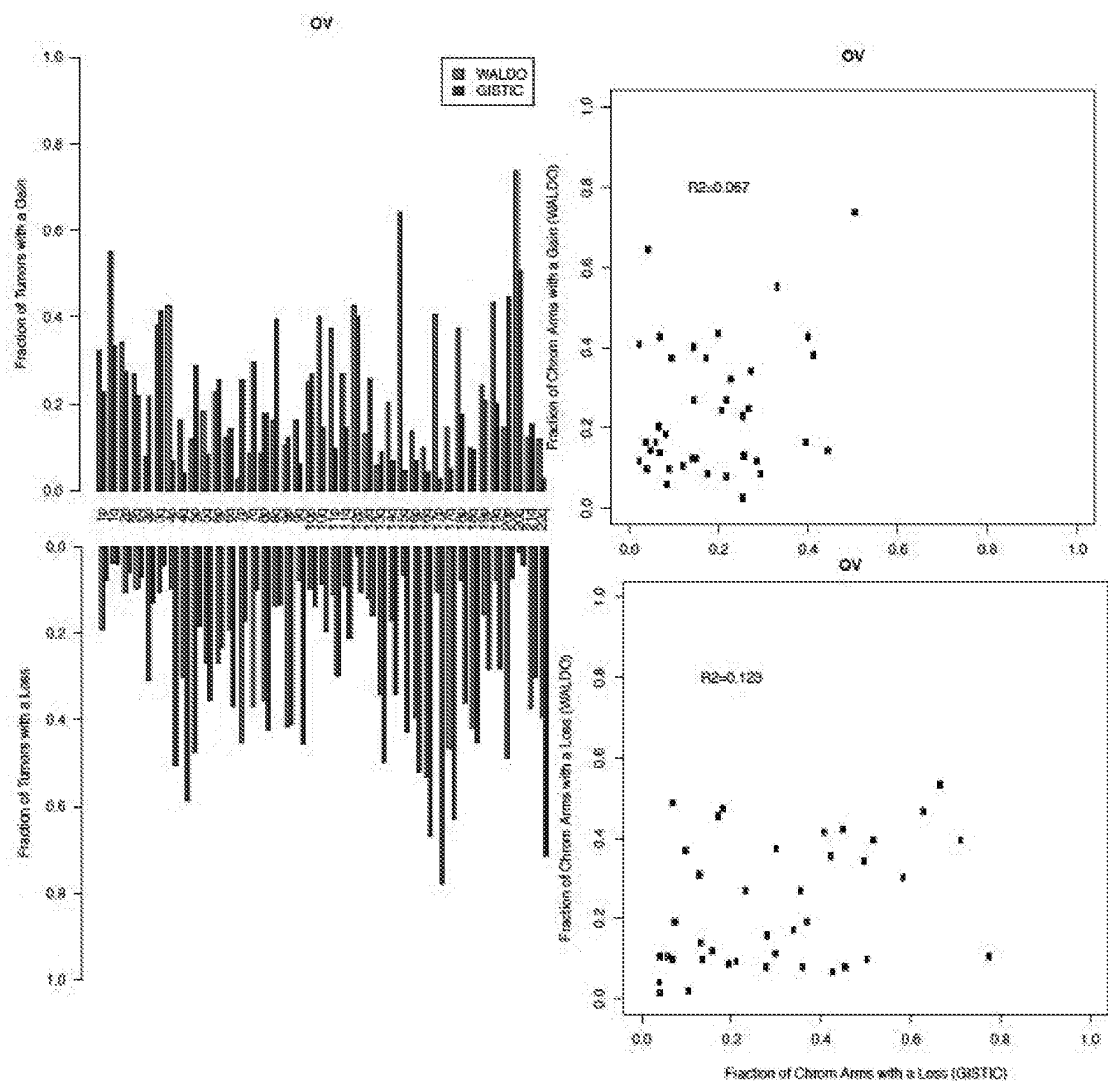


FIG. 40 (continued)

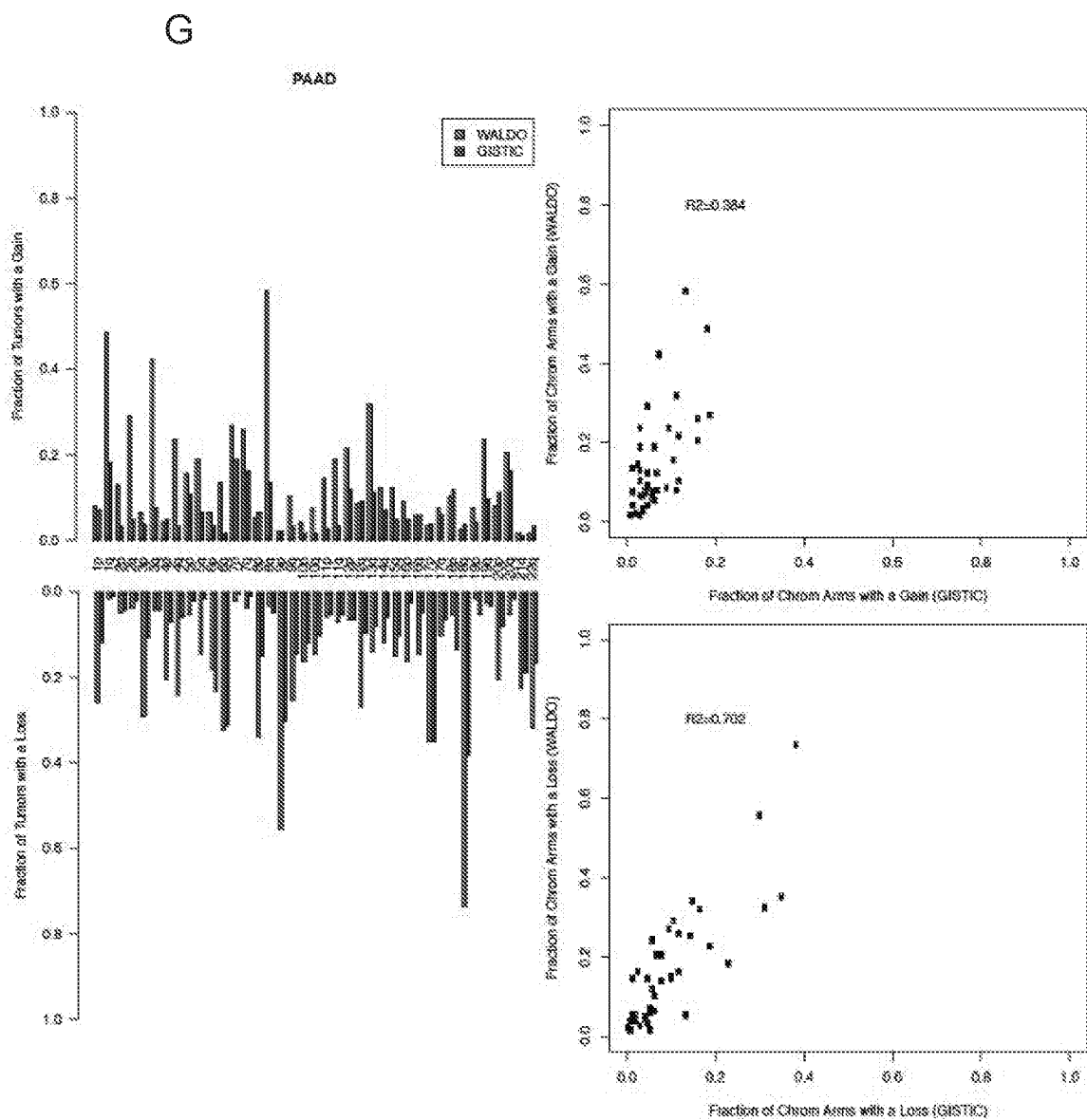


FIG. 40 (continued)

H

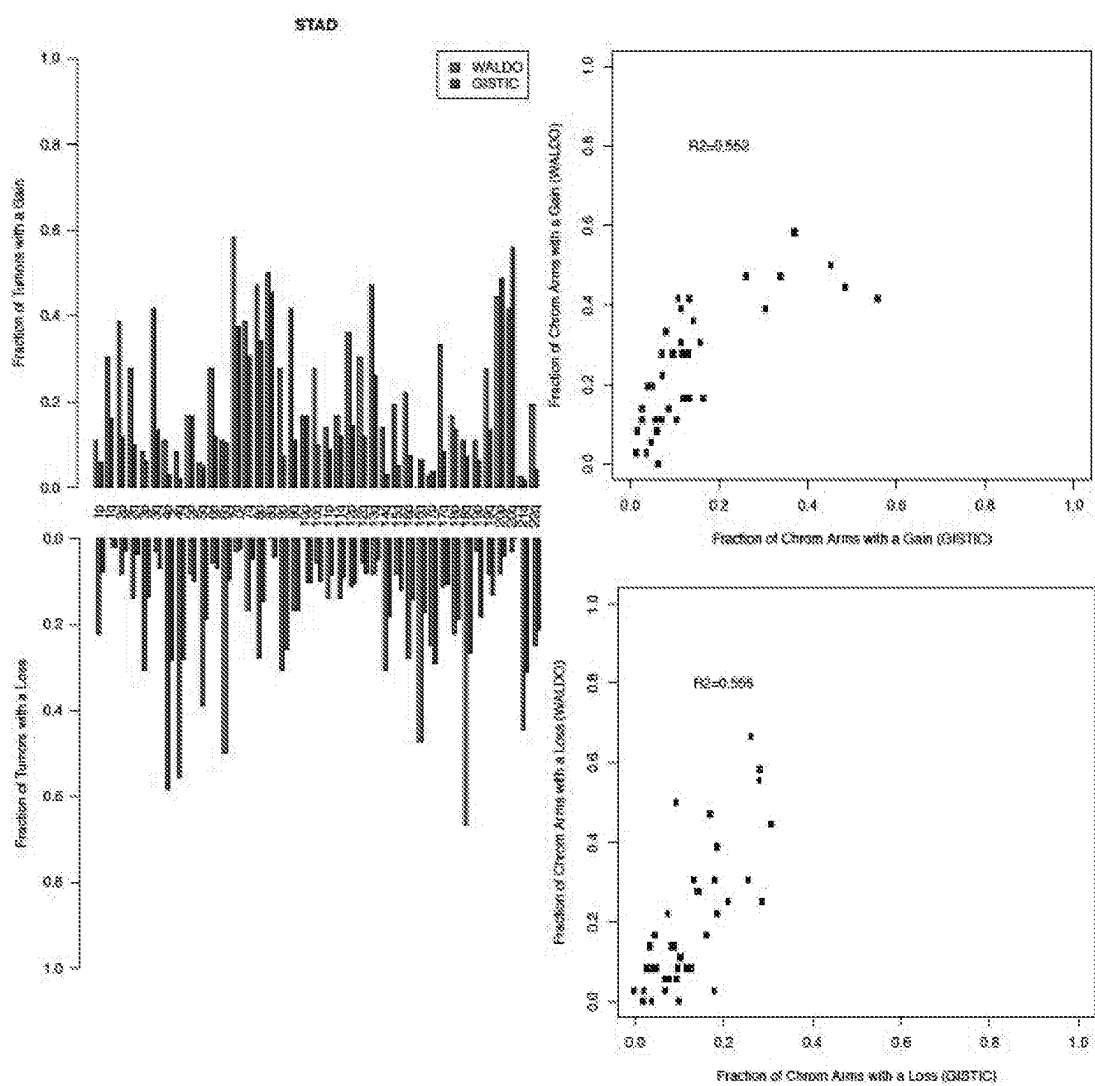


FIG. 40 (continued)

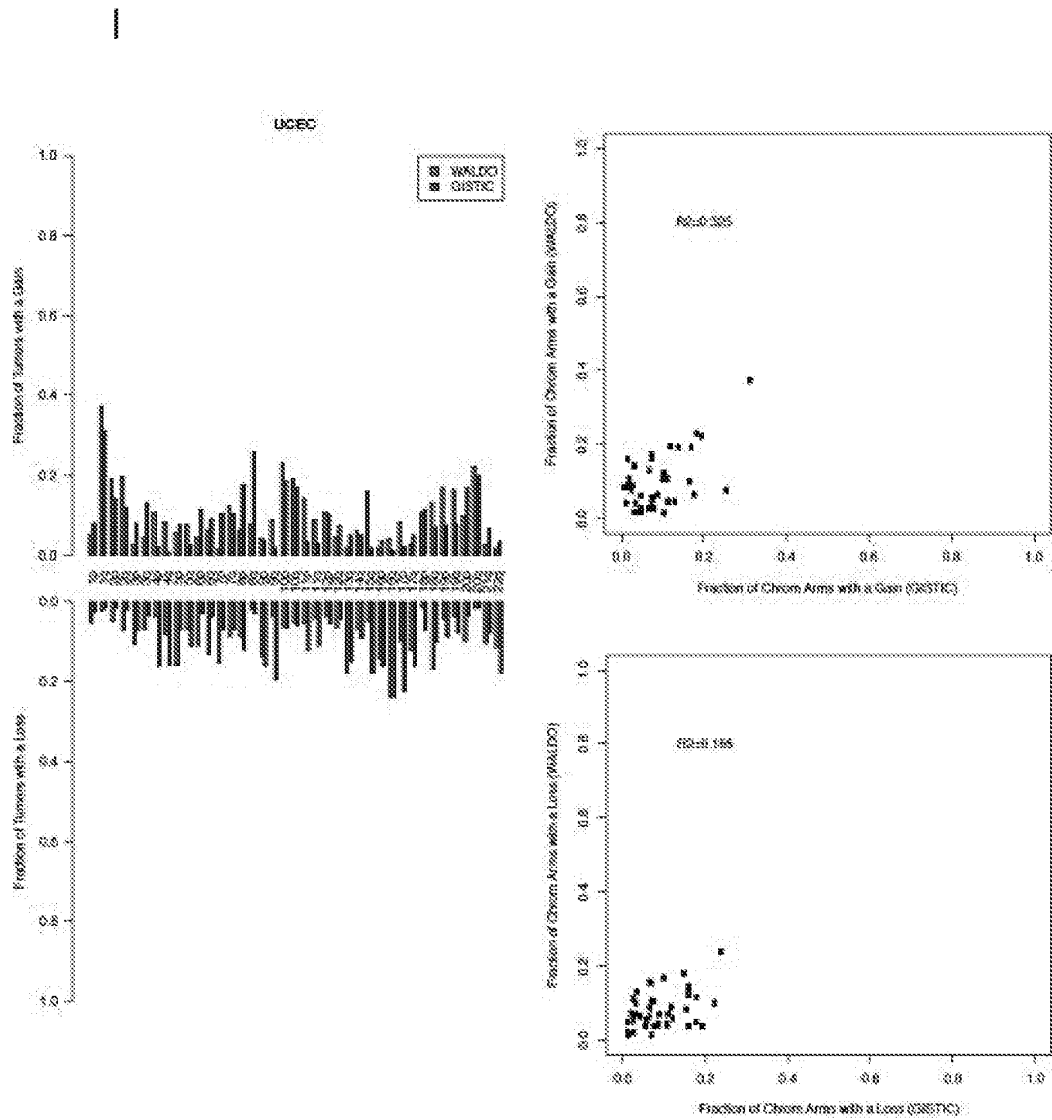


FIG. 40 (continued)

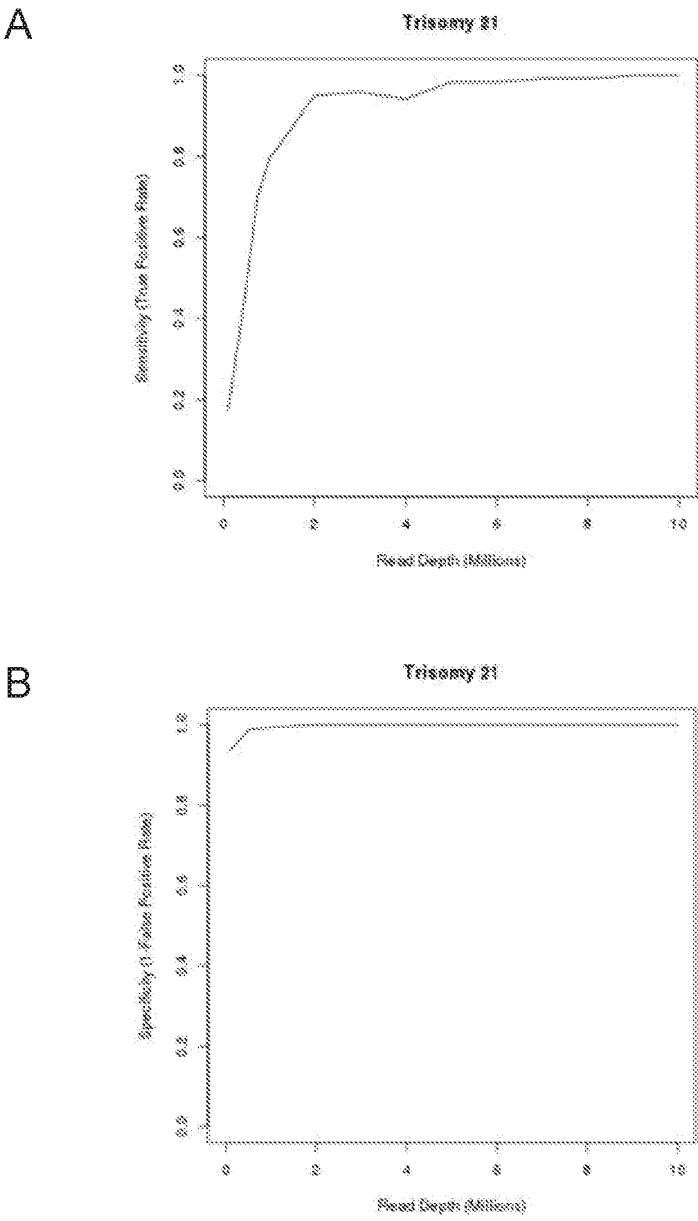


FIG. 41

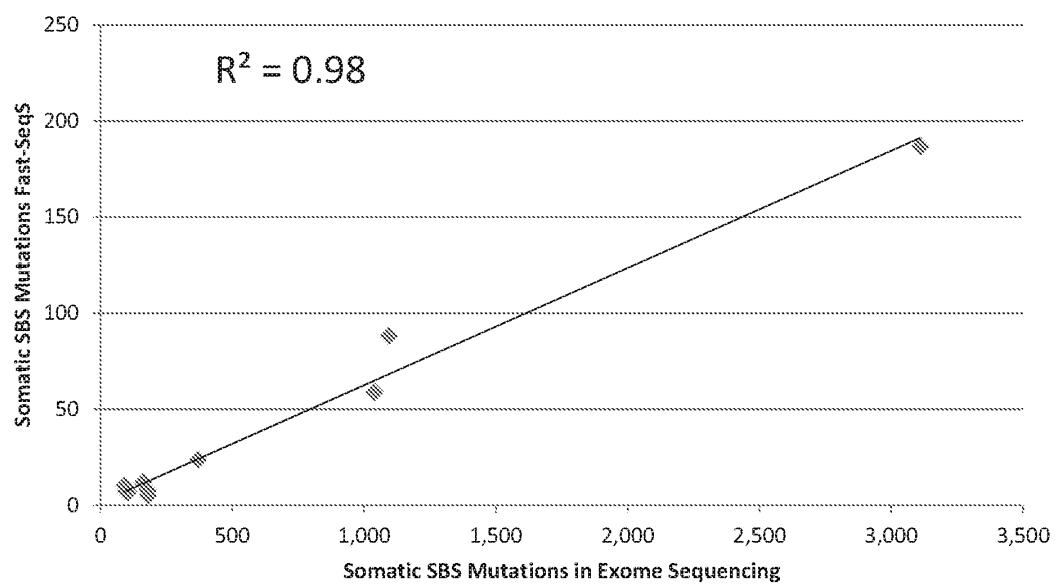


FIG. 42

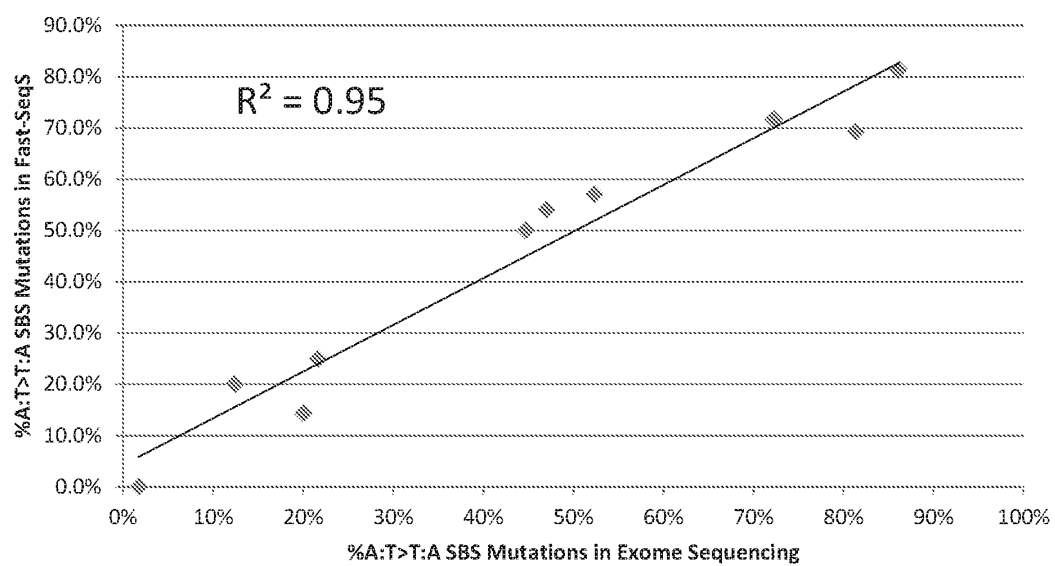


FIG. 43

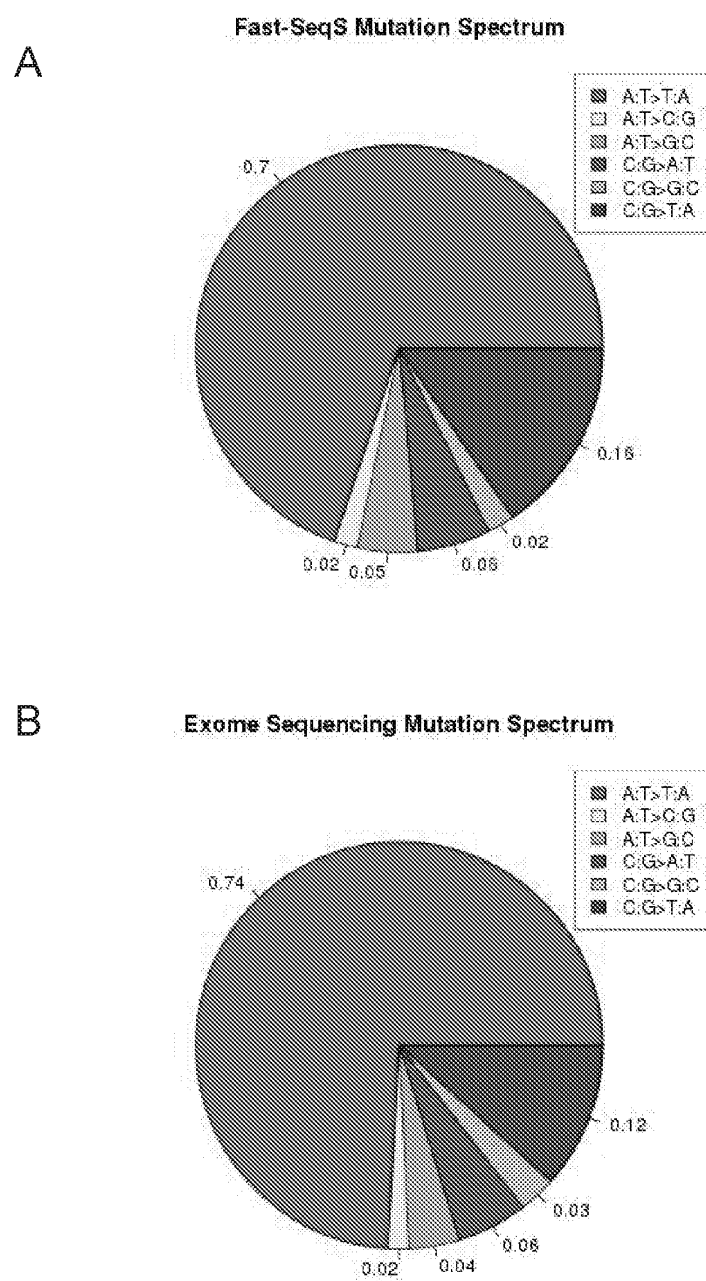


FIG. 44

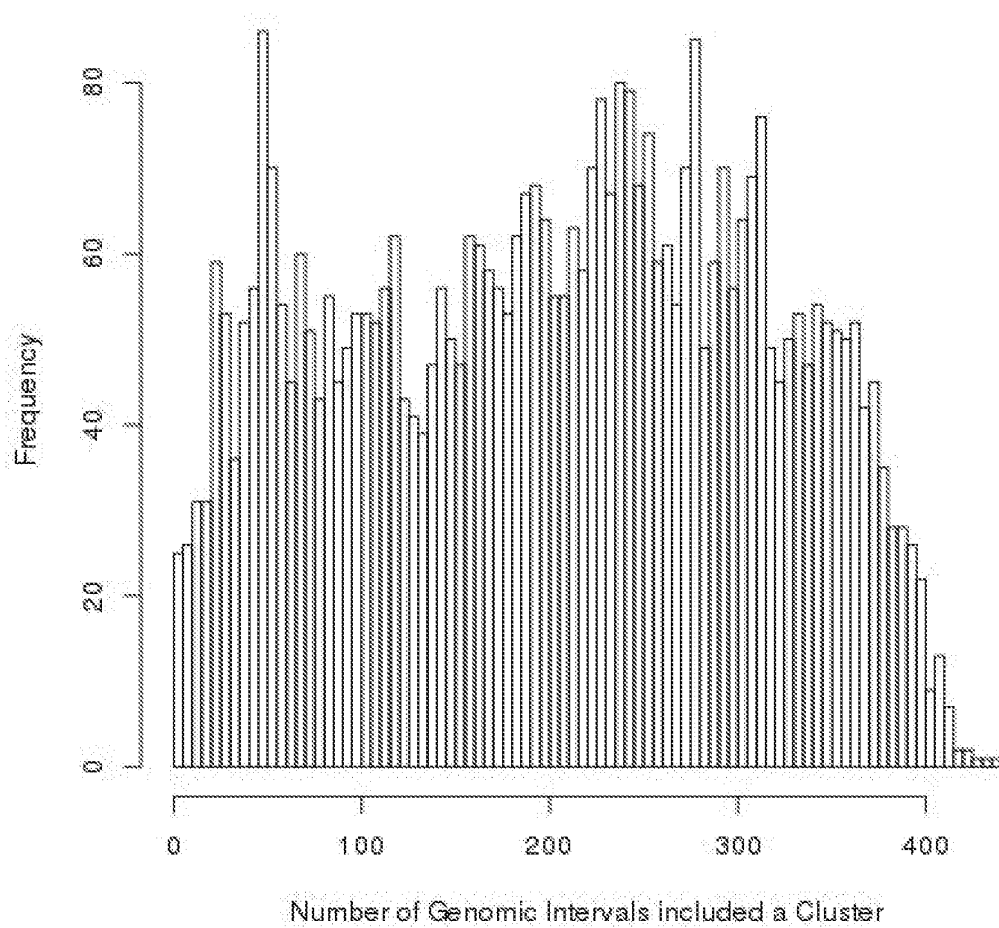


FIG. 45

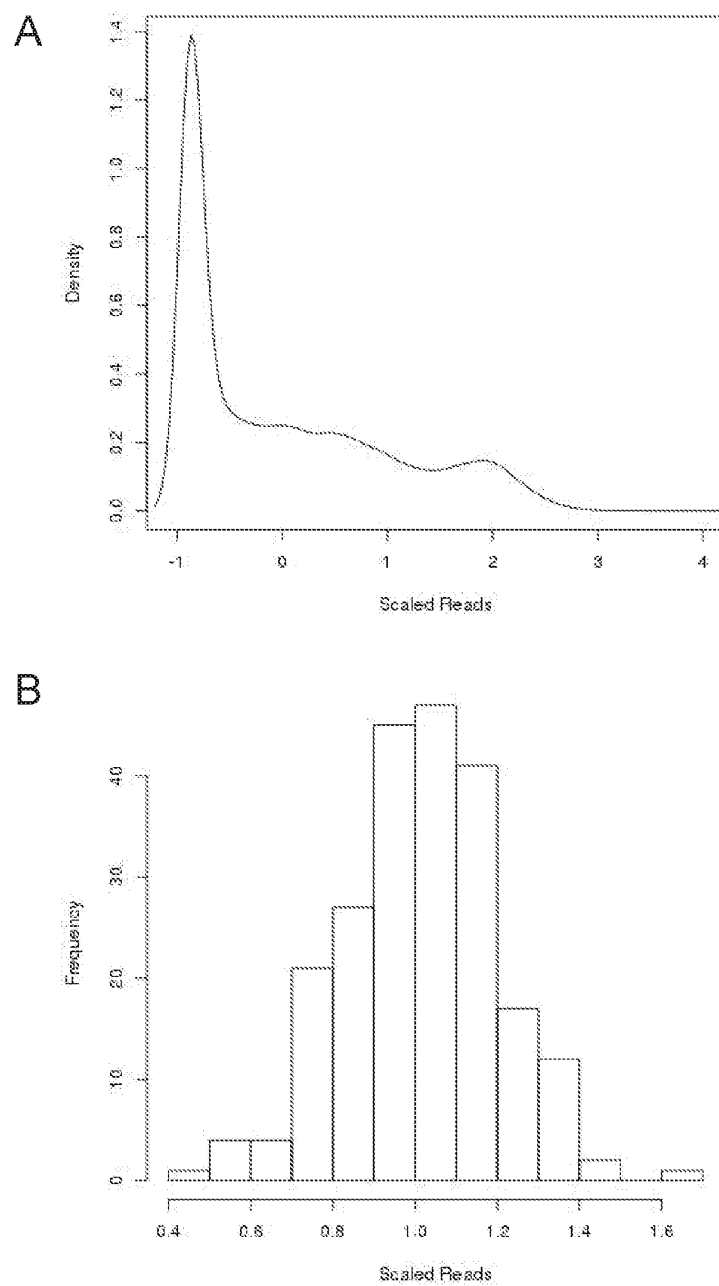


FIG. 46

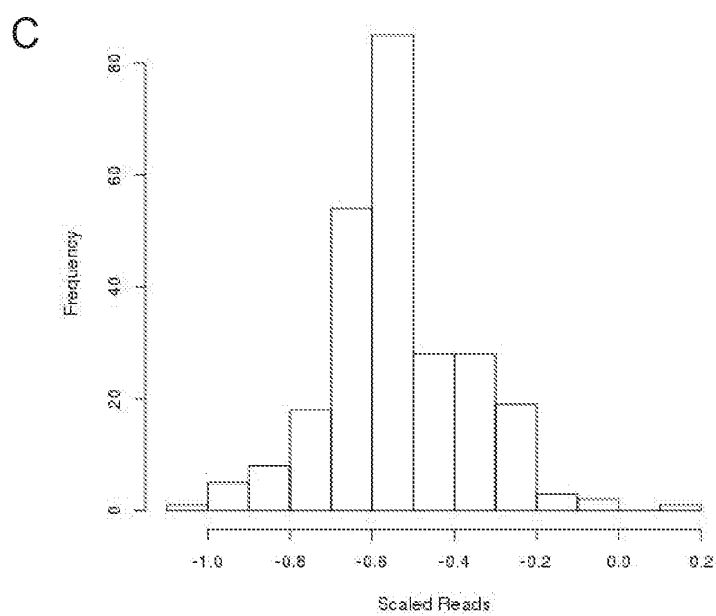
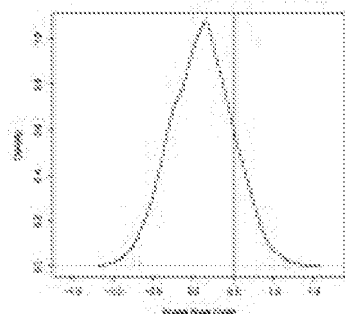


FIG. 46

Call Aneuploidy on a Chromosome Arm



500kb Genomic Interval 1

Estimate parameters:

$$R_1 = N(\mu_1, \sigma_1^2)$$

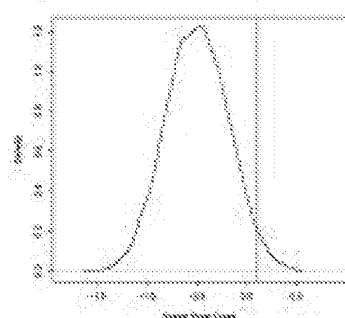
$$R_2 = N(\mu_2, \sigma_2^2)$$

$$\mu_1 = 0.5$$

$$\mu_2 = 0.5$$

$$\sigma_1^2 = 0.15$$

$$P^*(1) = 0.001 \mu_1 = 0.15, \sigma_1^2 = 0.15 \mid R_1 \mid R_2 = 0.201 > 0.001 \text{ is outlier}$$



500kb Genomic Interval 2

Estimate parameters:

$$R_1 = N(\mu_1, \sigma_1^2)$$

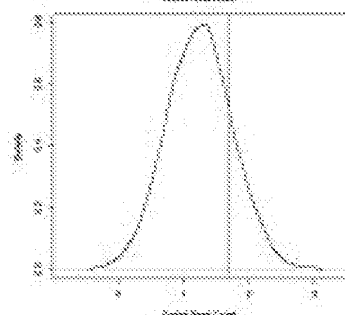
$$R_2 = N(\mu_2, \sigma_2^2)$$

$$\mu_1 = 0.5$$

$$\mu_2 = -0.5$$

$$\sigma_1^2 = 0.1$$

$$P^*(1) = 0.001 \mu_1 = -0.5, \sigma_1^2 = 0.1 \mid R_1 \mid R_2 = 0.201 > 0.001 \text{ is outlier}$$



500kb Genomic Interval 3

Estimate parameters:

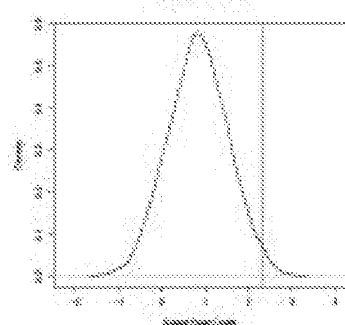
$$R_1 = N(\mu_1, \sigma_1^2)$$

$$\mu_1 = 1.7$$

$$\mu_2 = 1.25$$

$$\sigma_1^2 = 0.25$$

$$P^*(1) = 0.001 \mu_1 = 1.25, \sigma_1^2 = 0.25 \mid R_1 \mid R_2 = 0.201 > 0.001 \text{ is outlier}$$



Chromosome Arm

$$\sum_{i=1}^3 R_i = N(\sum_{i=1}^3 \mu_i, \sum_{i=1}^3 \sigma_i^2)$$

$$\sum_{i=1}^3 \mu_i = (0.5 + 0.5 + 1.7) = 2.7$$

$$\sum_{i=1}^3 \sigma_i^2 = 0.15 + (-0.5) + 0.25 = 0.05$$

$$\sum_{i=1}^3 \sigma_i^2 = 0.15 + 0.1 + 0.25 = 0.5$$

$$5.004(\sum_{i=1}^3 \mu_i = 0.05, \sum_{i=1}^3 \sigma_i^2 = 0.5) \mid R_1 \mid R_2 = 2.0) = 0.0201 < 0.001$$

$$p = 0.05$$

Chromosome arm is gained

FIG. 47

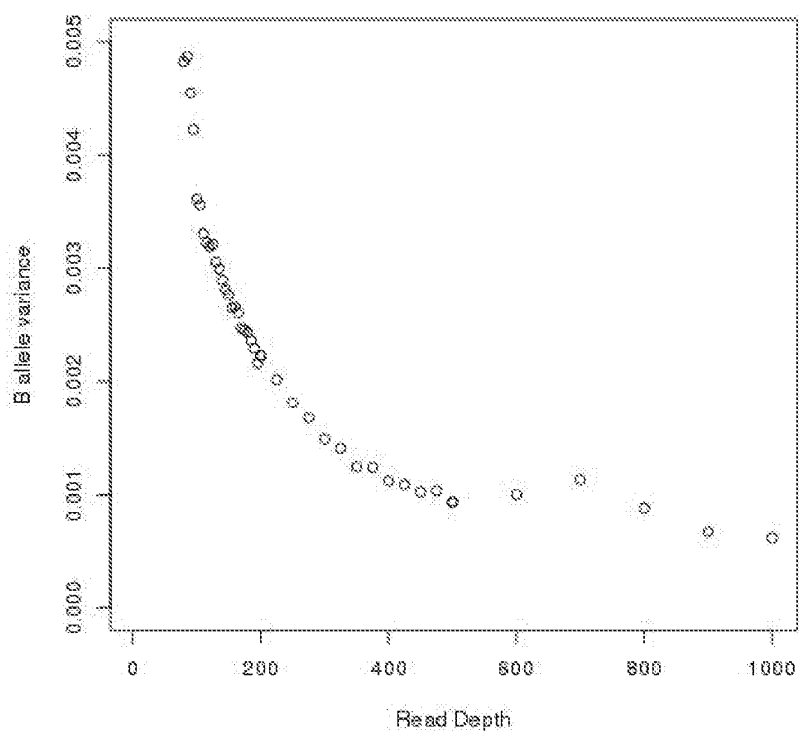


FIG. 48

Synthetics with one arm alterations

```

N <- number of normal WBC samples used
f <- desired neoplastic cell fraction
r <- chromosome arm (1p..22q)
r_type <- alteration type (gain or loss)
p(r) <- probability of arm gain (0.5 default)
rho <- desired unique read depth of synthetic sample (9M default)
j<-desired number of repeats

For j=1:5
  For i=1:N
    s <- sample[i]
    U <- get UIDs from s
    For f in 1:F
      For r in 1p:22q #select alteration types and spike-in to t
        For r_type in gain loss
          h <- new synthetic sample
          t <- copy of s
          u_ref <- get all reference allele UIDs that that map
                    to chr arm r
          if r_type== gain
            t <- add duplicate u_ref reads
          else (r_type == loss)
            t <- subtract u_ref reads
          Endif
          h_normal<-random select (1-f)*rho UIDs from s
          h_aneuploid<-random select f*rho UIDs from t
          h<-h_normal+h_aneuploid
        End
      End
    End
  End
End

```

FIG. 49

```

N <- number of normal WBC samples used
d <- desired degree of aneuploidy (number arms altered)
f <- desired neoplastic cell fraction
r <- chromosome arm (1p..22q)
r_type <- arm alteration type (gain or loss)
p(r) <- probability of arm gain (0.5 default)
rho <- desired unique read depth of synthetic sample (9M default)
j <- desired number of repeats

For j=1:5
  For i=1:N
    s <- sample[i]
    U <- get UIDs from s
    For f in 1:F
      h <- new synthetic sample
      For d in (5,10,15,20,25) #desired numbers of altered arms
        t <- copy of s
        For a in 1:d #select alteration types and spike-in to t
          r <- random select arm
          r_type <- alteration type with p(r)
          u_ref <- get all reference allele UIDs that that map
                    to chr arm r
          if r_type== gain
            t <- add duplicate u_ref reads
          else (r_type == loss)
            t <- subtract u_ref reads
          Endif
          h_normal <- random select (1-f)*rho UIDs from s
          h_aneuploid <- random select f*rho UIDs from t
          h <- h_normal + h_aneuploid
        End
      End
    End
  End
End

```

FIG. 50

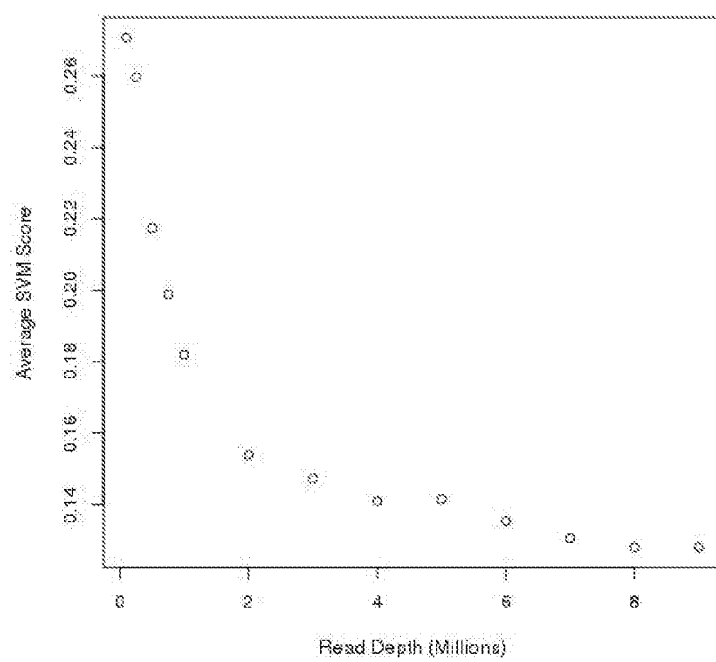


FIG. 51

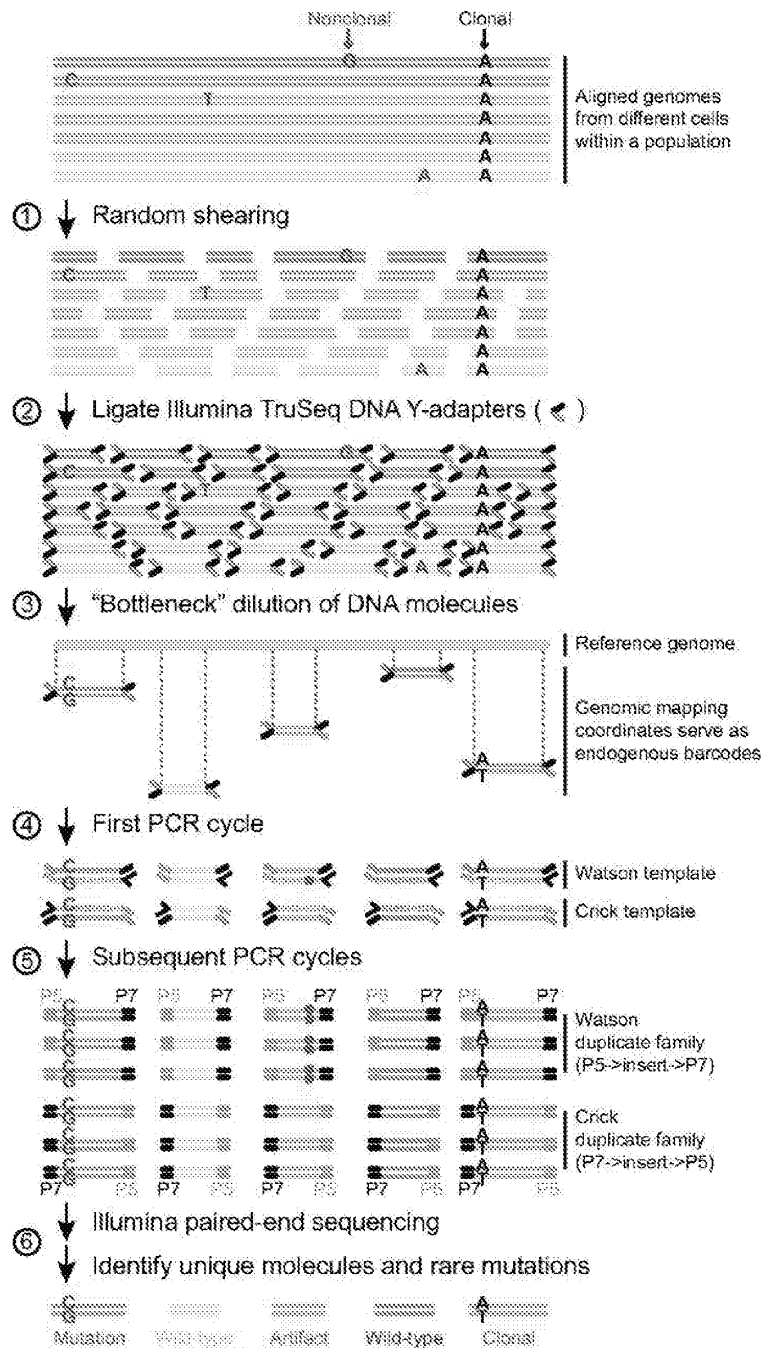


FIG. 52

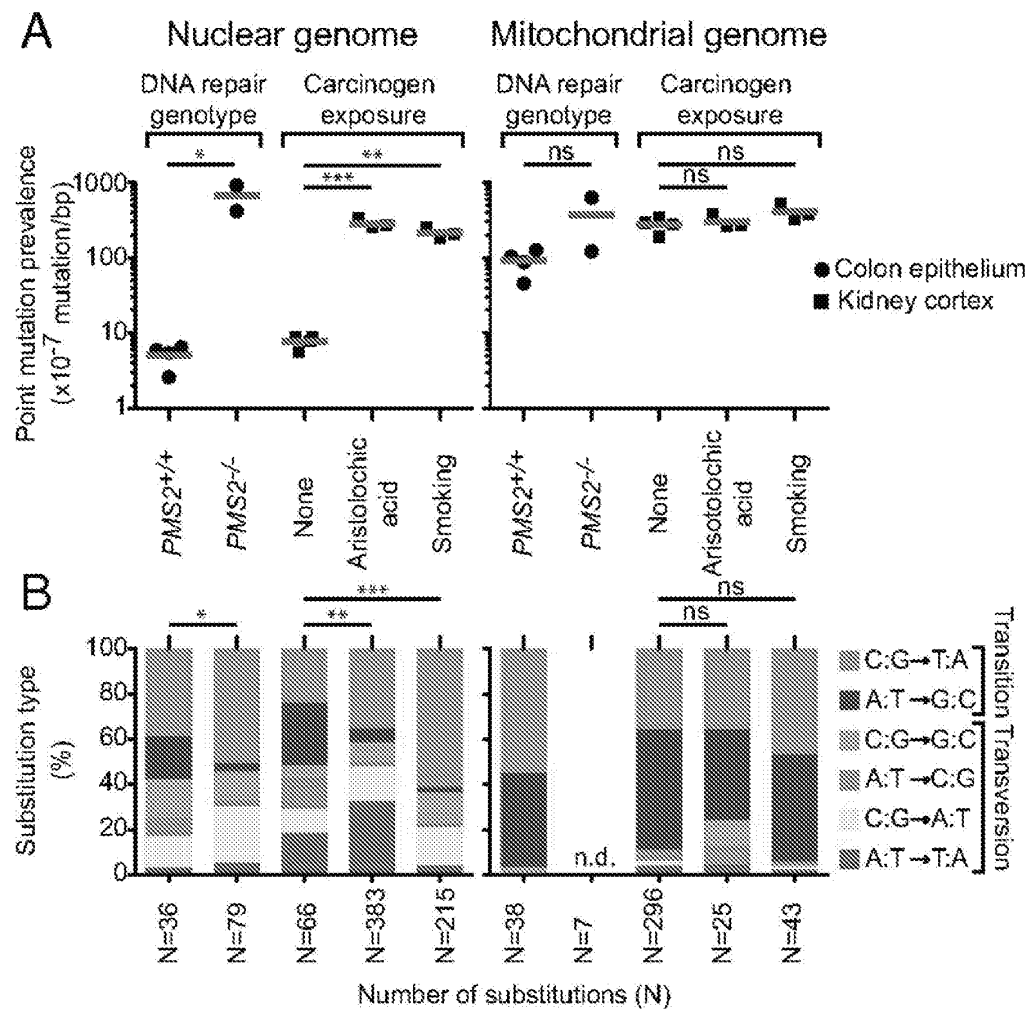


FIG. 53

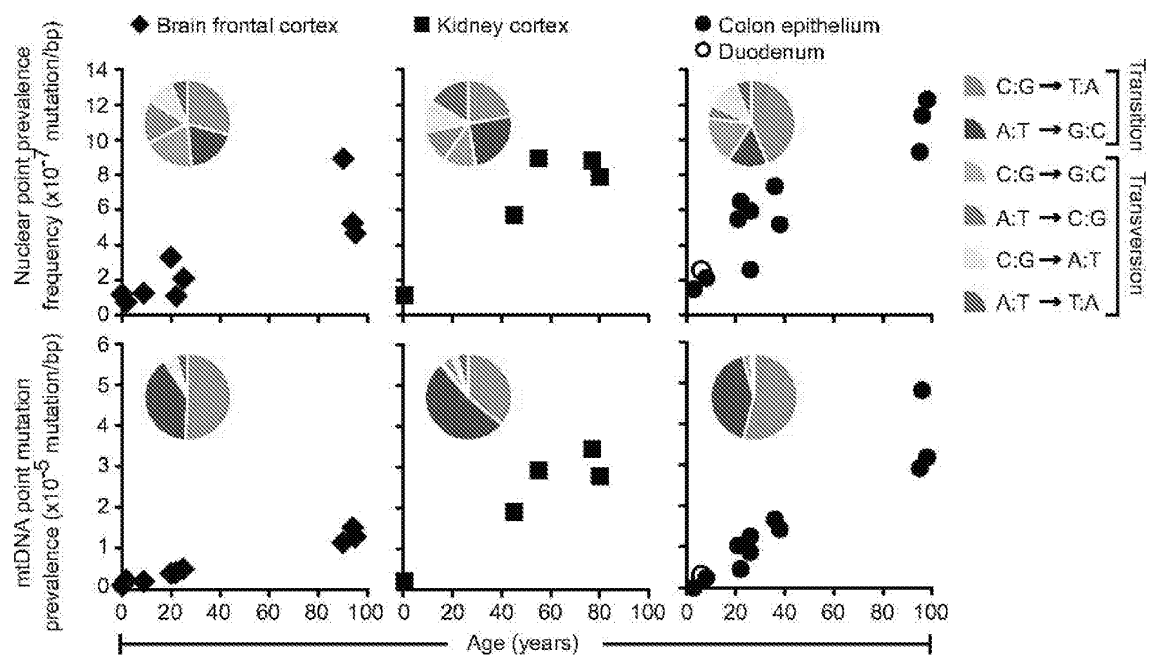


FIG. 54

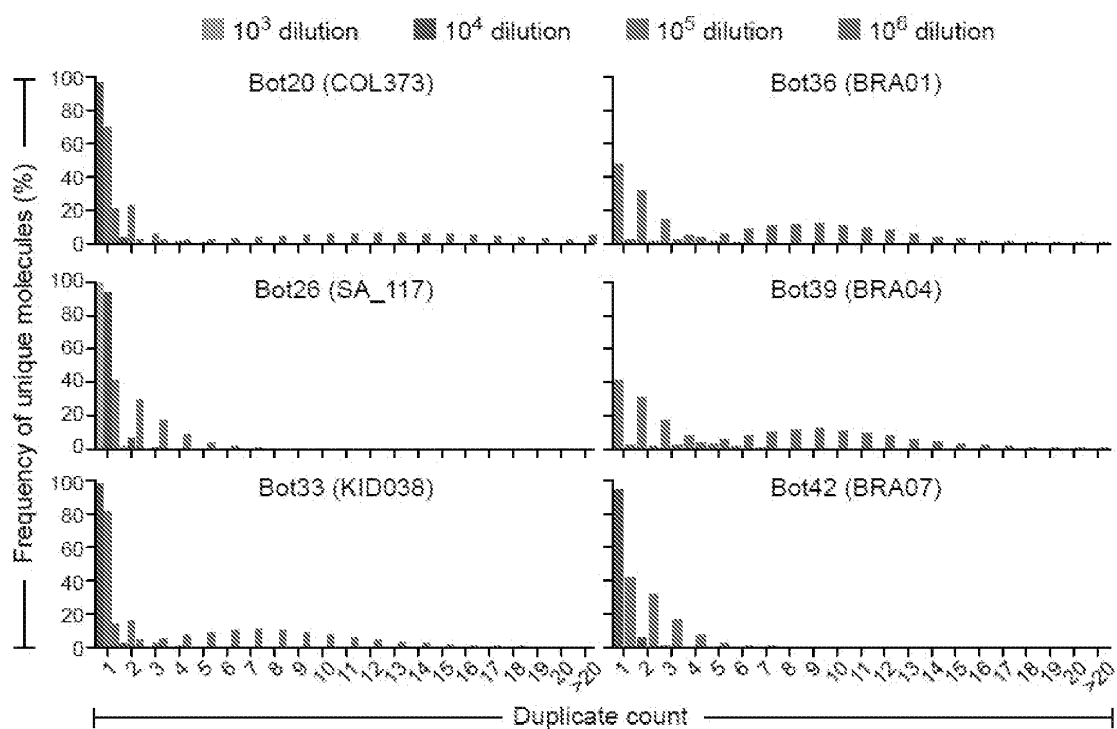


FIG. 55

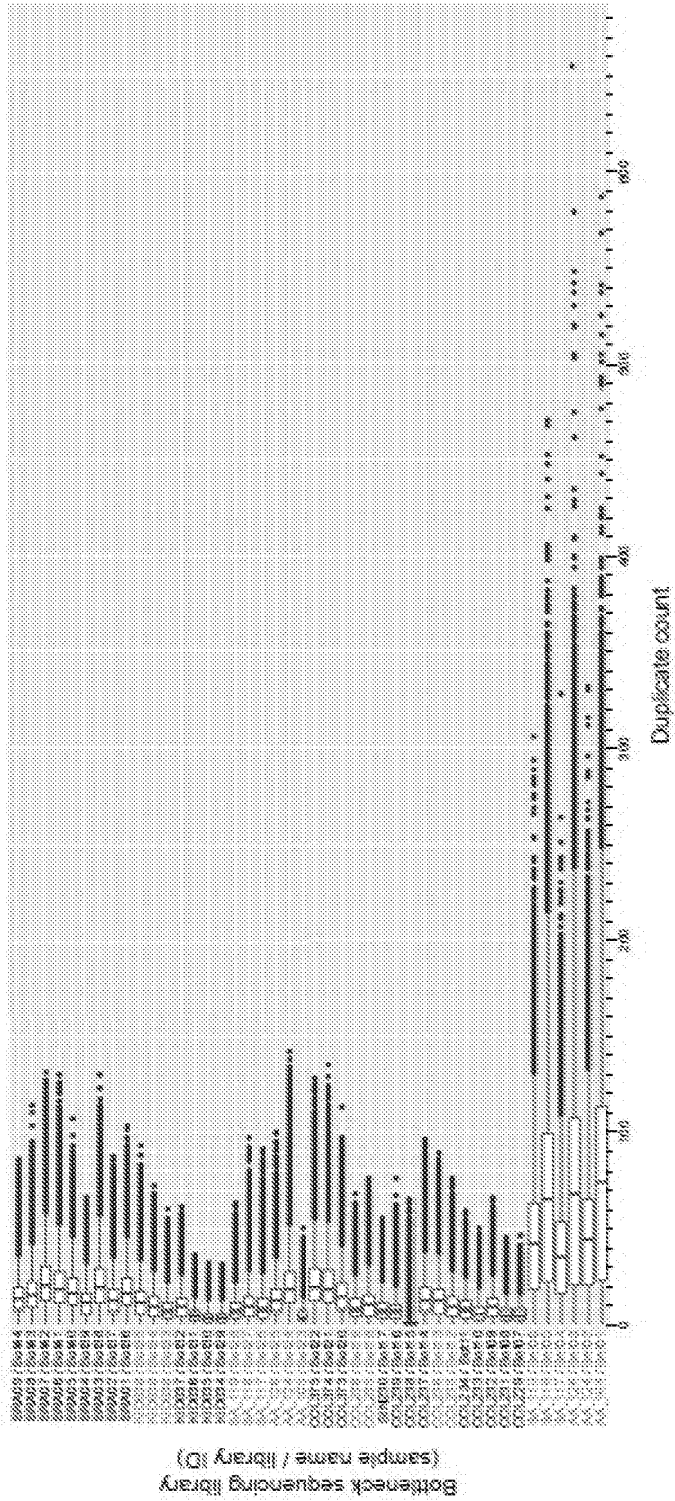


FIG. 56

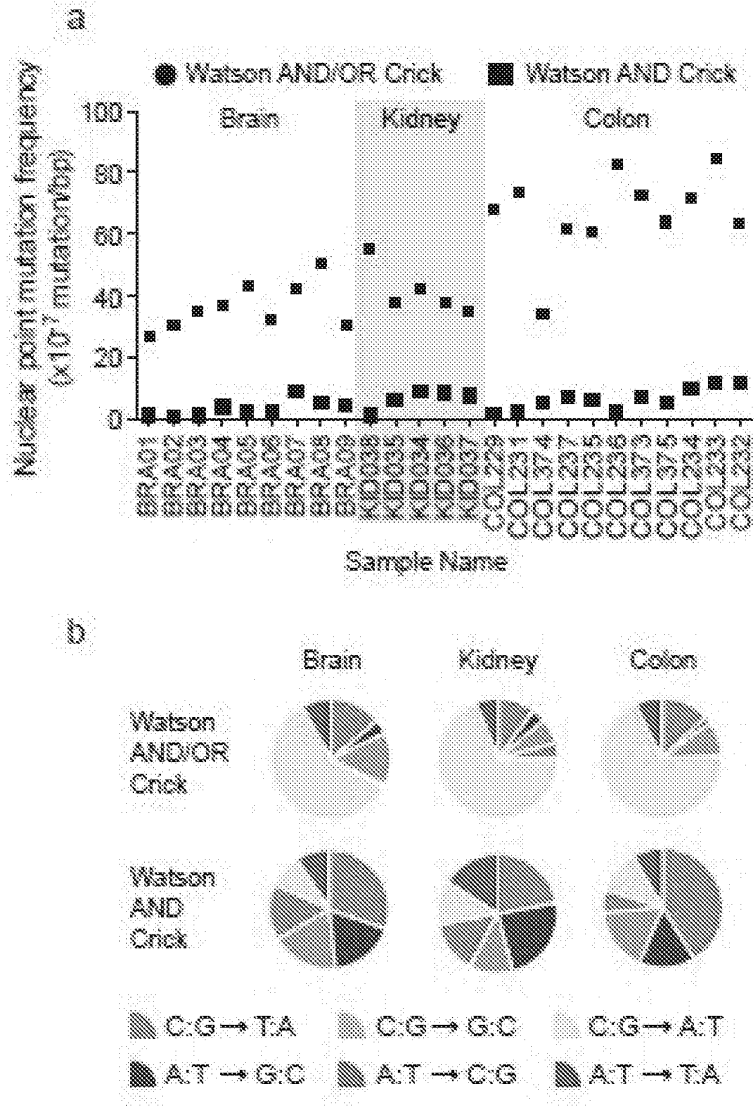


FIG. 57

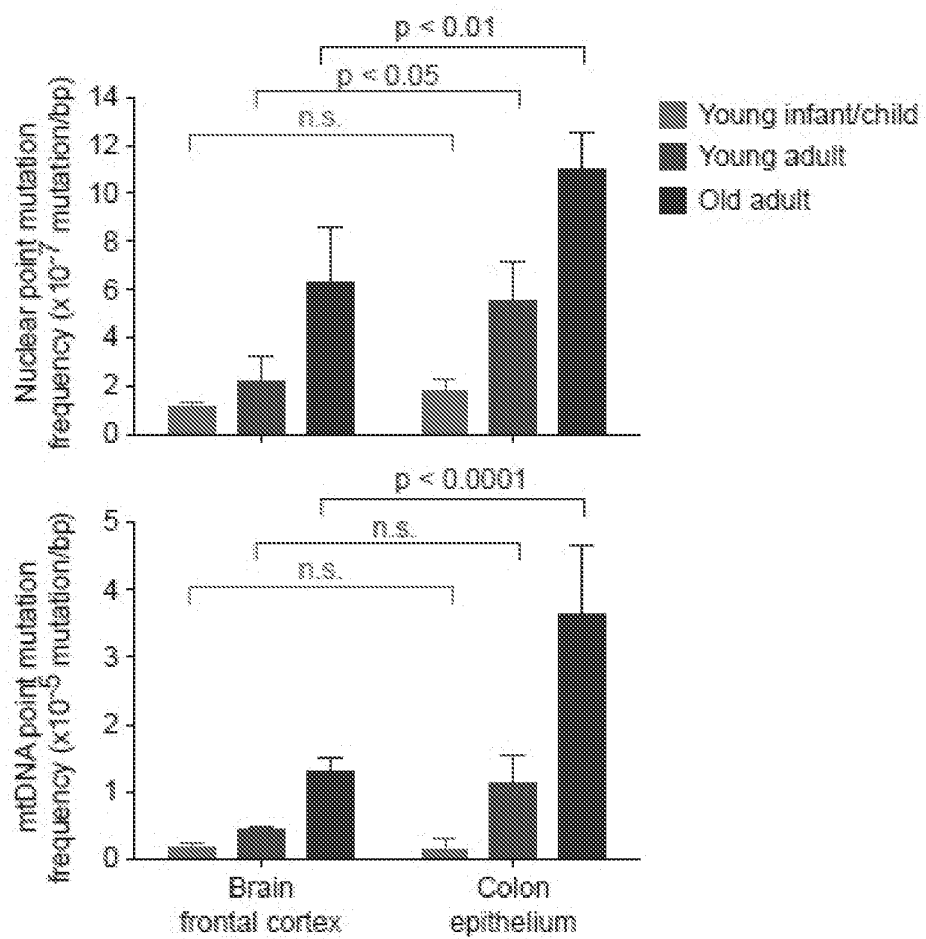


FIG. 58

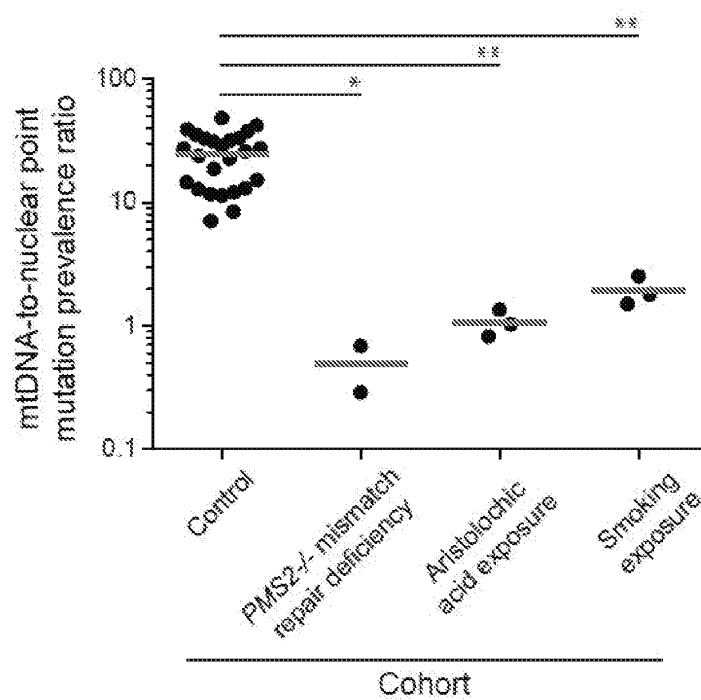


FIG. 59

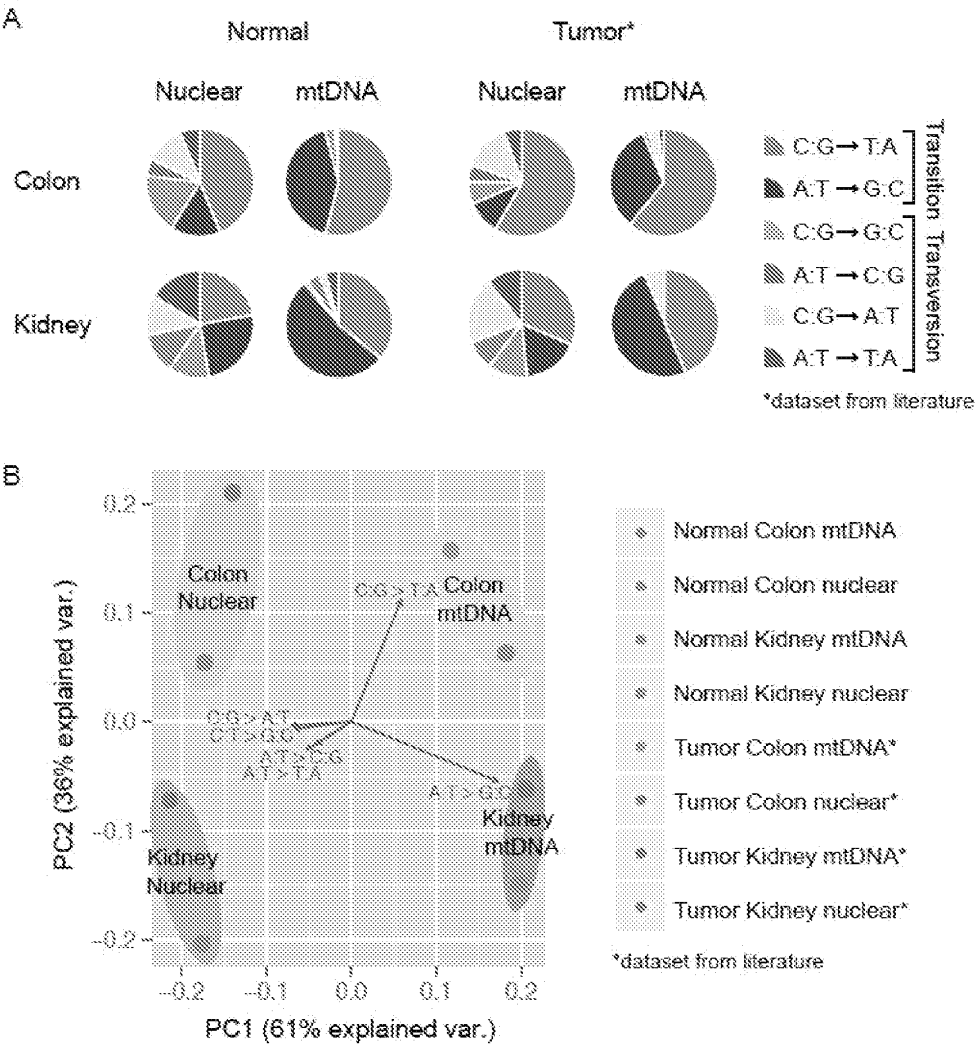


FIG. 60

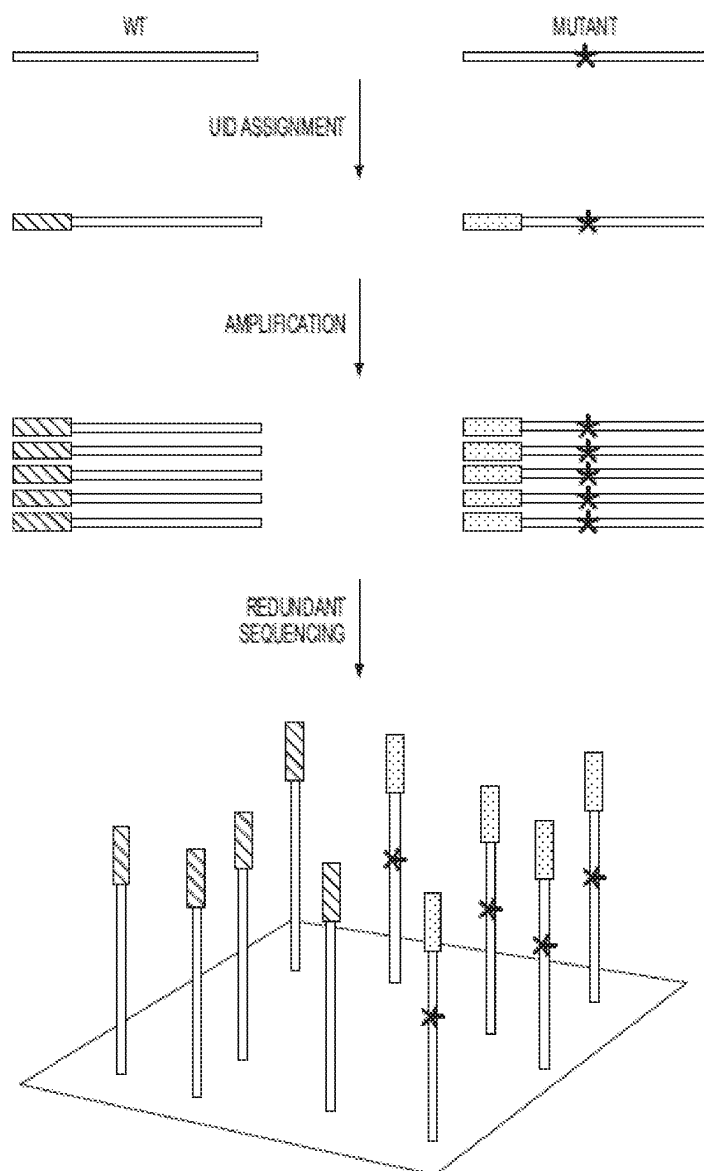


FIG. 61

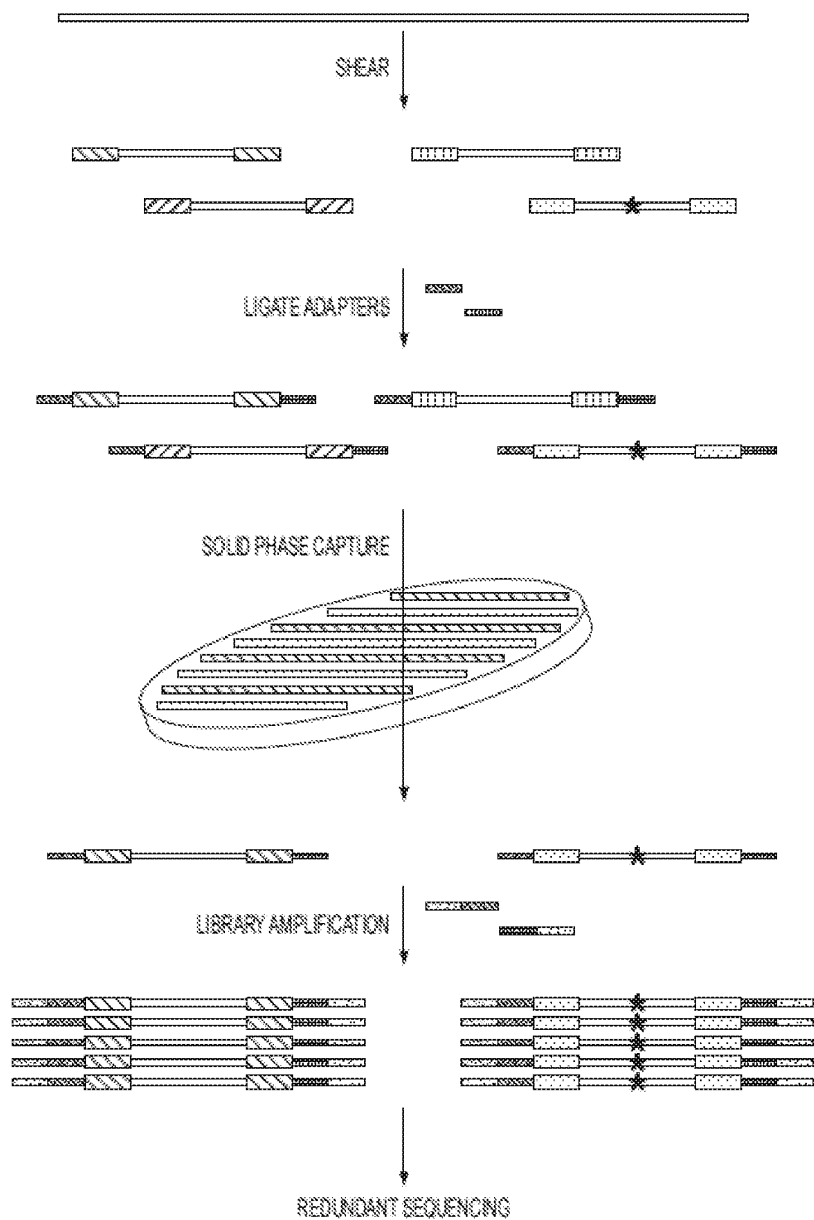


FIG. 62

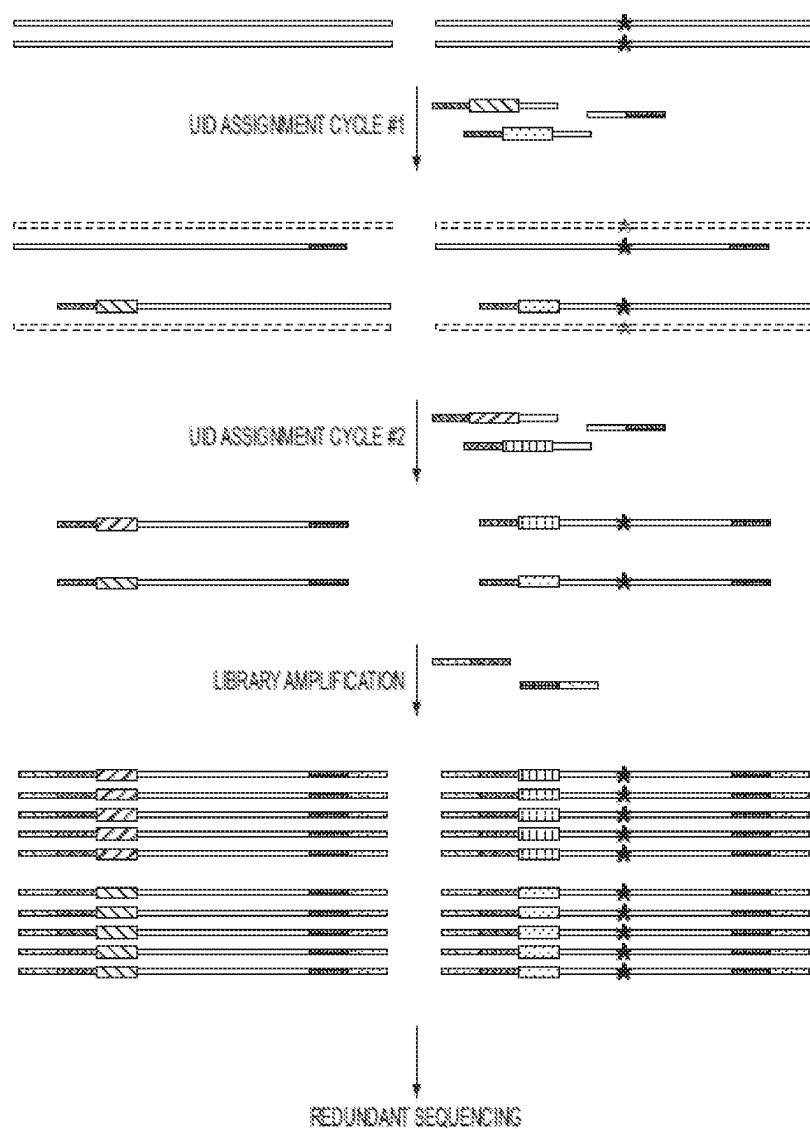
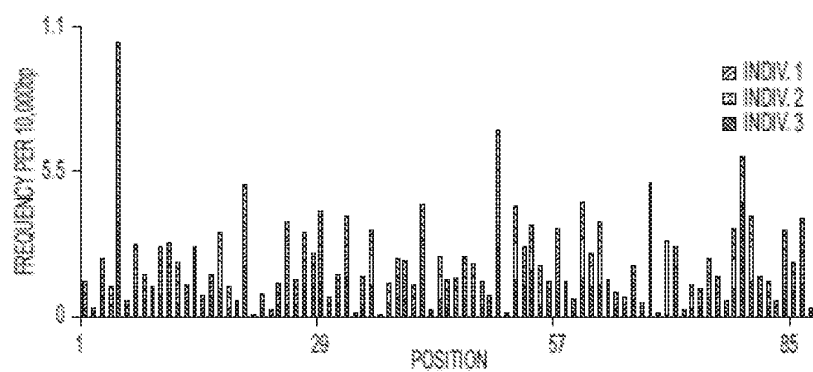


FIG. 63

A



B

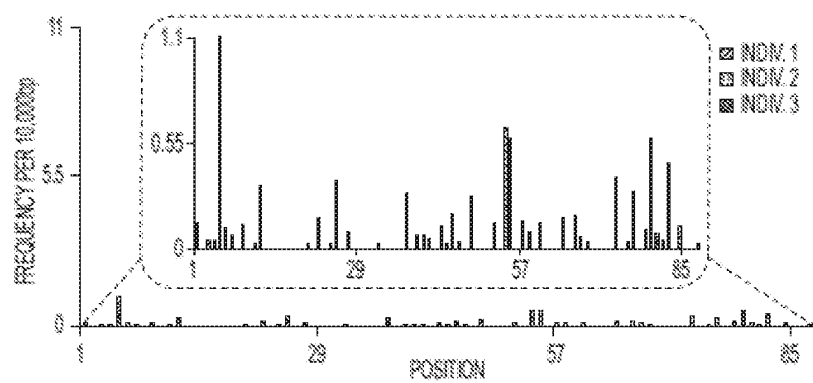


FIG. 64

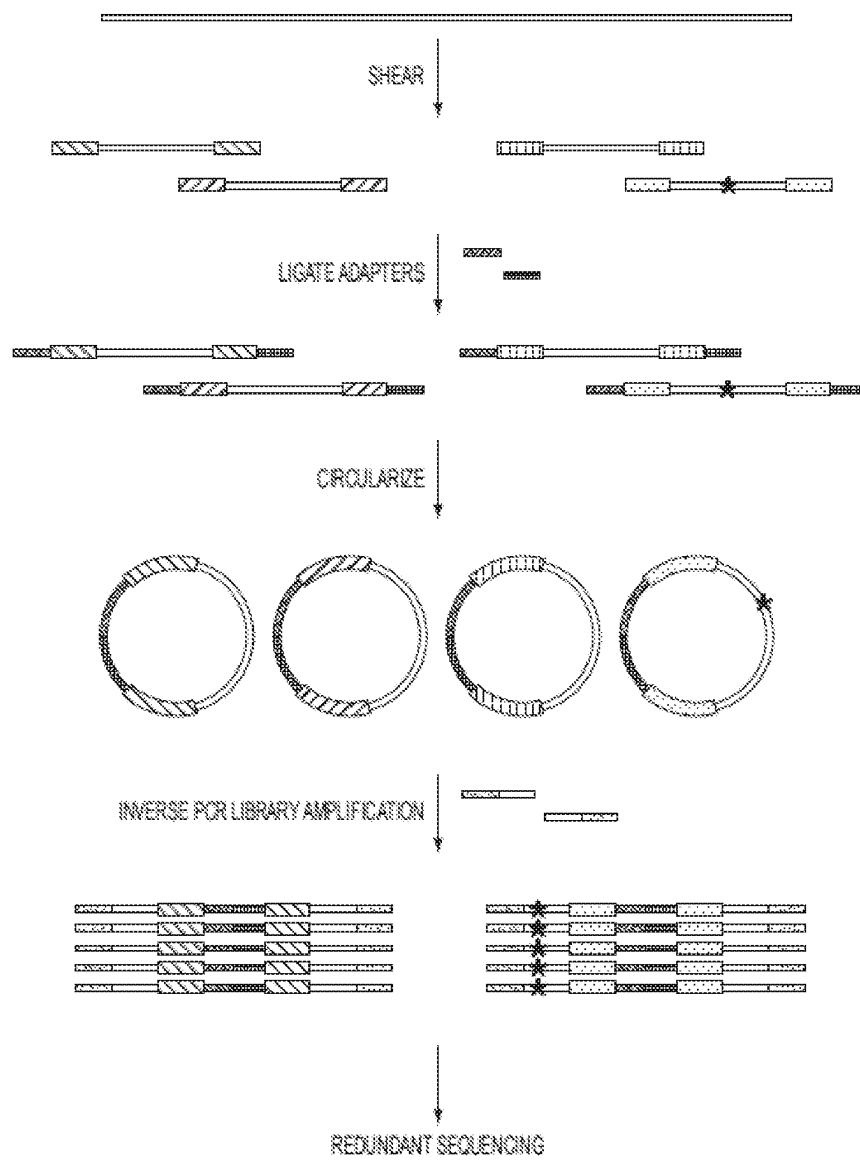


FIG. 65

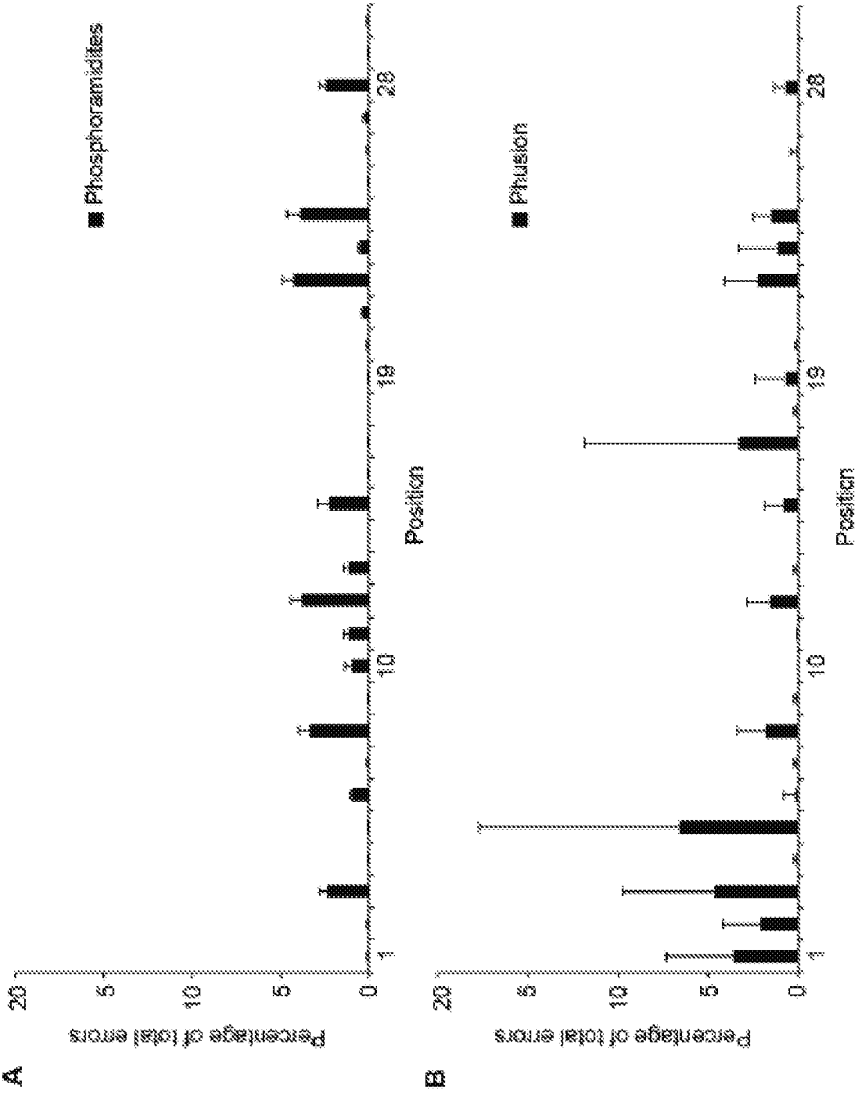


FIG. 66

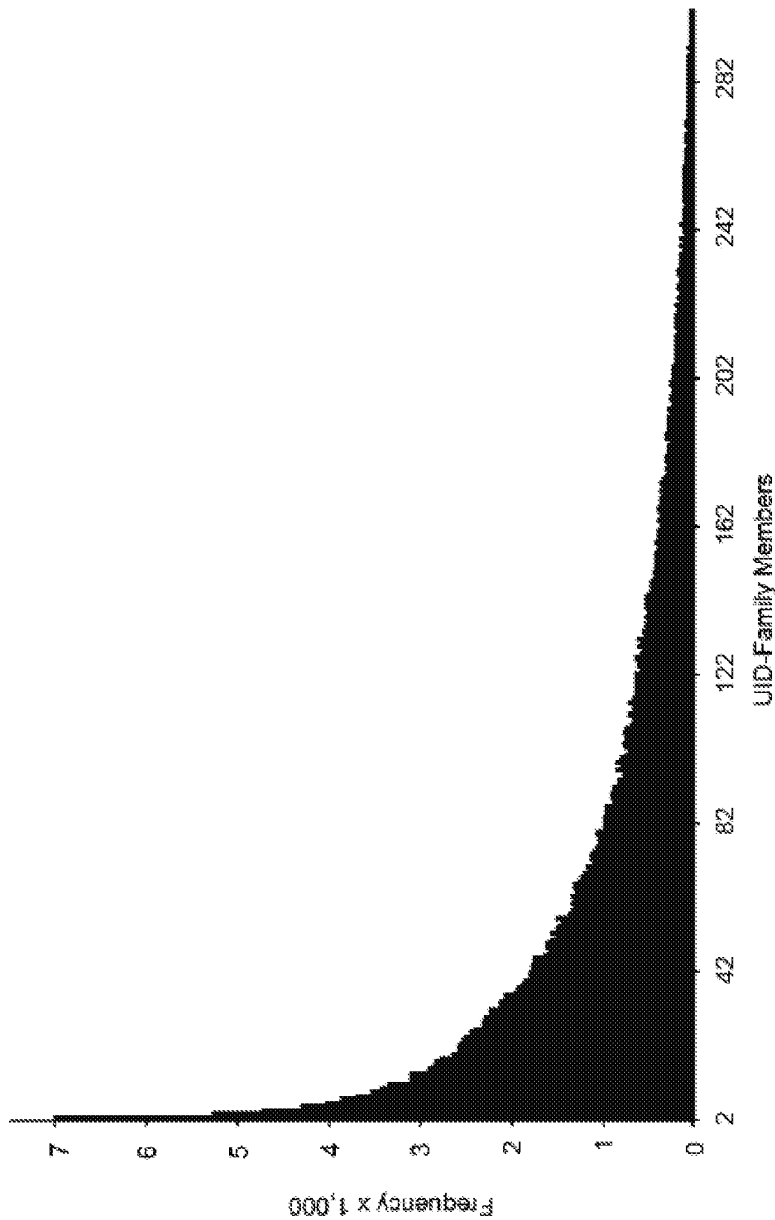
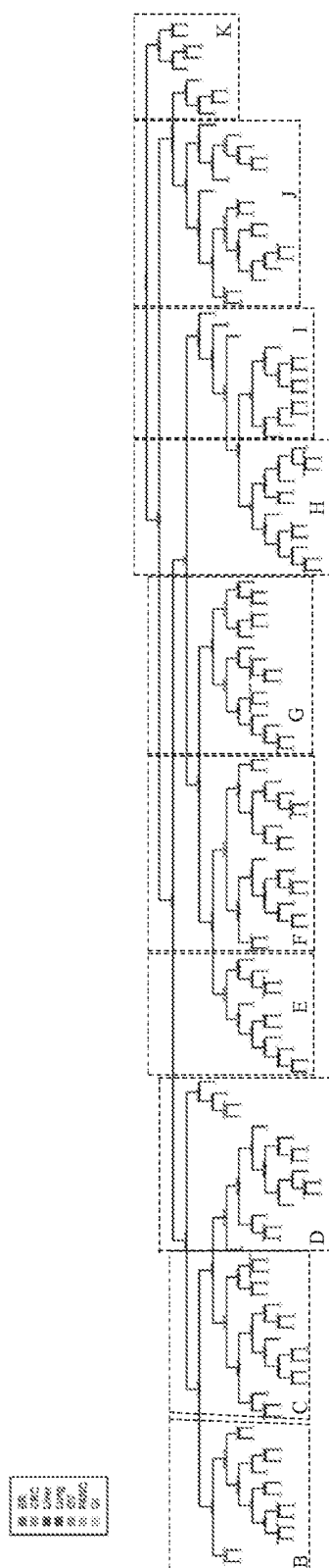


FIG. 67



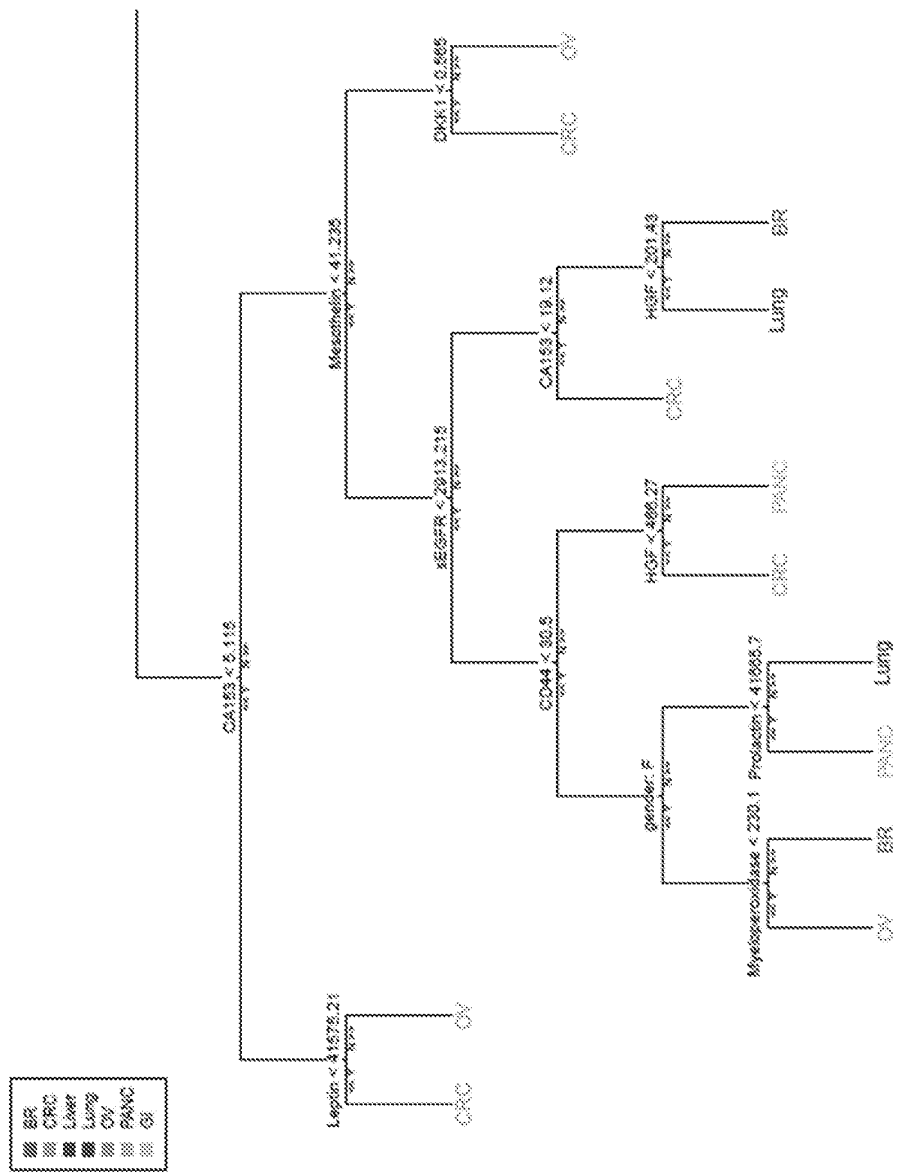
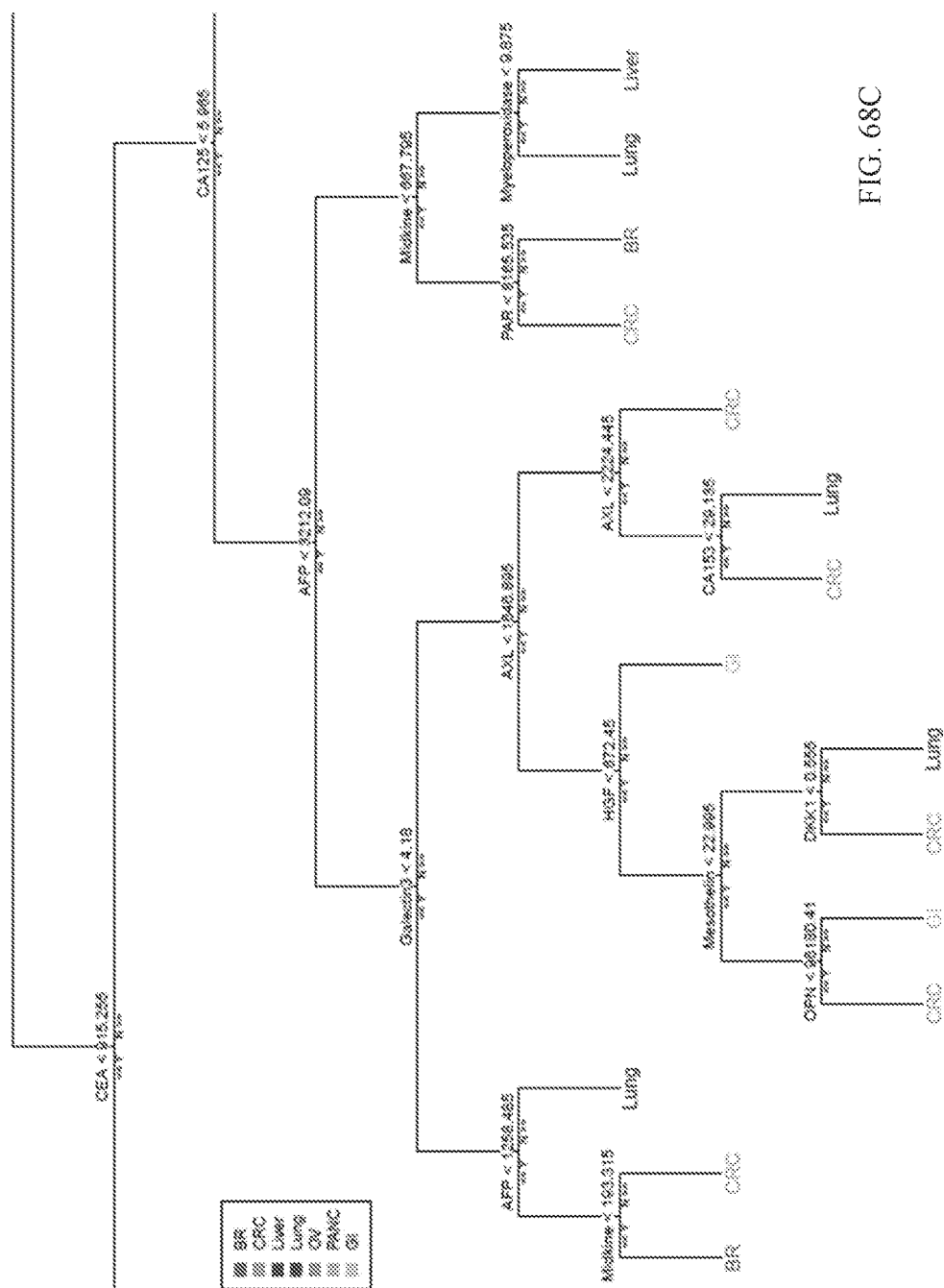


FIG. 68B



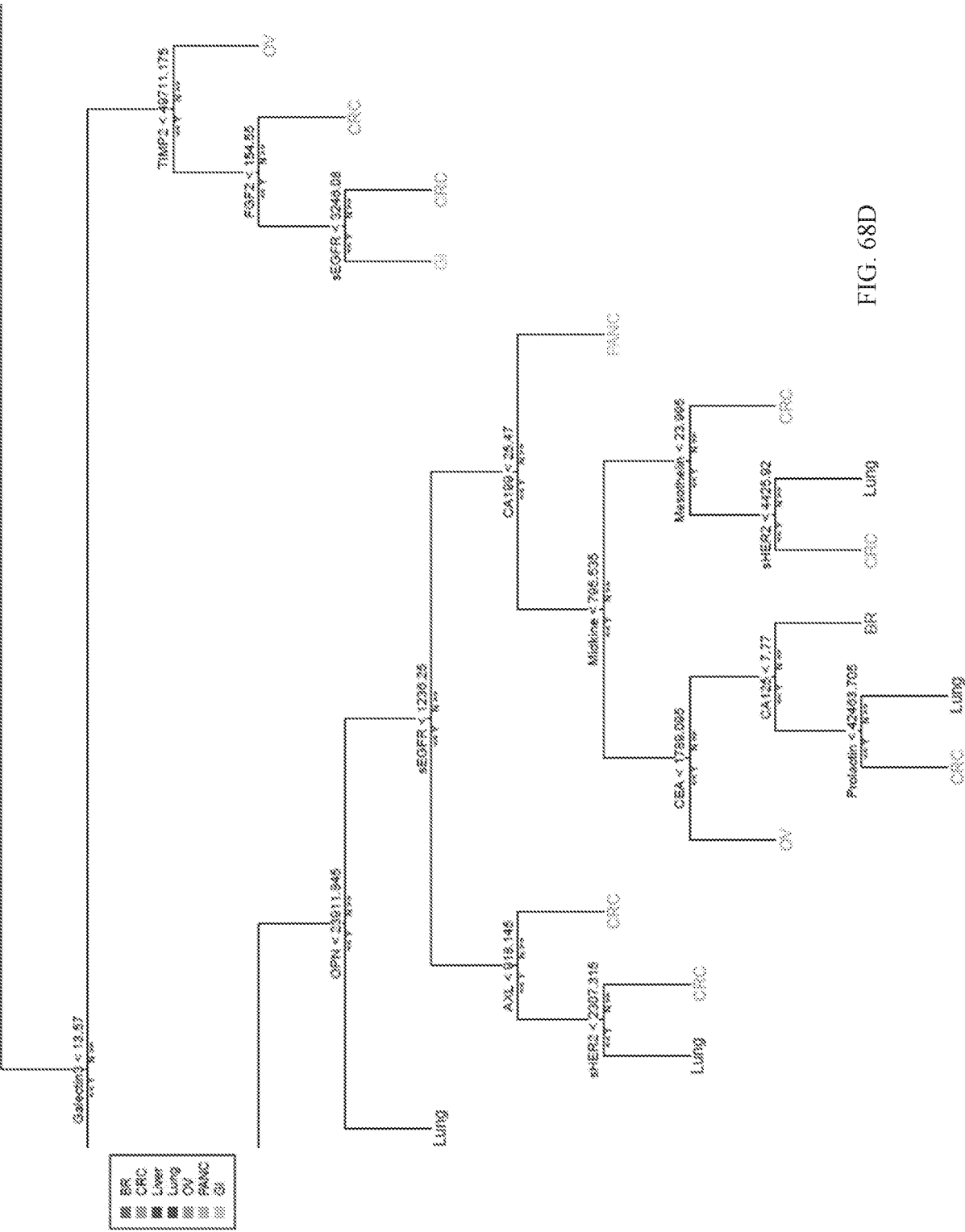


FIG. 68D

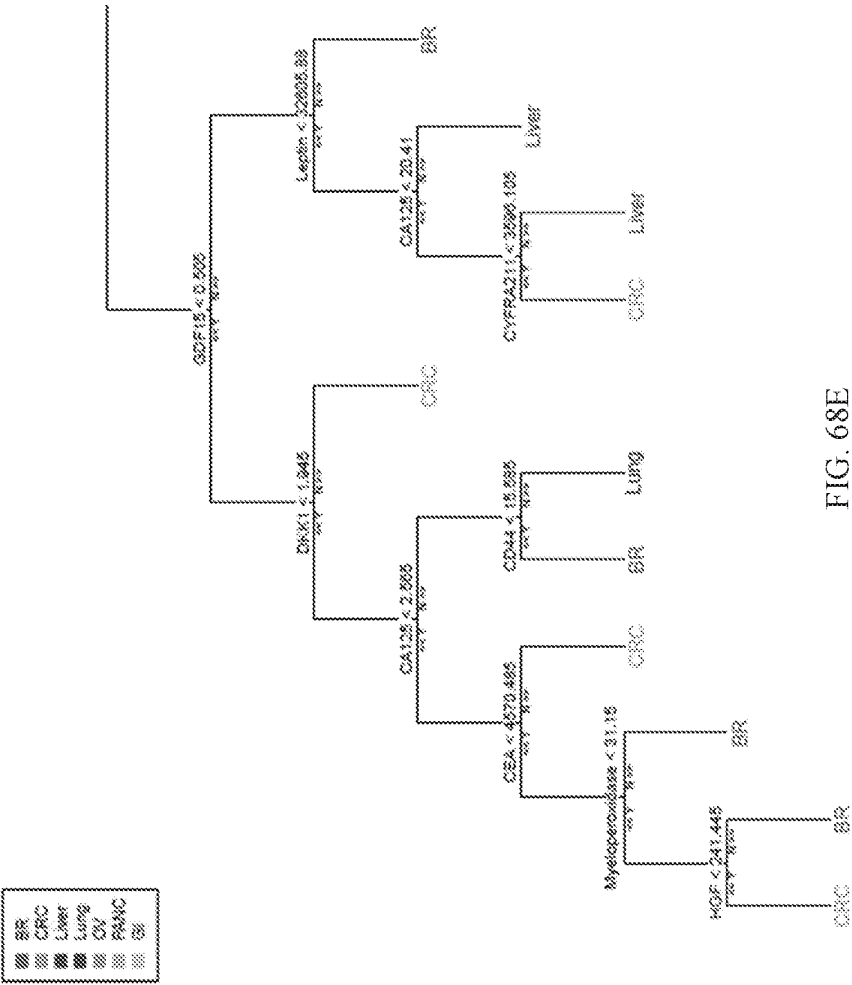


FIG. 68E

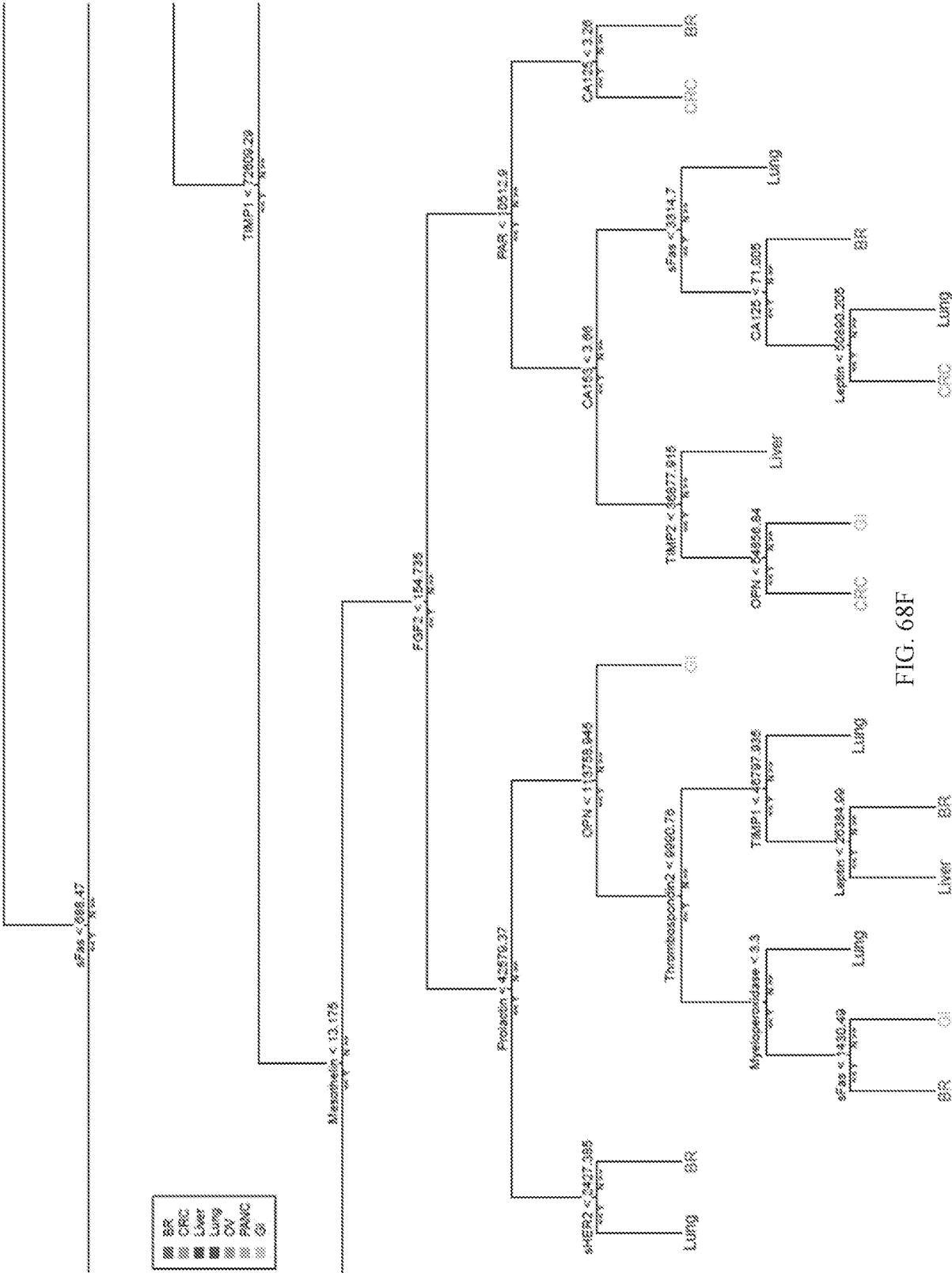


FIG. 68F

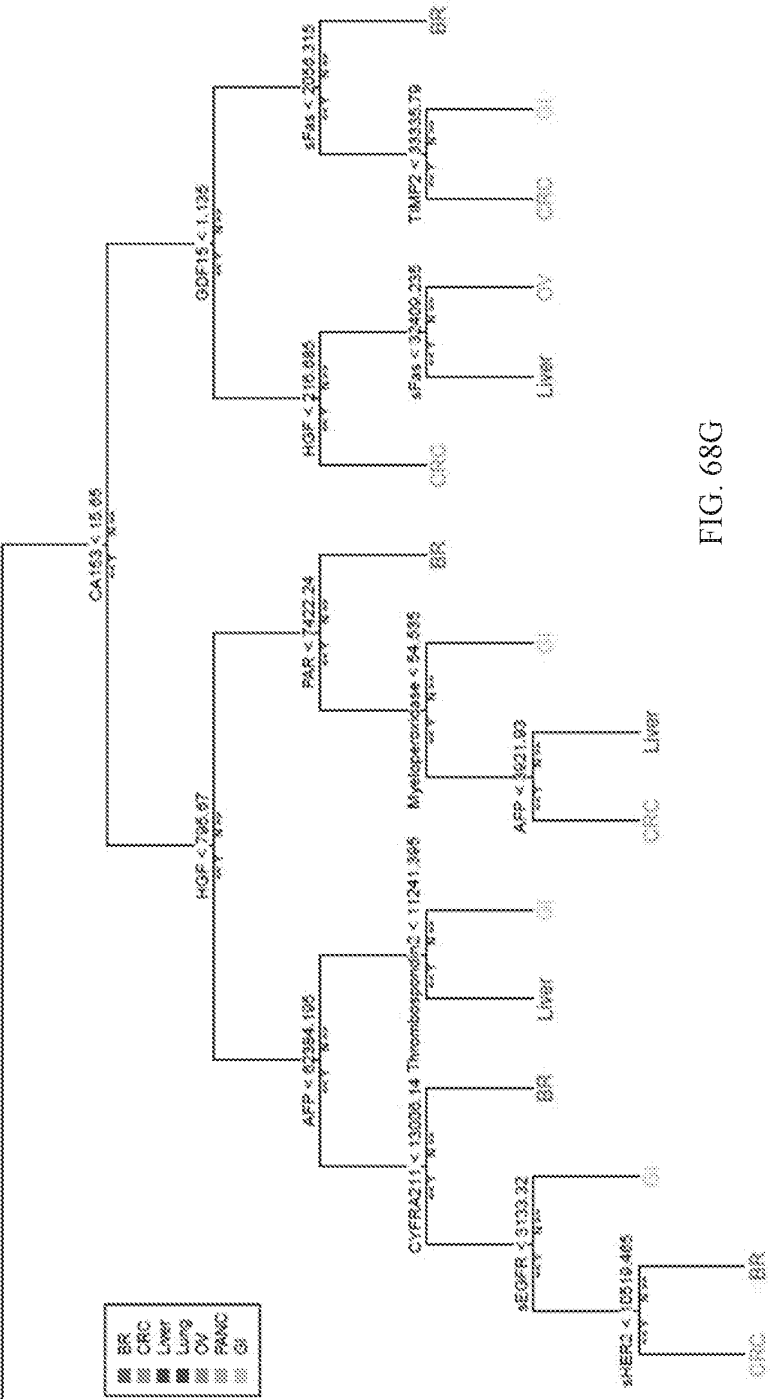


FIG. 68G

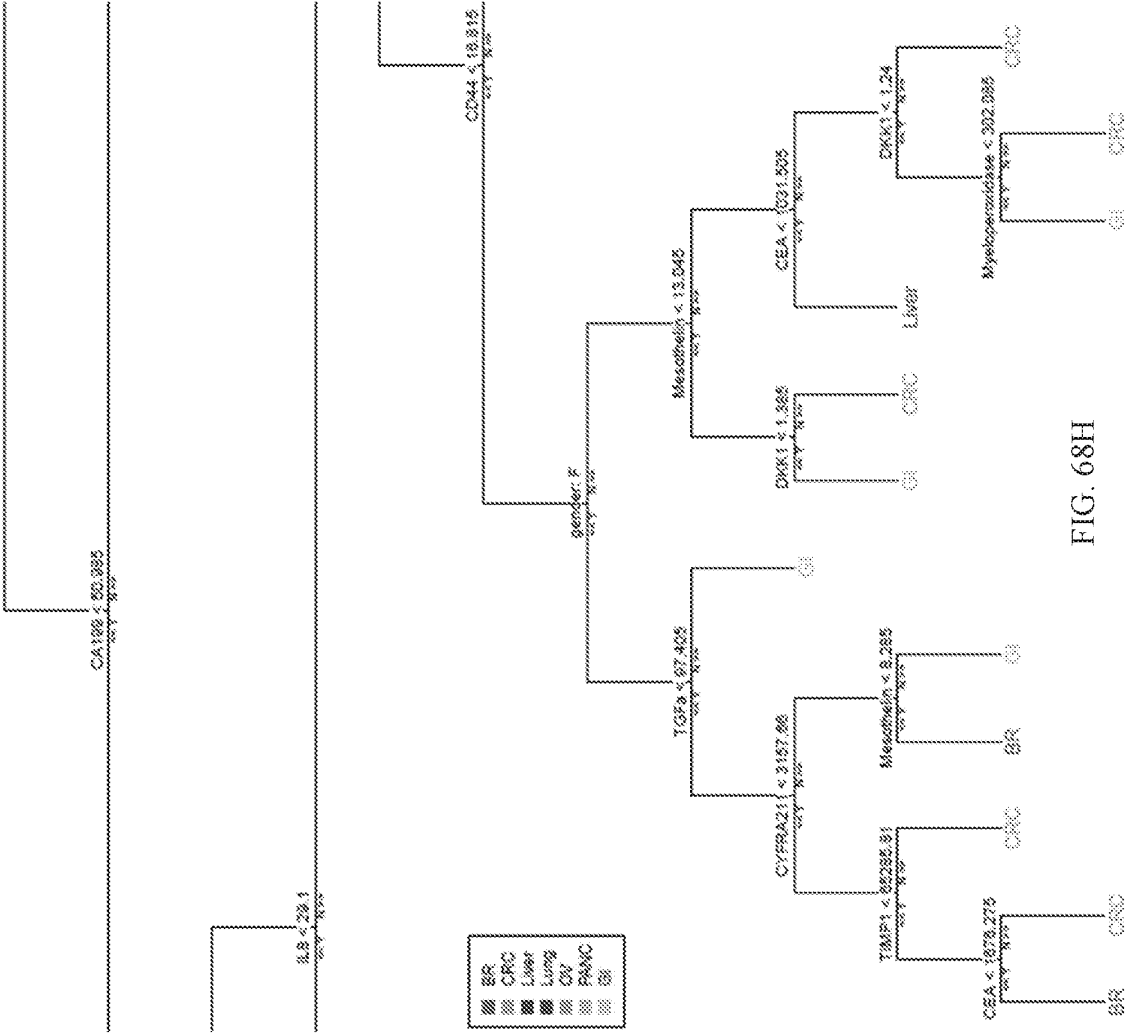
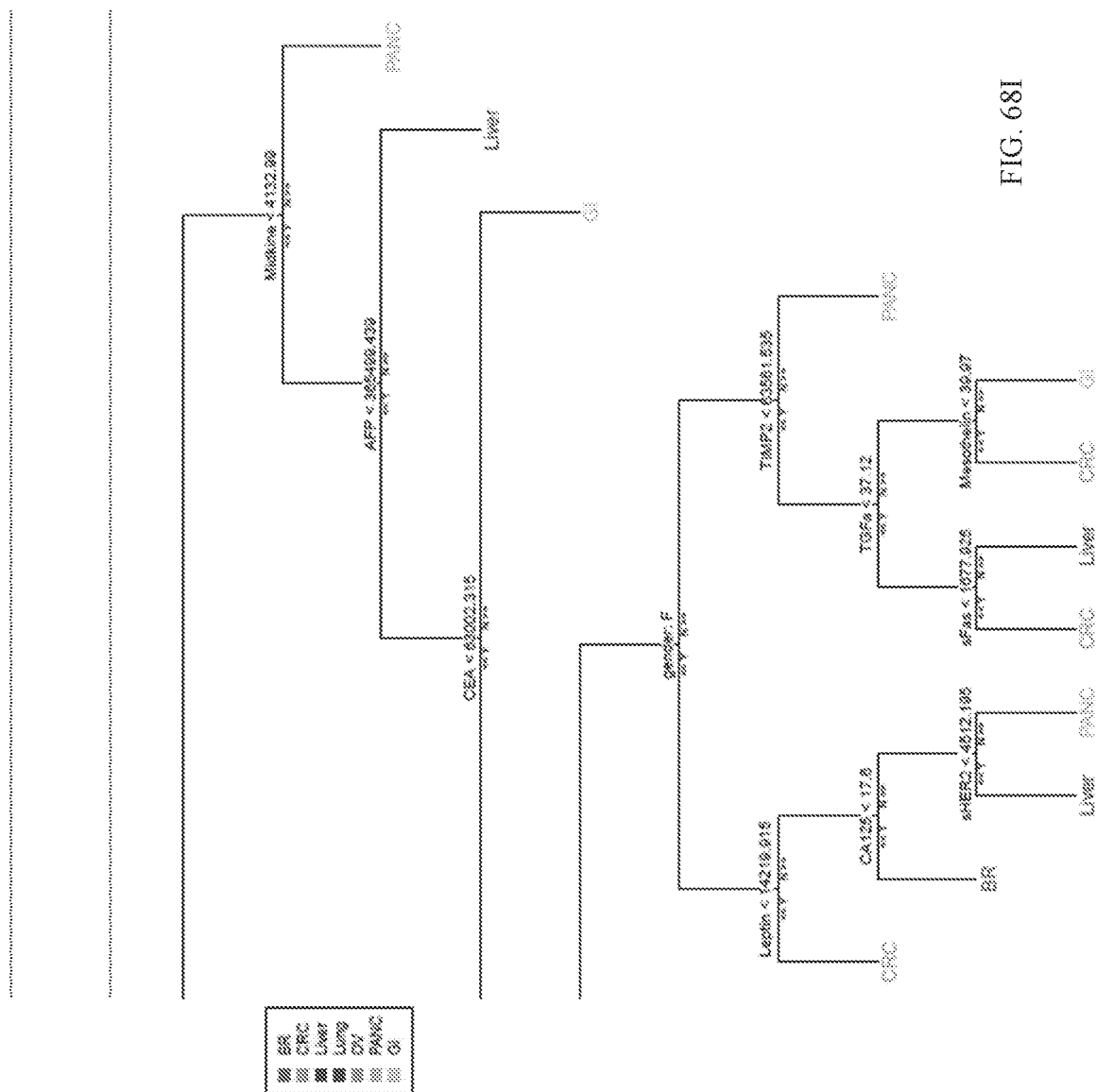
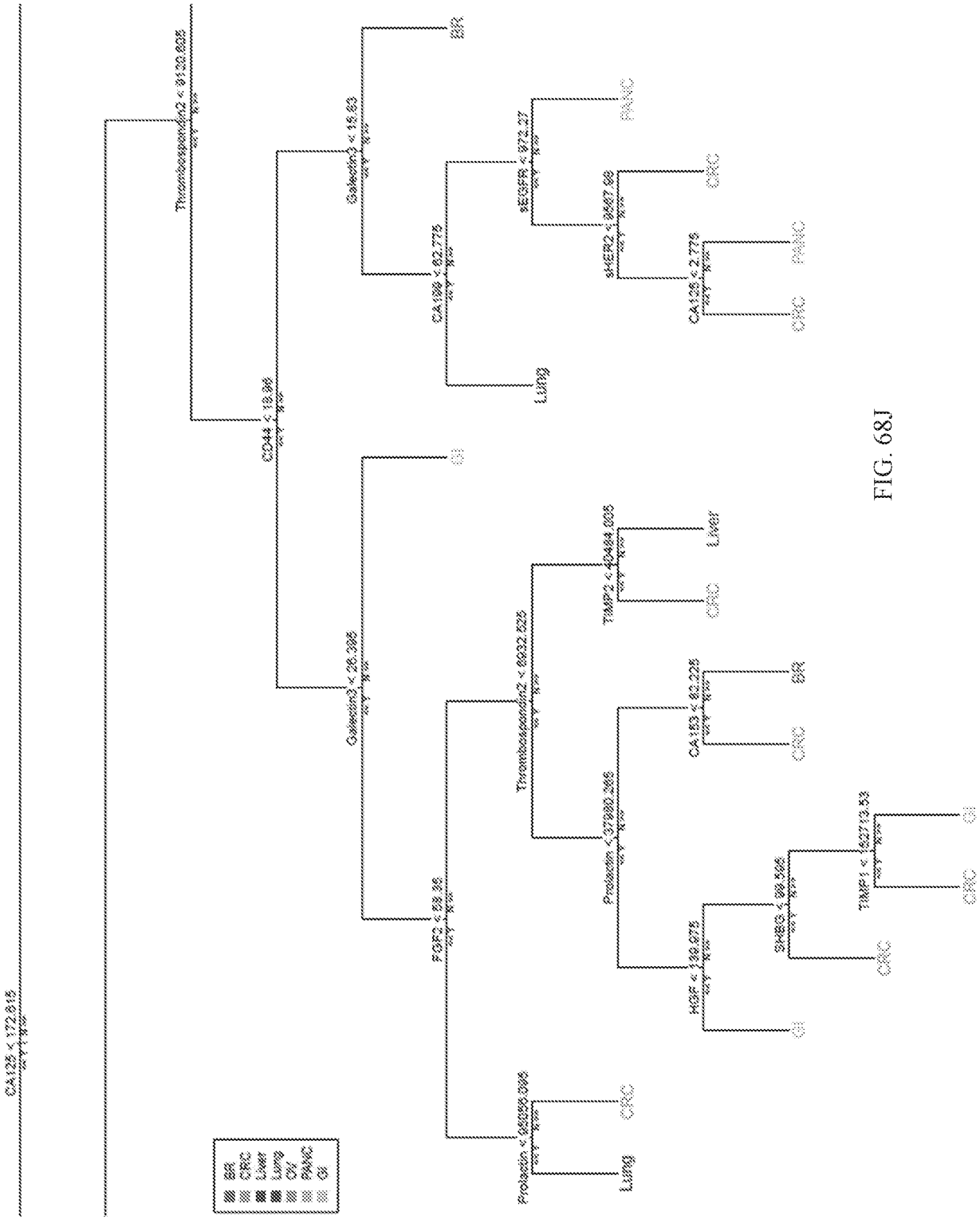


FIG. 68H





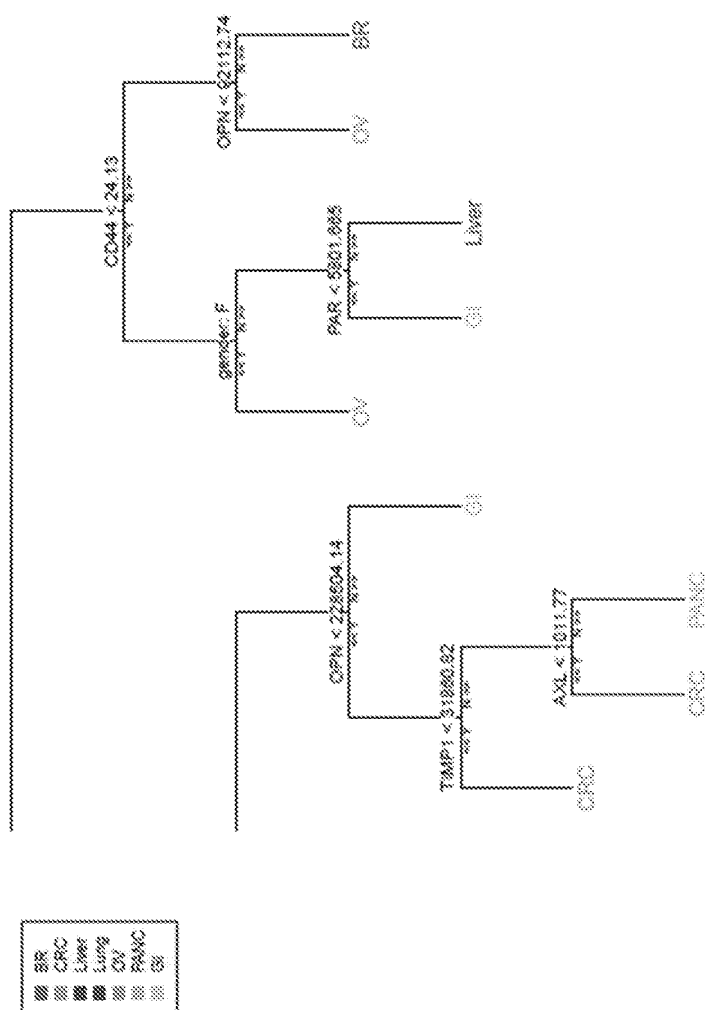


FIG. 68K

```
## R version 3.5.0 (2018-04-23)
## Platform: x86_64-w64-mingw32/x64 (64-bit)
## Running under: Windows 10 x64 (build 17134)
##
## Matrix products: default
##
## locale:
## [1] LC_COLLATE=English_United States.1252
## [2] LC_CTYPE=English_United States.1252
## [3] LC_MONETARY=English_United States.1252
## [4] LC_NUMERIC=C
## [5] LC_TIME=English_United States.1252
##
## attached base packages:
## [1] stats      graphics  grDevices  utils      datasets  methods   base
##
## other attached packages:
## [1] caret_4.0-60      ggplot2_2.2.0      lattice_0.20-35
## [4] rInTrees_1.2       DT_0.4             randomForest_4.5-14
##
## loaded via a namespace (and not attached):
## [1] magic_1.5-2        ddtalpha_1.2.4      tidyr_0.8.1
## [4] rfmisc_1.1-2       splines_3.5.0       foreach_1.4.4
## [7] prodlim_2018.04.18 Assesswhat_0.2.0  stats4_3.5.0
## [10] arules_1.6-1       DRR_0.0.2           yaml_2.2.19
## [13] rsubspace_0.93-1   apred_0.9-6         pillar_1.2.3
## [16] backports_1.1.2    glue_1.2.0          digest_0.6.18
## [19] RFF_1.8            colorspace_1.2-2    recipes_0.1.3
## [22] gtm_2.1.3          hemlucola_0.3.6     Matrix_1.2-14
## [25] plyr_1.8.4         psych_1.8.4         timeDate_2014.102
## [28] pkgconfig_2.0.1    CVST_0.2-2         broom_0.4.5
## [31] gower_0.2.5        xtable_1.5-2        scales_0.5.0
## [34] gower_0.1.2        lava_1.6.2          rttble_1.4.2
## [37] rghost_0.71.2      withr_2.1.2         rnet_7.3-12
## [40] lassoval_0.2.1     rnorm_1.5-5         survival_2.42-6
## [43] magrittr_1.5       evaluate_0.11       rime_3.2-137
## [46] MASS_7.3-50        dimRed_0.1.0        foreign_0.8-70
## [49] class_7.3-14       tools_3.5.0         data.table_1.11.4
## [52] rstring_1.3.1      kernlab_0.9-36      ransell_0.5.0
## [55] bindrepp_0.3.2     pls_2.8-0            compiler_2.6.0
## [58] RcppRoll_0.3.0     rlang_0.2.1         grid_3.5.0
## [61] iterators_1.0.9    hemlucola_1.2       xmarkdown_1.10
## [64] geometry_0.3-6     qtable_0.2.0        ModelMetrics_1.1.0
## [67] codetools_0.2-15   akima_1.4-5         reshape2_1.4.3
## [70] R6_2.2.2           lubridate_1.7.4     knitr_1.20
## [73] dplyr_0.7.6        bindr_0.1.1          xprojroot_1.3-2
## [76] stringi_1.2.3      parallel_3.5.0      Rcpp_0.12.17
## [79] xpart_4.1-13       DEoptim_1.0-8        tidyselct_0.2.4
```

FIG. 69

```

## [1] "Tumor samples: n = 626"

## [1] "Proteins: AFP, CA125, CA153, CA199, CEA, CYPRA31, EGFR, ER4, GGF, IL6, IL8, Leptin, OPN, Prolactin, sFa
s, TGFa, TIMP1, TIMP2, Galectin3, Myeloperoxidase, SHBG, SIK1, GDF15, NSE, OPN, CD44, Mesothelin, Midkine, AXL, F
AP, sB2M, sKIM2, Thrombospondin2"

## [1] "Predictions from randomForest:"

##
## Breast CRC Liver Lung Ovarian Pancreas st_esc
## 48 332 17 25 41 52 41

## 15638 rules (length=6) were extracted from the first 500 trees.

## [1] "Predictions from inTrees classifier:"

## predictions
## Breast CRC Liver Lung Ovarian Pancreas st_esc
## 57 352 19 36 45 50 57

## [1] "Accuracies from the paper:"

## Breast CRC Liver Lung Ovarian Pancreas st_esc
## 0.63 0.64 0.44 0.39 0.79 0.61 0.46

## [1] "Accuracies from inTrees classifier:"

## Breast CRC Liver Lung Ovarian Pancreas st_esc
## 0.6376812 0.8393281 0.3863636 0.4095238 0.8113208 0.7058824 0.6325641

```

FIG. 69 (continued)