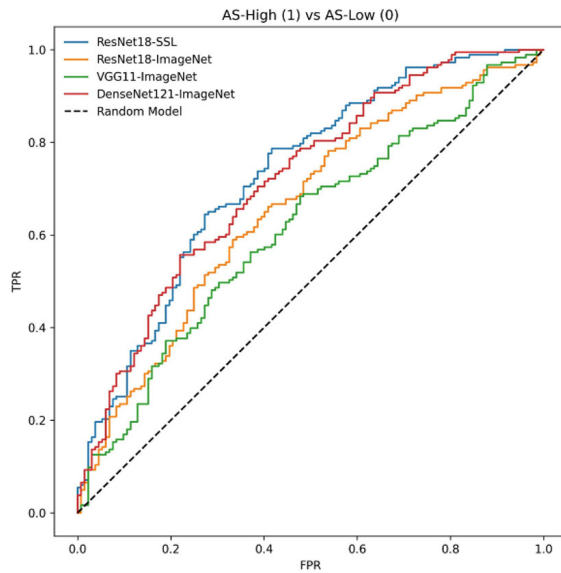
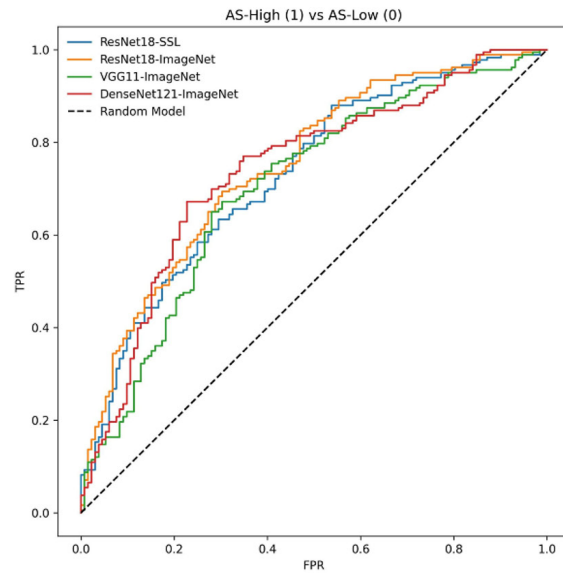


Predicting CRC chromosomal instability

A Tile-based approach



B Nucleus-based approach



Supplementary Figure 1. AUROC for tile-based and nucleus-based models. A: Comparing AUROC for each model in tile-based predictions. B: Comparing AUROC for each model in nucleus-based predictions.

Supplementary Table 1. Prediction of binarized FGA and WGD

	Accuracy (SD)	AUROC (SD)	F1 (SD)
WGD	0.673 (0.046)	0.750 (0.053)	0.632 (0.055)
FGA (threshold 0.2)	0.679 (0.038)	0.695 (0.039)	0.749 (0.032)

Supplementary Table 2. Prediction of AS with and without MSI-high cases

	MSI cases	Accuracy (SD)	AUROC (SD)	F1 (SD)	Sensitivity (SD)	Specificity (SD)
Tile-based Model	Include	0.686 (0.038)	0.742 (0.023)	0.735 (0.040)	0.754 (0.069)	0.592 (0.066)
	Exclude	0.667 (0.020)	0.635 (0.076)	0.782 (0.012)	0.906 (0.025)	0.203 (0.084)
Nuclei-based Model (M+D)	Include	0.721 (0.053)	0.750 (0.054)	0.766 (0.050)	0.792 (0.071)	0.622 (0.087)
	Exclude	0.696 (0.061)	0.731 (0.076)	0.786 (0.051)	0.856 (0.091)	0.387 (0.097)

Supplementary Table 3. Distribution of AS-H and AS-L with or without MSI-high cases

	Include MSI	Exclude MSI
AS-H	183 cases (58%)	180 cases (66%)
AS-L	132 cases (42%)	93 cases (34%)
Total	315 cases (100%)	273 cases (100%)

Predicting CRC chromosomal instability

Supplementary Table 4. Tumor cellularity and tumor cell to immune cell ratio in representative slides for correctly predicted AS-H, falsely predicted AS-H, and falsely predicted AS-L cases in tile-based models (TCGA-CM-4752, TCGA-A6-6653, and TCGA-F5-6464, respectively)

TCGA-CM-4752-01Z-00-DX1								
AS-H prediction score (%)	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of tiles	48	107	137	169	151	100	39	6
Tumor cells	8391	21317	29867	38968	37304	26068	10072	1748
Immune cells	523	634	818	789	535	291	87	5
Tumor cells ratio	174.81	199.22	218.01	230.58	247.05	260.68	258.26	291.33
Immune cell ratio	10.90	5.93	5.97	4.67	3.54	2.91	2.23	0.83
Tumor cell/immune cell	16.04	33.62	36.51	49.39	69.73	89.58	115.77	349.60
TCGA-A6-6653-01Z-00-DX1								
AS-H prediction score (%)	10-20	20-30	30-40	40-50	50-60	60-70	70-80	80-90
No of tiles	2	15	35	27	45	41	35	5
Tumor cells	447	2973	7028	5887	11953	12209	11019	1873
Immune cells	6	68	230	172	293	272	229	8
Tumor cells ratio	223.50	198.20	200.80	218.04	265.62	297.78	314.83	374.60
Immune cell ratio	3.00	4.53	6.57	6.37	6.51	6.63	6.54	1.60
Tumor cell/immune cell	74.50	43.72	30.56	34.23	40.80	44.89	48.12	234.13
TCGA-F5-6464-01Z-00-DX1								
AS-H prediction score (%)	0-10	10-20	20-30	30-40	40-50	50-60	60-70	70-80
No of tiles	15	123	234	132	52	24	3	1
Tumor cells	3005	19675	31582	17176	7122	4034	841	292
Immune cells	180	1011	1521	1052	453	151	5	0
Tumor cells ratio	200.33	159.96	134.97	130.12	136.96	168.08	280.33	292.00
Immune cell ratio	12.00	8.22	6.50	7.97	8.71	6.29	1.67	0.00
Tumor cell/immune cell	16.69	19.46	20.76	16.33	15.72	26.72	168.20	inf

Supplementary Table 5. The 10 morphologic features with a high c-index of univariate Cox proportional hazard models for overall survival (OS) and disease-free survival (DFS)

OS	c-index	DFS	c-index
Std major axis length	0.607814	Mean intensity B sd	0.688108
Mean intensity B sd	0.593903	Std minor axis length	0.682124
Mean intensity A sd	0.590499	Mean intensity gray mean	0.673897
Mean intensity gray sd	0.587243	Std area	0.65445
Std intensity A sd	0.579621	Mean intensity B mean	0.650711
Std intensity S sd	0.579029	Std perimeter	0.646971
Std intensity A mean	0.576883	Std intensity A sd	0.635752
Std intensity gray sd	0.571333	Mean intensity A mean	0.622289
Std intensity gray mean	0.566894	Mean intensity S mean	0.606582
Mean circularity	0.563046	Mean intensity gray sd	0.600598

Predicting CRC chromosomal instability

Supplementary Table 6. Morphologic features of tumor nuclei significantly correlated with copy number signatures ($P < 0.05$)

Signature	Morphology
CX1	Mean eccentricity, std intensity A mean
CX2	Mean area, mean major axis length, mean minor axis length, mean perimeter, mean intensity B mean, std area, std minor axis length, std perimeter, std intensity gray sd
CX3	Std minor axis length
CX4	Mean area, mean major axis length, mean minor axis length, mean perimeter
CX6	Mean major axis length, mean circularity, mean eccentricity, mean intensity gray mean, mean intensity S mean, mean intensity B mean, std area, std major axis length, std minor axis length, std perimeter, std intensity A mean, std intensity A sd, std intensity B mean
CX8	Mean area, mean major axis length, mean minor axis length, mean perimeter, std area, std minor axis length
CX9	Std intensity A mean
CX10	Mean intensity gray mean, mean intensity gray sd, mean intensity A sd, mean intensity B mean, mean intensity B sd, std intensity A mean, std intensity A sd
CX11	Mean solidity, mean intensity S sd, std area, std minor axis length, std perimeter, std intensity gray sd, std intensity S mean, std intensity S sd, std intensity B sd
CX16	Mean intensity B sd
CX17	Std area, std major axis length, std minor axis length, std perimeter

Supplementary Table 7. Correlation between clusters and clinical and molecular features

		Cluster 1 (%)	Cluster 2 (%)	P value
Age	< 65	48 (14.4%)	103 (33.1%)	0.56
	≥ 65	46 (14.8%)	114 (36.7%)	
Sex	Female	48 (15.4%)	107 (34.4%)	0.776
	Male	46 (14.8%)	110 (35.4%)	
pTNM	I/II	41 (14.0%)	114 (39.0%)	0.085
	III/IV	49 (16.8%)	88 (30.1%)	
pT stage	pT1-2	15 (4.8%)	42 (13.5%)	0.488
	pT3-4	79 (25.3%)	176 (56.4%)	
pN stage	pN0	44 (14.1%)	126 (40.4%)	0.074
	pN1-2	50 (16.0%)	92 (29.5%)	
pM stage	pM0	63 (24.5%)	152 (59.1%)	0.259
	pM1	16 (6.2%)	26 (10.1%)	
MSI	Stable/low	91 (29.1%)	181 (57.8%)	0.002*
	High	4 (1.3%)	37 (11.8%)	
CIMP	Negative/low	73 (25.4%)	174 (60.6%)	0.009*
	High	4 (1.4%)	36 (12.5%)	
BRAF	Wildtype	85 (27.5%)	184 (59.5%)	0.136
	Mutant	8 (2.6%)	32 (10.4%)	
KRAS	Wildtype	53 (17.2%)	122 (39.5%)	0.934
	Mutant	40 (12.9%)	94 (30.4%)	
TP53	Wildtype	24 (7.8%)	81 (26.2%)	0.047*
	Mutant	69 (22.3%)	135 (43.7%)	
AS	Low	22 (7.0%)	109 (34.8%)	< 0.001**
	High	73 (23.3%)	109 (34.8%)	

* $P < 0.05$, ** $P < 0.001$.