Appendix 1 Methods

Patients and CT Images Acquisition

The inclusion criteria for this study were as follows: (i) aged 18 years or older; (ii) histopathologically confirmed adenocarcinoma or squamous cell carcinoma; (iii) an Eastern Cooperative Oncology Group performance status of 0-1; (iv) presence of measurable tumors; (v) resectable stage IB (tumor diameter greater than 4 cm) to stage IIIB tumors. Exclusion criteria were as follows: (i) any genetic mutations in EGFR, ALK, ROS1 and MET, identified through genetic testing; (ii) presence of unresectable tumors or metastases detected during exploratory surgery; (iii) receipt of other neoadjuvant targeted therapies; (iv) refusal to participate in follow-up; (v) lack of available pre-treatment enhanced chest computed tomography (CT) scans for radiomics analysis. Our selection of CT images adhered to strict predefined criteria to ensure both consistency across all imaging platforms and relevance to the study's objectives. These criteria included the evaluation of image quality, maintaining a uniform slice thickness of 5 mm, and the verification of the absence of artifacts that could potentially skew radiomics analysis. Each image underwent a thorough review by experienced radiologists, confirming their suitability for inclusion in our study. As part of our exclusion criteria, we have omitted patients without pre-treatment enhanced chest CT scans. Specifically, for our analysis, we selected a set of pre-treatment chest CT images for each patient, which resulted in 146 sets for the training cohort, 61 sets for the validation cohort, and 36 sets for the test cohort. This study adhered to the 8th edition of the American Joint Committee on Cancer (AJCC) Tumor, Node, Metastasis (TNM) staging system.

The study employed five distinct CT scanners to acquire diagnostic images: two Philips Brilliance models (128 and 256), the Siemens Somatom Definition AS 128, the Philips IQon Spectral CT and the GE Revolution 256 CT. Each scanner was calibrated to a tube voltage of 120 kV and a tube current of 220 mA, with a detector collimation set at 64×0.625 mm. To ensure consistency across all imaging platforms, a slice thickness of 5 mm was maintained. The field of view was standardized at 20×20 cm, facilitating a comprehensive scan area. All CT scanners were used to perform contrast-enhanced scans, utilizing a high-pressure injector to administer a non-ionic iodine-based contrast agent, Iohexol (containing 300 mg/ml of iodine). The dosage of the contrast agent was 1.5 ml/kg, and it was administered at a flow rate of 3.0 ml/s. All images were imported into ITK-SNAP (http://www.itksnap.org/pmwiki/pmwiki.php) for annotation. The region of interest (ROI) was delineated by bounding boxes encompassing the entire tumor volume. Tumor annotations were independently performed by two radiologists with five years of experience. Discrepancies between observers were resolved through consultation with a senior radiologist with more than ten years of experience.

We incorporated two preprocessing approaches to augment the reliability and uniformity of the data. Initially, we applied pixel value truncation to constrain intensity values within a specified range of -1400 to 200 Hounsfield Units, determined by a window level of -600 and a width of 1600. Furthermore, we harmonized the voxel spacing across various volumes of interest, setting a uniform resolution of $1 \text{ mm} \times 1 \text{ mm}$ by employing a fixed resolution resampling technique.

Treatment Program

In this study, the primary preoperative chemotherapy for squamous cell carcinoma patients included intravenous paclitaxellike and platinum-based drugs, while adenocarcinoma patients received intravenous pemetrexed combined with platinumbased drugs. The platinum-based drugs comprised carboplatin (area under the curve of 5) or cisplatin (25 mg/m² on days 1–3). Paclitaxel regimens were either paclitaxel (135–175 mg/m²) or albumin-bound paclitaxel (260 mg/m²), along with pemetrexed at 500 mg/m². The preoperative immunotherapy involved programmed cell death protein 1 inhibitors: tislelizumab, pembrolizumab, camrelizumab, sintilimab (all at 200 mg) or toripalimab (240 mg). Patients typically received 1-3 doses of these regimens every three weeks, averaging two cycles.

Patients in this study underwent radical surgical resection for lung cancer under general anesthesia, 4 to 6 weeks after their last neoadjuvant treatment. Surgical techniques included video-assisted thoracoscopic surgery and thoracotomy. Depending on patient-specific factors and tumor characteristics, the resection varied among lobectomy, sleeve resection and total unilateral pneumonectomy.

Follow Up

In this study, we closely monitored all participating patients from their initial admission. This monitoring included routine outpatient visits and telephonic follow-ups. Throughout the follow-up period, patients underwent comprehensive physical assessments and various diagnostic tests based on medical necessity. These tests included chest-enhanced CT scans, positron emission tomography-CT when required, ultrasound examinations, tracheoscopies, magnetic resonance imaging and whole-body bone scans. For those patients whose most recent medical records in our case system were older than one month at the time of our study's cut-off date, we conducted additional telephone follow-ups. These calls were focused on gathering updated information regarding the patients' health progression and survival status. The follow-up period for our study concluded on December 6, 2023. Across all patients, the median duration of follow-up was 19 months, ranging from 5 to 39 months. Disease-free survival was defined as as the duration from the completion of lung cancer's radical resection to the first instance of recurrence, any-cause mortality, or the last follow-up. Overall survival was determined as the time span from the start of the first neoadjuvant therapy cycle until death from any cause or the final follow-up visit.



Figure S1 Flow chart for patients and CT image selection.



Figure S2 LASSO regression analysis of tumor radiomic features. (A) Intra-tumor features: LASSO coefficients and MSE with tenfold cross-validation. (B-D) Peri-tumoral regions at 2, 4, and 6 mm: LASSO coefficients and MSE with tenfold cross-validation. (E-H) Main feature coefficients in radiomic scores for intra-tumor and peri-tumoral regions (2, 4, 6 mm) represented in bar graphs.

Signature	Feature name	Non-MPR (M ± SD)	MPR (M ± SD)	P value
Habitat	original_shape_Elongation_h1	0.09±0.89	-0.04±1.05	0.64
	original_shape_Flatness_h1	0.08±0.86	-0.04±1.06	0.57
	original_shape_LeastAxisLength_h1	0.07±1.08	-0.03±0.97	0.73
	original_shape_MajorAxisLength_h1	-0.14±1.00	0.06±1.00	0.24
	original_shape_Maximum2DDiameterColumn_h1	0.04±1.08	-0.02±0.97	0.86
	original_shape_Maximum2DDiameterRow_h1	-0.15±0.97	0.07±1.01	0.23
	original_shape_Maximum2DDiameterSlice_h1	-0.13±1.07	0.06±0.97	0.27
	original_shape_Maximum3DDiameter_h1	-0.12±1.01	0.05±1.00	0.35
	original_shape_MeshVolume_h1	0.05±1.07	-0.02±0.97	0.98
	original_shape_MinorAxisLength_h1	-0.01±1.09	0.01±0.96	0.82
	original shape Sphericity h1	0.04±1.01	-0.02±1.00	0.70
	original shape SurfaceArea h1	0.04±1.17	-0.02±0.92	0.73
	original shape SurfaceVolumeBatio h1	0.05+0.94	-0.02+1.03	0.55
	original_shape_canacevolume.ht	0.05+1.07	-0.02+0.97	0.00
atra tumoral		0.05±1.07	-0.02±0.97	0.33
illa-lumorai		0.00±0.85	-0.03±1.07	0.73
	Intra_original_snape_Flatness	0.12±0.91	-0.05±1.04	0.32
	intra_original_shape_LeastAxisLength	0.06 ± 1.06	-0.03 ± 0.98	0.72
	intra_original_shape_MajorAxisLength	-0.11 ± 0.99	0.05 ± 1.01	0.48
	intra_original_shape_Maximum2DDiameterColumn	0.02±1.07	-0.01±0.97	0.89
	intra_original_shape_Maximum2DDiameterRow	-0.16±0.97	0.07±1.01	0.21
	intra_original_shape_Maximum2DDiameterSlice	-0.13±1.08	0.06 ± 0.96	0.25
	intra_original_shape_Maximum3DDiameter	-0.11±1.01	0.05±1.00	0.37
	intra_original_shape_MeshVolume	0.07±1.21	-0.03±0.89	0.84
	intra_original_shape_MinorAxisLength	-0.04±1.10	0.02±0.96	0.64
	intra_original_shape_Sphericity	0.13±0.98	-0.06±1.01	0.22
	intra_original_shape_SurfaceArea	-0.01±1.10	0.00±0.96	0.63
	intra original shape SurfaceVolumeRatio	0.04 ± 1.13	-0.02±0.94	0.83
	intra_original_shape_contact/olume	0.07+1.21	_0.03+0.89	0.84
leri2mm	peri original shape Elongation	0.15±0.80	-0.07±1.08	0.04
enzmin		0.15±0.80	-0.07±1.08	0.22
	peri_original_snape_riatness	0.21±0.97	-0.10 ± 1.00	0.08
	peri_original_shape_LeastAxisLength	0.10±1.05	-0.04±0.98	0.44
	peri_original_shape_MajorAxisLength	-0.13 ± 1.02	0.06 ± 0.99	0.30
	peri_original_shape_Maximum2DDiameterColumn	0.01 ± 1.07	-0.01 ± 0.97	0.91
	peri_original_shape_Maximum2DDiameterRow	-0.16 ± 0.97	0.07 ± 1.01	0.19
	peri_original_shape_Maximum2DDiameterSlice	-0.13 ± 1.08	0.06 ± 0.96	0.27
	peri_original_shape_Maximum3DDiameter	-0.12±1.01	0.05±1.00	0.34
	peri_original_shape_MeshVolume	0.05 ± 1.07	-0.02±0.97	0.68
	peri_original_shape_MinorAxisLength	-0.06±1.06	0.03±0.98	0.64
	peri_original_shape_Sphericity	0.09 ± 1.11	-0.04±0.95	0.48
	peri_original_shape_SurfaceArea	0.04±1.09	-0.02±0.96	0.77
	peri_original_shape_SurfaceVolumeRatio	-0.09±0.88	0.04±1.05	0.44
	peri original shape VoxelVolume	0.05 ± 1.07	-0.02±0.97	0.71
Peri4mm	peri original shape Flongation	0 18+0 79	_0.08+1.08	0.15
	peri original shape Elatness	0.22±0.97	-0.10+1.00	0.10
	pen_onginal_shape_riatness	0.22±0.97	-0.10±1.00	0.07
	peri_original_shape_LeastAxisLength	0.10±1.05	-0.05±0.98	0.40
	peri_original_shape_MajorAxisLength	-0.13±1.02	0.06±0.99	0.28
	peri_original_shape_Maximum2DDiameterColumn	0.01±1.07	-0.00 ± 0.97	0.95
	peri_original_shape_Maximum2DDiameterRow	-0.17±0.96	0.08 ± 1.01	0.17
	peri_original_shape_Maximum2DDiameterSlice	-0.13±1.07	0.06±0.96	0.28
	peri_original_shape_Maximum3DDiameter	-0.12±1.01	0.06±1.00	0.33
	peri_original_shape_MeshVolume	0.05±1.07	-0.02±0.97	0.70
	peri_original_shape_MinorAxisLength	-0.06±1.06	0.03±0.98	0.62
	peri_original_shape_Sphericity	0.10±1.12	-0.05±0.94	0.40
	peri_original_shape_SurfaceArea	0.02±1.08	-0.01±0.97	0.87
	peri_original_shape_SurfaceVolumeRatio	-0.09±0.94	0.04±1.03	0.48
	peri_original_shape_VoxelVolume	0.05±1.07	-0.02±0.97	0.71
eri6mm	peri_original_shape_Elongation	0.18±0.76	-0.08±1.09	0.14
	peri original shape Flatness	0.23±0.98	-0.10±1.00	0.06
	peri original shane LeastAvisl ength	0 11+1 05	-0.05+0.98	0.38
	peri_original_enape_ceaerniscengui	_0 1/+1 02	0.06+0.00	0.00
		-0.14±1.02	0.0010.33	0.27
		0.01±1.07	-0.00±0.97	0.96
	perl_original_shape_Maximum2DDiameterRow	-0.17±0.96	0.08±1.01	0.16
	peri_original_shape_Maximum2DDiameterSlice	-0.13±1.07	0.06±0.97	0.27
	peri_original_shape_Maximum3DDiameter	-0.12±1.01	0.06±1.00	0.31
	peri_original_shape_MeshVolume	0.04±1.06	-0.02±0.97	0.71
	peri_original_shape_MinorAxisLength	-0.07±1.06	0.03±0.98	0.59
	peri_original_shape_Sphericity	0.12±1.12	-0.06±0.94	0.33
	peri_original_shape_SurfaceArea	0.01±1.08	-0.00±0.97	0.93
	peri_original_shape SurfaceVolumeRatio	-0.09±0.96	0.04±1.02	0.45
	peri original shape VoxelVolume	0 04+1 07	-0 02+0 97	0 72
		0.0121.01	0.0220.01	0.12

Table S1 Statistical analysis of shape-related radiomic features for predicting major pathologic response to combined neoadjuvant chemoimmunotherapy in non-small cell lung cancer patients

© AME Publishing Company.

Model name	Cohort	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV
SVM	Training	0.658	0.706	0.612-0.799	0.640	0.696	0.821	0.471
	Validation	0.393	0.514	0.355-0.673	0.163	0.944	0.875	0.321
	Test	0.500	0.495	0.295–0.695	0.391	0.692	0.692	0.391
Random Forest	Training	0.610	0.806	0.735–0.877	0.470	0.913	0.922	0.442
	Validation	0.426	0.578	0.429–0.726	0.209	0.944	0.900	0.333
	Test	0.472	0.515	0.322-0.708	0.261	0.846	0.750	0.393
Extra-Trees	Training	0.836	0.880	0.823–0.937	0.890	0.717	0.873	0.750
	Validation	0.525	0.580	0.434-0.726	0.395	0.833	0.850	0.366
	Test	0.556	0.548	0.347-0.750	0.435	0.769	0.769	0.435
XGBoost	Training	0.644	0.755	0.675–0.834	0.570	0.804	0.864	0.462
	Validation	0.689	0.618	0.459–0.776	0.837	0.333	0.750	0.462
	Test	0.472	0.522	0.328-0.715	0.435	0.538	0.625	0.350
LightGBM	Training	0.781	0.844	0.778–0.909	0.800	0.739	0.870	0.630
	Validation	0.541	0.590	0.417–0.764	0.442	0.778	0.826	0.368
	Test	0.639	0.512	0.302-0.722	0.913	0.154	0.656	0.500
LR	Training	0.644	0.683	0.590-0.777	0.650	0.630	0.793	0.453
	Validation	0.672	0.687	0.550-0.825	0.674	0.667	0.829	0.462
	Test	0.472	0.508	0.310-0.707	0.217	0.923	0.833	0.400
Naive Bayes	Training	0.678	0.682	0.588–0.777	0.740	0.543	0.779	0.490
	Validation	0.639	0.683	0.549–0.818	0.558	0.833	0.889	0.441
	Test	0.472	0.512	0.316–0.708	0.217	0.923	0.833	0.400
GBM	Training	0.822	0.908	0.861–0.956	0.790	0.891	0.940	0.661
	Validation	0.656	0.627	0.471-0.784	0.674	0.611	0.806	0.440
	Test	0.556	0.512	0.298-0.725	0.565	0.538	0.684	0.412
AdaBoost	Training	0.726	0.858	0.799–0.917	0.660	0.870	0.917	0.541
	Validation	0.639	0.608	0.448-0.768	0.674	0.556	0.784	0.417
	Test	0.528	0.478	0.278-0.678	0.522	0.538	0.667	0.389

Table S2 Performance of different machine learning algorithms for the intra-tumoral signature

Model name	Cohort	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV
SVM	Training	0.747	0.784	0.693–0.874	0.730	0.783	0.880	0.571
	Validation	0.475	0.484	0.338–0.631	0.302	0.889	0.867	0.348
	Test	0.361	0.375	0.171–0.578	0.043	0.923	0.500	0.353
Random Forest	Training	0.315	0.615	0.552-0.679	0.000	1.000	0.000	0.315
	Validation	0.672	0.472	0.397–0.546	0.930	0.056	0.702	0.250
	Test	0.611	0.435	0.302-0.568	0.957	0.000	0.629	0.000
Extra-Trees	Training	0.315	0.538	0.496–0.581	0.000	1.000	0.000	0.315
	Validation	0.656	0.465	0.427-0.504	0.930	0.000	0.690	0.000
	Test	0.361	0.517	0.430-0.603	0.000	1.000	0.000	0.361
XGBoost	Training	0.596	0.697	0.603–0.792	0.520	0.761	0.825	0.422
	Validation	0.295	0.534	0.385–0.682	0.000	1.000	0.000	0.295
	Test	0.500	0.542	0.367–0.717	0.217	1.000	1.000	0.419
LightGBM	Training	0.527	0.744	0.665–0.823	0.320	0.978	0.970	0.398
	Validation	0.328	0.547	0.401-0.692	0.047	1.000	1.000	0.305
	Test	0.361	0.520	0.333–0.707	0.000	1.000	0.000	0.361
LR	Training	0.767	0.723	0.628-0.817	0.900	0.478	0.789	0.687
	Validation	0.475	0.482	0.335–0.629	0.256	1.000	1.000	0.360
	Test	0.444	0.391	0.200-0.583	0.217	0.846	0.714	0.379
Naive Bayes	Training	0.671	0.734	0.647–0.820	0.630	0.761	0.851	0.486
	Validation	0.459	0.463	0.306-0.619	0.326	0.778	0.778	0.326
	Test	0.444	0.480	0.281–0.679	0.174	0.923	0.800	0.387
GBM	Training	0.890	0.939	0.903–0.975	0.910	0.848	0.929	0.812
	Validation	0.475	0.522	0.364–0.680	0.372	0.722	0.762	0.325
	Test	0.611	0.348	0.156–0.539	0.913	0.077	0.636	0.333
AdaBoost	Training	0.795	0.891	0.839-0.942	0.790	0.804	0.898	0.638
	Validation	0.508	0.618	0.453–0.782	0.372	0.833	0.842	0.357
	Test	0.389	0.396	0.202-0.590	0.130	0.846	0.600	0.355

Table S3 Performance of different machine learning algorithms for the Peri2mm signature

Model name	Cohort	Accuracy	AUC	95% Cl	Sensitivity	Specificity	PPV	NPV
SVM	Training	0.760	0.731	0.636–0.826	0.830	0.609	0.822	0.622
	Validation	0.459	0.487	0.324–0.650	0.349	0.722	0.750	0.317
	Test	0.417	0.468	0.267–0.669	0.130	0.923	0.750	0.375
Random Forest	Training	0.315	0.706	0.616-0.796	0.000	1.000	0.000	0.315
	Validation	0.541	0.552	0.401-0.704	0.535	0.556	0.742	0.333
	Test	0.528	0.532	0.332-0.731	0.435	0.692	0.714	0.409
Extra-Trees	Training	0.329	0.628	0.544–0.712	0.020	1.000	1.000	0.319
	Validation	0.557	0.526	0.387-0.666	0.674	0.278	0.690	0.263
	Test	0.361	0.371	0.204–0.538	0.000	1.000	0.000	0.361
XGBoost	Training	0.315	0.659	0.568–0.750	0.000	1.000	0.000	0.315
	Validation	0.443	0.482	0.322-0.642	0.442	0.444	0.655	0.250
	Test	0.361	0.540	0.362-0.718	0.000	1.000	0.000	0.361
LightGBM	Training	0.699	0.770	0.689–0.851	0.710	0.674	0.826	0.517
	Validation	0.295	0.435	0.285-0.585	0.000	1.000	0.000	0.295
	Test	0.500	0.381	0.185–0.577	0.696	0.154	0.593	0.222
LR	Training	0.568	0.699	0.611–0.787	0.430	0.870	0.878	0.412
	Validation	0.459	0.461	0.308–0.614	0.302	0.833	0.812	0.333
	Test	0.417	0.472	0.272-0.671	0.130	0.923	0.750	0.375
Naive Bayes	Training	0.582	0.688	0.599–0.778	0.450	0.870	0.882	0.421
	Validation	0.721	0.446	0.277-0.614	0.953	0.167	0.732	0.600
	Test	0.444	0.535	0.336-0.734	0.130	1.000	1.000	0.394
GBM	Training	0.863	0.947	0.912-0.982	0.820	0.957	0.976	0.710
	Validation	0.393	0.445	0.291-0.600	0.140	1.000	1.000	0.327
	Test	0.639	0.478	0.271-0.686	0.913	0.154	0.656	0.500
AdaBoost	Training	0.733	0.853	0.793–0.912	0.670	0.870	0.918	0.548
	Validation	0.443	0.532	0.368–0.696	0.256	0.889	0.846	0.333
	Test	0.361	0.445	0.245-0.644	0.130	0.769	0.500	0.333

Table S4 Performance of different machine learning algorithms for the Peri4mm signature

Model name	Cohort	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV
SVM	Training	0.760	0.717	0.616–0.819	0.850	0.565	0.810	0.634
	Validation	0.443	0.465	0.304–0.626	0.302	0.778	0.765	0.318
	Test	0.361	0.445	0.247-0.642	0.000	1.000	0.000	0.361
Random Forest	Training	0.315	0.612	0.540–0.684	0.000	1.000	0.000	0.315
	Validation	0.656	0.472	0.379–0.564	0.907	0.056	0.696	0.200
	Test	0.361	0.528	0.413-0.644	0.000	1.000	0.000	0.361
Extra-Trees	Training	0.747	0.730	0.650-0.810	0.850	0.522	0.794	0.615
	Validation	0.738	0.554	0.421-0.688	0.977	0.167	0.737	0.750
	Test	0.583	0.609	0.424–0.793	0.783	0.231	0.643	0.375
XGBoost	Training	0.315	0.630	0.548–0.711	0.000	1.000	0.000	0.315
	Validation	0.672	0.488	0.372-0.605	0.930	0.056	0.702	0.250
	Test	0.361	0.567	0.392-0.741	0.000	1.000	0.000	0.361
LightGBM	Training	0.644	0.663	0.567–0.758	0.660	0.609	0.786	0.452
	Validation	0.590	0.413	0.265-0.561	0.791	0.111	0.680	0.182
	Test	0.361	0.458	0.245-0.672	0.217	0.615	0.500	0.308
LR	Training	0.712	0.645	0.541-0.749	0.800	0.522	0.784	0.545
	Validation	0.672	0.422	0.257-0.588	0.907	0.111	0.709	0.333
	Test	0.444	0.498	0.301-0.696	0.174	0.923	0.800	0.387
Naive Bayes	Training	0.699	0.640	0.539–0.742	0.820	0.435	0.759	0.526
	Validation	0.492	0.490	0.326-0.654	0.419	0.667	0.750	0.324
	Test	0.472	0.468	0.269–0.667	0.304	0.769	0.700	0.385
GBM	Training	0.795	0.837	0.762-0.912	0.810	0.761	0.880	0.648
	Validation	0.705	0.461	0.309–0.613	0.977	0.056	0.712	0.500
	Test	0.444	0.492	0.293–0.690	0.174	0.923	0.800	0.387
AdaBoost	Training	0.705	0.846	0.779–0.913	0.640	0.848	0.901	0.520
	Validation	0.705	0.483	0.333–0.632	0.977	0.056	0.712	0.500
	Test	0.472	0.552	0.362-0.742	0.174	1.000	1.000	0.406

Table S5 Performance of different machine learning algorithms for the Peri6mm signature

Model name	Cohort	Accuracy	AUC	95% CI	Sensitivity	Specificity	PPV	NPV
SVM	Training	0.623	0.743	0.659–0.828	0.510	0.870	0.895	0.449
	Validation	0.738	0.820	0.714-0.927	0.698	0.833	0.909	0.536
	Test	0.722	0.635	0.422-0.848	0.913	0.385	0.724	0.714
Random Forest	Training	0.856	0.920	0.877-0.963	0.890	0.783	0.899	0.766
	Validation	0.738	0.809	0.698–0.921	0.698	0.833	0.909	0.536
	Test	0.722	0.699	0.508–0.890	0.826	0.538	0.760	0.636
Extra-Trees	Training	0.719	0.840	0.773–0.907	0.650	0.870	0.915	0.533
	Validation	0.689	0.724	0.579–0.869	0.628	0.833	0.900	0.484
	Test	0.639	0.669	0.470-0.867	0.609	0.692	0.778	0.500
XGBoost	Training	0.712	0.826	0.754–0.897	0.660	0.826	0.892	0.528
	Validation	0.574	0.822	0.716-0.929	0.395	1.000	1.000	0.409
	Test	0.639	0.769	0.614–0.925	0.522	0.846	0.857	0.500
LightGBM	Training	0.774	0.824	0.749–0.900	0.780	0.761	0.876	0.614
	Validation	0.623	0.716	0.574–0.859	0.558	0.778	0.857	0.424
	Test	0.667	0.686	0.499–0.873	0.696	0.615	0.762	0.533
LR	Training	0.671	0.665	0.571-0.759	0.720	0.565	0.783	0.481
	Validation	0.574	0.683	0.542-0.825	0.442	0.889	0.905	0.400
	Test	0.528	0.589	0.383–0.794	0.348	0.846	0.800	0.423
Naive Bayes	Training	0.603	0.674	0.582-0.766	0.540	0.739	0.818	0.425
	Validation	0.689	0.664	0.506-0.822	0.744	0.556	0.800	0.476
	Test	0.722	0.645	0.442-0.849	0.913	0.385	0.724	0.714
GBM	Training	0.815	0.888	0.834–0.943	0.820	0.804	0.901	0.673
	Validation	0.754	0.799	0.682-0.917	0.744	0.778	0.889	0.560
	Test	0.639	0.687	0.500-0.874	0.609	0.692	0.778	0.500
AdaBoost	Training	0.774	0.823	0.753–0.892	0.820	0.674	0.845	0.633
	Validation	0.721	0.722	0.572-0.872	0.814	0.500	0.795	0.529
	Test	0.639	0.749	0.589–0.910	0.522	0.846	0.857	0.500

Table S6 Performance of different machine learning algorithms for the habitat signature



Figure S3 Comparative receiver operating characteristic curves for signature prediction accuracy across training (A), validation (B) and test cohorts (C).