Cell Reports Medicine, Volume *5*

Supplemental information

Multi-modal cell-free DNA genomic and fragmentomic patterns enhance cancer survival and recurrence analysis

Norbert Moldovan, Ymke van der Pol, Tom van den Ende, Dries Boers, Sandra Verkuijlen, Aafke Creemers, Jip Ramaker, Trang Vu, Sanne Bootsma, Kristiaan J. Lenos, Louis Vermeulen, Marieke F. Fransen, Michiel Pegtel, Idris Bahce, Hanneke van Laarhoven, and Florent Mouliere

Multi-modal cell-free DNA genomic and fragmentomic patterns enhance cancer survival and recurrence analysis.

Norbert Moldovan, Ymke van der Pol, Tom van den Ende, Dries Boers, Sandra Verkuijlen, Aafke Creemers, Jip Ramaker, Trang Vu, Sanne Bootsma, Kristiaan J. Lenos, Louis Vermeulen, Marieke F. Fransen, Michiel Pegtel, Idris Bahce, Hanneke van Laarhoven, Florent Mouliere

SUPPLEMENTARY DATA

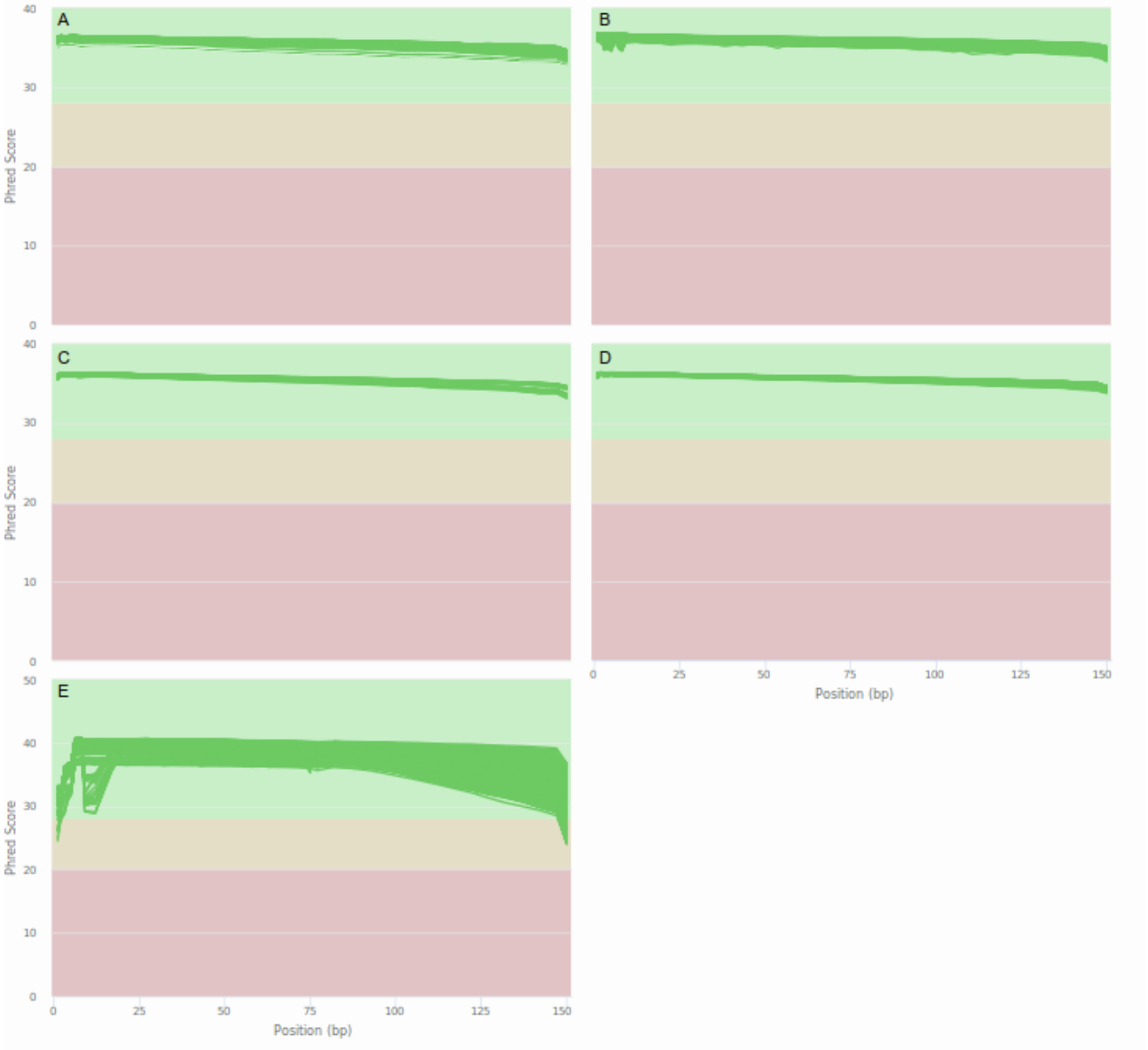

Figure S1: Average base quality (phred scores) of reads per sample after adapter trimming. Related to Figure 1. (A) sequencing batch 1, (B) batch 2, (C) batch 3, (D) batch 4, (E) batch 5. A phred score >30 represents 1 incorrect base call per 1000 bases.

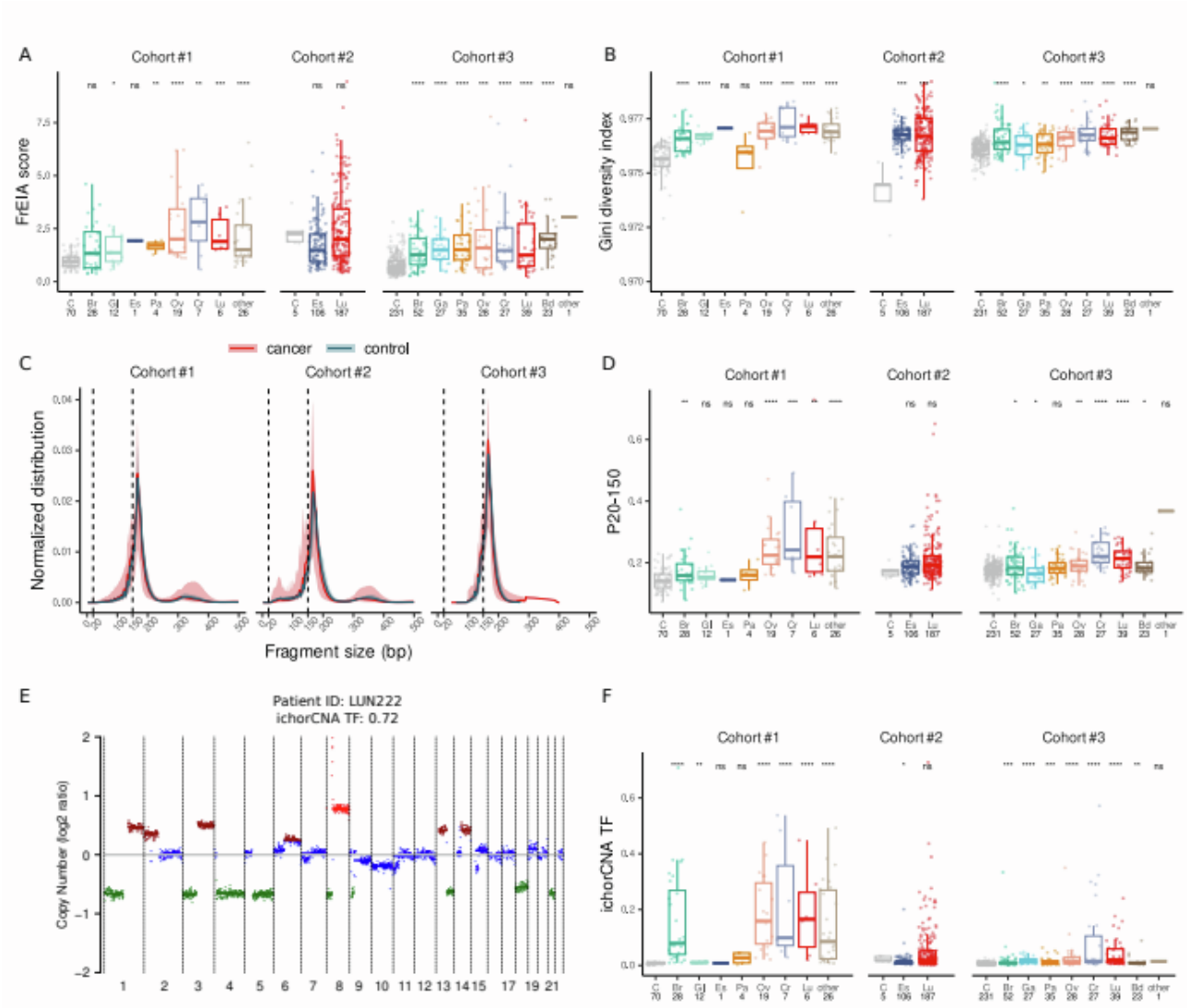

Figure S2: Measures of cfDNA biological features are altered in cancer for each cohort. Related to Figure 1. The increase in (A) the FrEIA score, (B) the Gini diversity index by cancer type in pre-treatment samples. (C) Aberrant normalized size distribution of cfDNA fragments in pre-treatment cancer samples compared to control. The vertical dashed lines outline the size interval used to calculate the P20-150 measure. The (D) P20-150 increased by cancer type in pre-treatment samples. (E) An example of the copy number alterations from the pre-treatment sample of lung cancer patient LUN222 derived by ichorCNA. Red dots: gains, blue dots: copy-neutral regions, green dots: losses. (F) The ichorCNA TF increased by cancer type in pre-treatment samples. Bd: bile duct cancer, Br: breast cancer, Cr: colorectal cancer, Es: esophageal cancer, Ga: gastric cancer, Gl: glioblastoma, Lu: lung cancer, Ov: ovarian cancer, Pa: pancreatic cancer. Numbers below the cancer type abbreviation represent the sample count. Cancer types with less than 10 samples are in the "other" category. P-values calculated using two-sided Mann-Whitney U test: ns: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.005$, ****: $p < 0.001$. When multiple hypotheses were tested, alpha values were adjusted using the Bonferroni method.

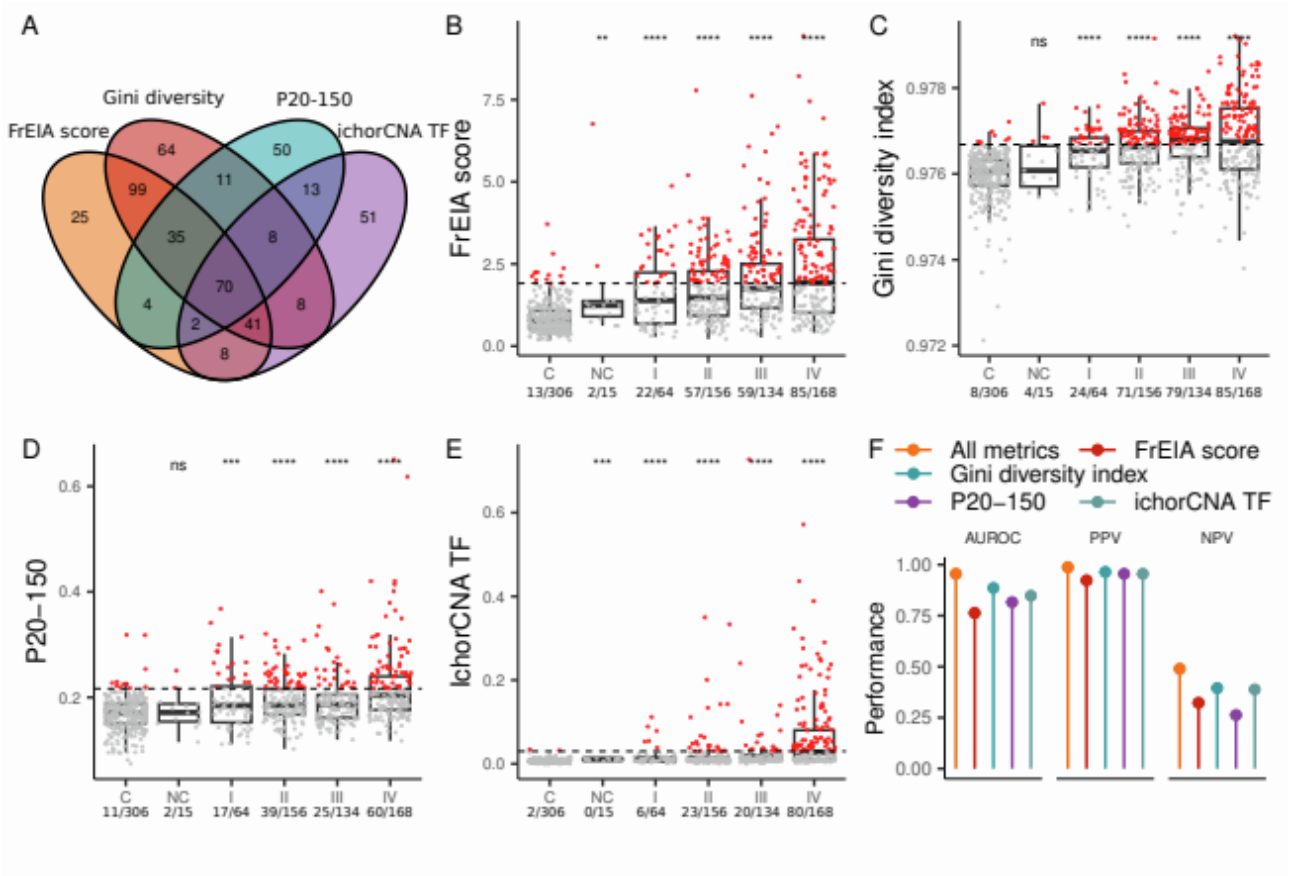

Figure S3: Testing for multiple measures of cfDNA biological features improves cancer detection. Related to Figure 4. (A) The number of pre-treatment cancer samples (total n = 628) detected by one measure or a combination of the measures. Detection rate of cancer stages of pre-treatment samples by (B) the FrEIA score, (C) the Gini diversity index, (D) the P20-150 and (E) the ichorCNA TF. Horizontal dashed line: detection threshold. Red dots represent samples higher than the detection threshold. P-values calculated using two-sided Mann-Whitney U test: ns: not significant, *: $p < 0.05$, **: $p < 0.01$, ***: $p < 0.005$, ****: $p < 0.001$. (F) Performance metrics of the logistic regression classifier based on a single measure or a combination of measures ("All metrics") tested on an independent cohort. AUROC: area under the receiver operating characteristic curve, PPV: positive predictive value, NPV: negative predictive value.

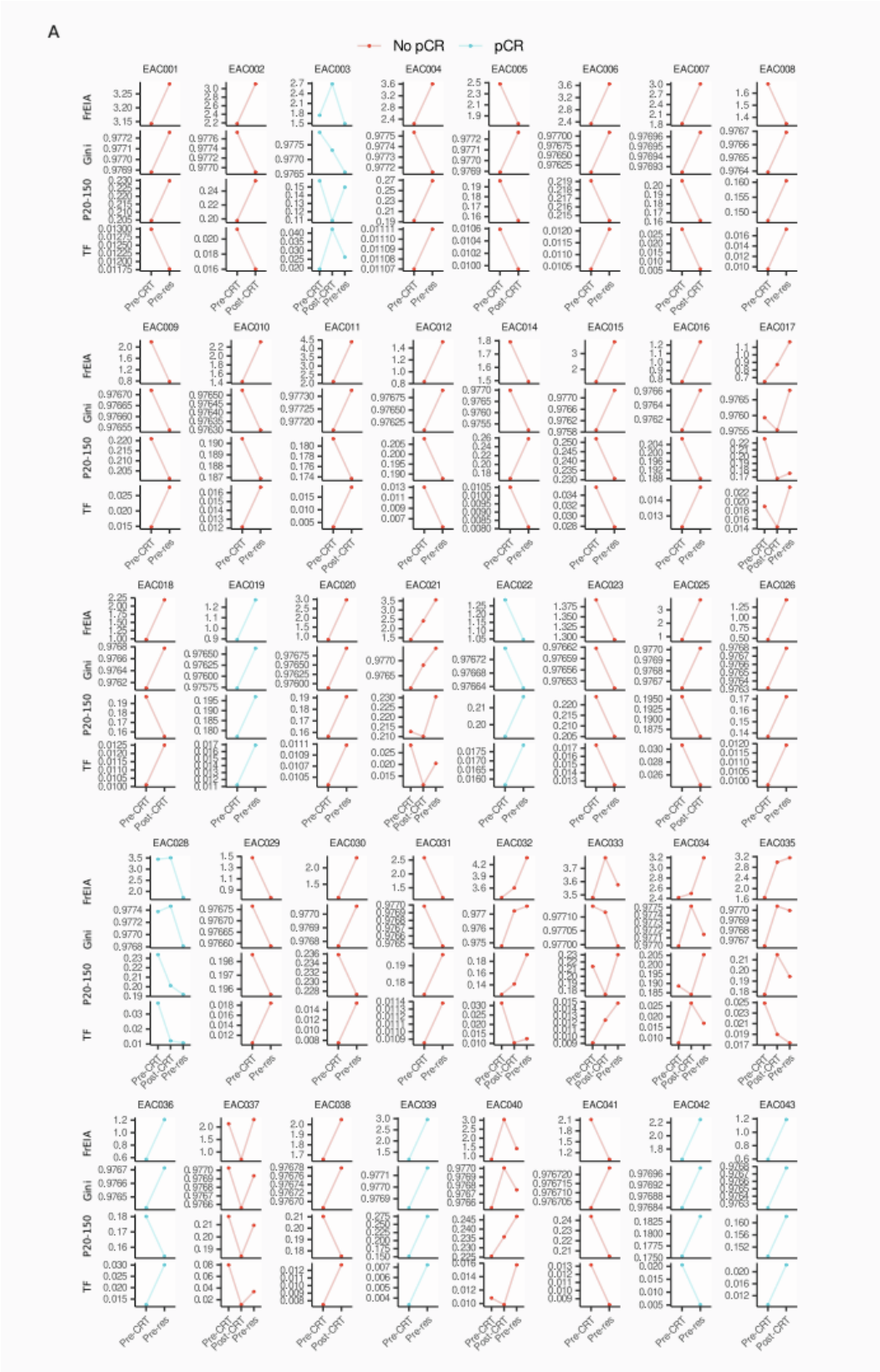

Figure S4: The change in the measures of cfDNA biological features throughout the clinical timeline of EAC patients. Related to Figure 5. pCR: complete pathological response corresponding to pT0N0 as determined by a pathologist at surgery. CRT: chemoradiotherapy, res: surgery.

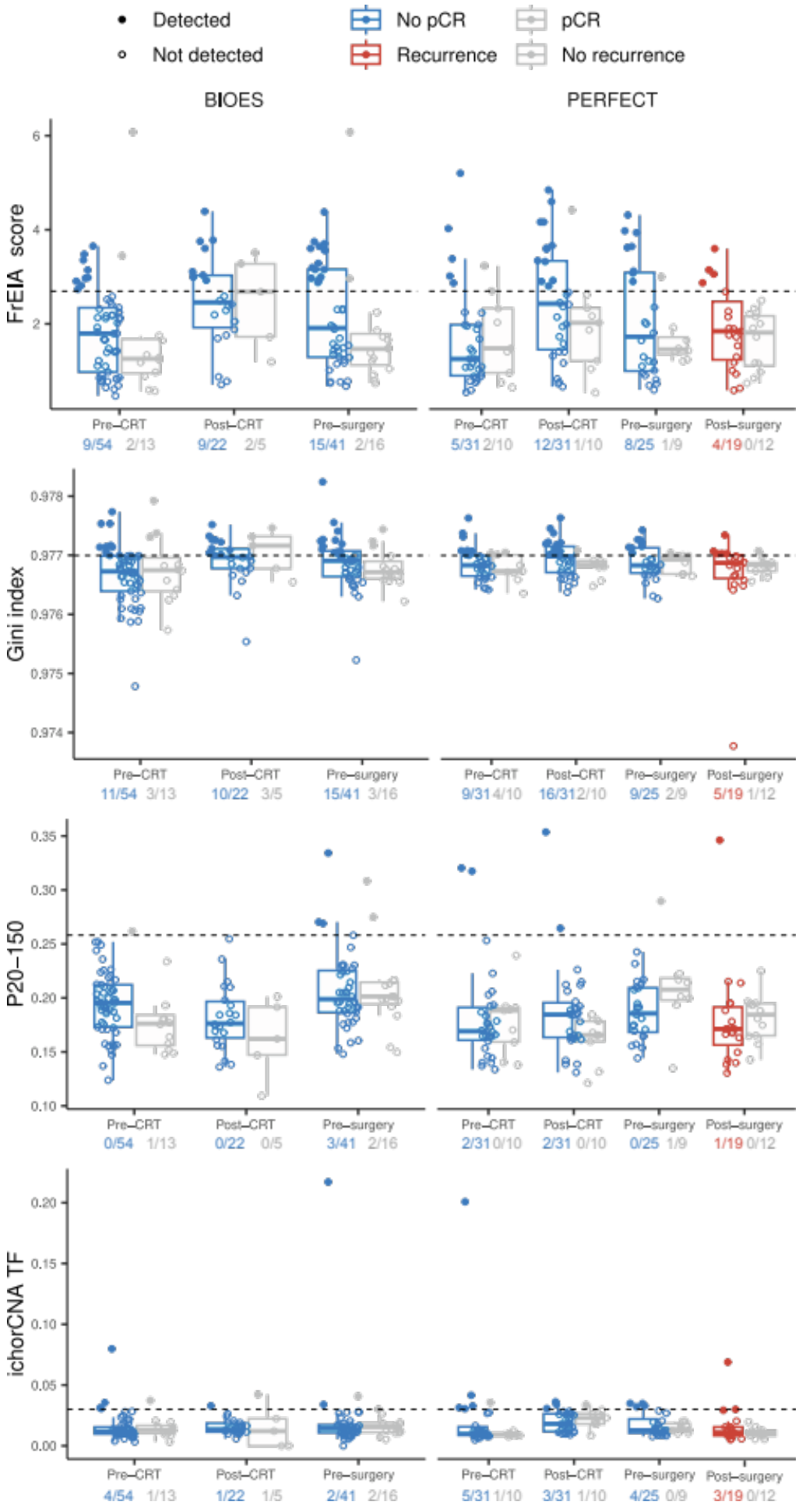
Detected
Not detected
No pCR
Recurrence
pCR
No recurrence
BIOES
PERFECT
FrEIA score
Gini index
P20−150
ichorCNA TF
Pre−CRT
Post−CRT
Pre−surgery
Post−surgery
9/54 2/13
9/22 2/5
15/41 2/16
5/31 2/10
12/31 1/10
8/25 1/9
4/19 0/12
11/54 3/13
10/22 3/5
15/41 3/16
9/31 4/10
16/31 2/10
9/25 2/9
5/19 1/12
0/54 1/13
0/22 0/5
3/41 2/16
2/31 0/10
2/31 0/10
0/25 1/9
1/19 0/12
4/54 1/13
1/22 1/5
2/41 2/16
5/31 1/10
3/31 1/10
4/25 0/9
3/19 0/12

Figure S5: Measures of cfDNA biological features of the rEAC samples. Related to Figure 5. Pre-CRT, post-CRT and pre-surgery samples are split by the patient's pathological complete response (pCR), while for pos-surgery by the recurrence status in 2 years. The horizontal dashed line represents the detection threshold (FrEIA score:) Numbers below the plots show the detected/total count of samples.

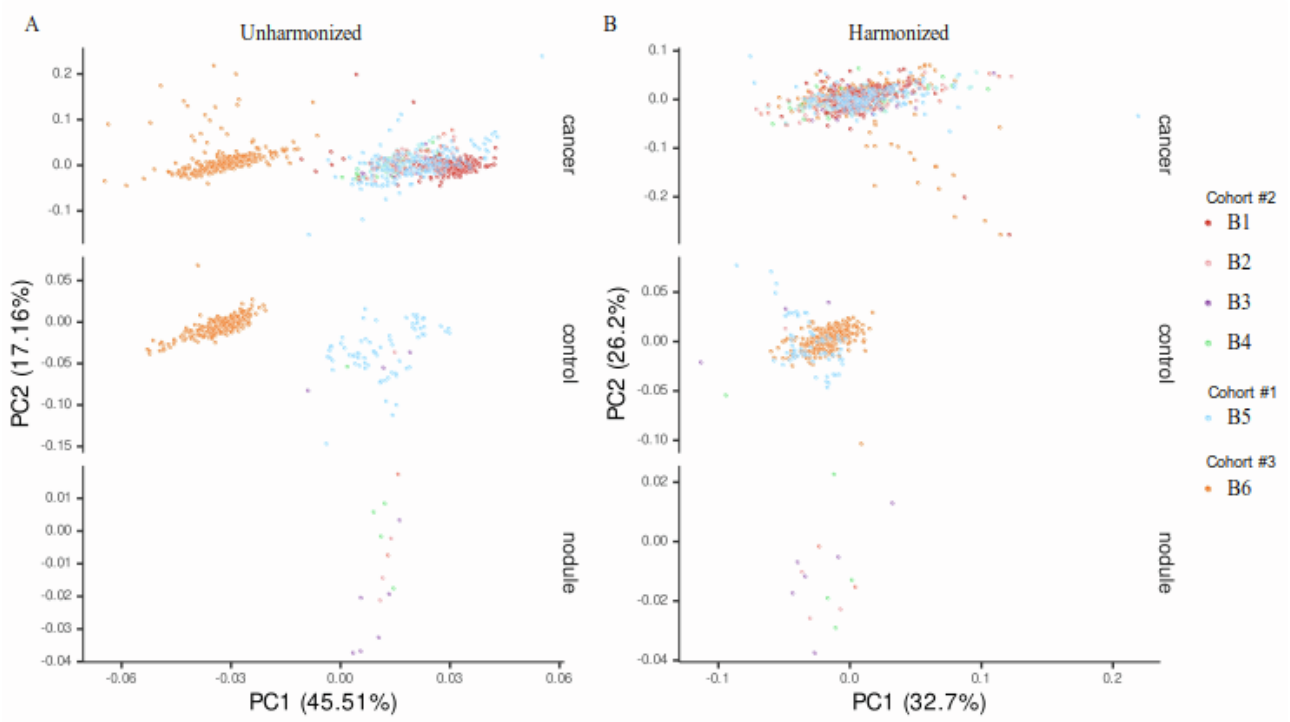

Figure S6: Harmonization of fragment end trinucleotide sequence counts. Related to Figure 1. Principal component analysis (PCA) of the 64 fragment end trinucleotide sequence counts (A) before and (B) after harmonization. The 6 batches (B1 to B6) represent 4 rounds of sequencing in for Cohort #2, and Cohort #1 and #3, which were considered as separate batches.