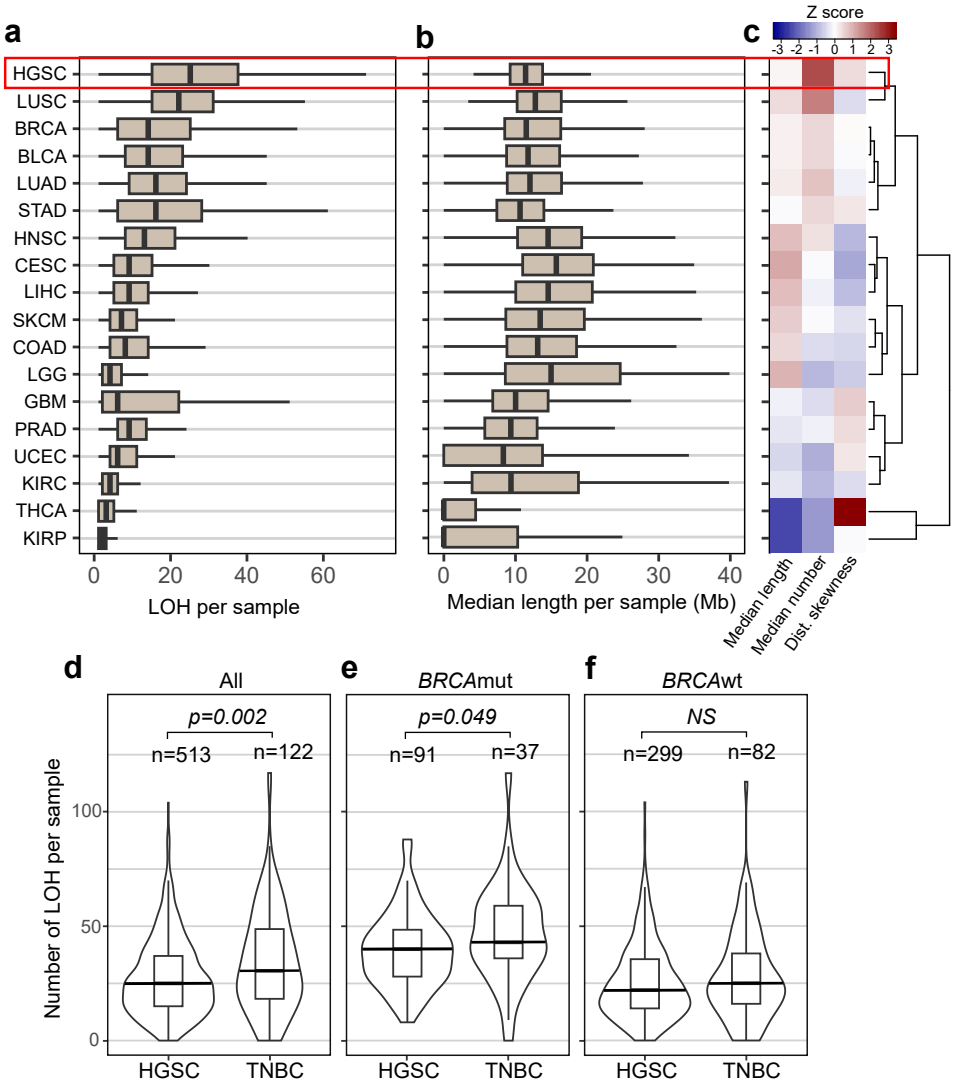
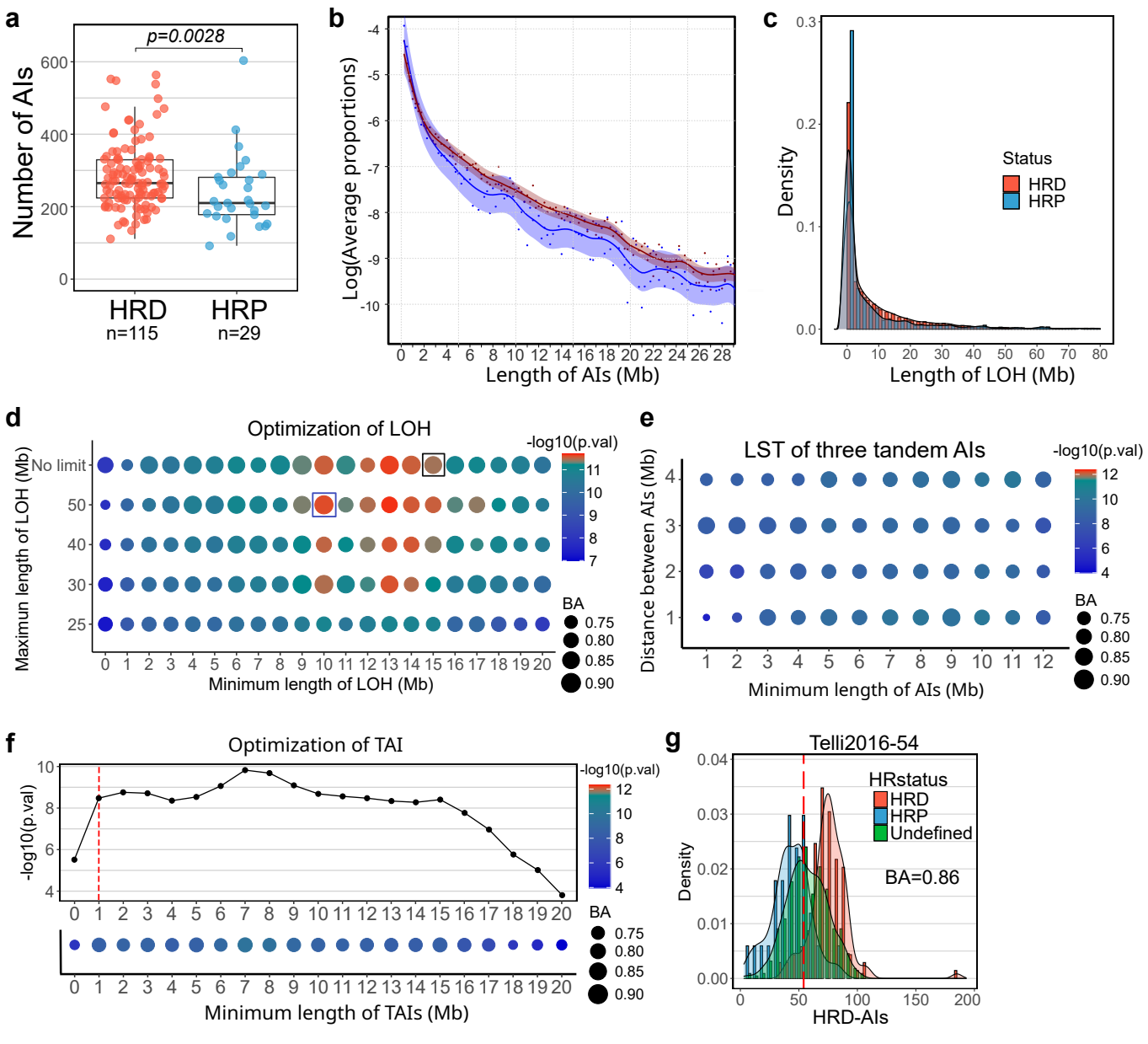
Supplementary Figure 1



Supplementary Figure 1. Pan-cancer characterization of LOH shows unique patterns in HGSC. **a** Box plots showing the number of LOH events larger than 3Mb and smaller than 50Mb in the different cancer types. **b** Box plots showing the median length of LOH events (longer than 3Mb and smaller than 50Mb) in the cancer types. The black vertical lines represent the sample medians, the boxes extend from first to third quartile and whiskers indicate the values at 1.5 times the interquartile range, no outliers are shown. **c** Hierarchical clustering of the cancer types using the median length, median number of LOH events per sample, and the skewness of the distribution of LOH length. **d** Violin- and box plots representing the number of LOH events in all HGSC samples as compared to TNBC (U test, p=0.005). Long horizontal lines represent the medians further notations are similar as in panel a .**e** Violin and box plots representing the number of LOH events in *BRCA*mut HGSC samples as compared to TNBC (U test, p=0.021). **f** Violin and box plots representing the number of LOH events in *BRCA*wt HGSC samples as compared to the TNBC (U test, NS).

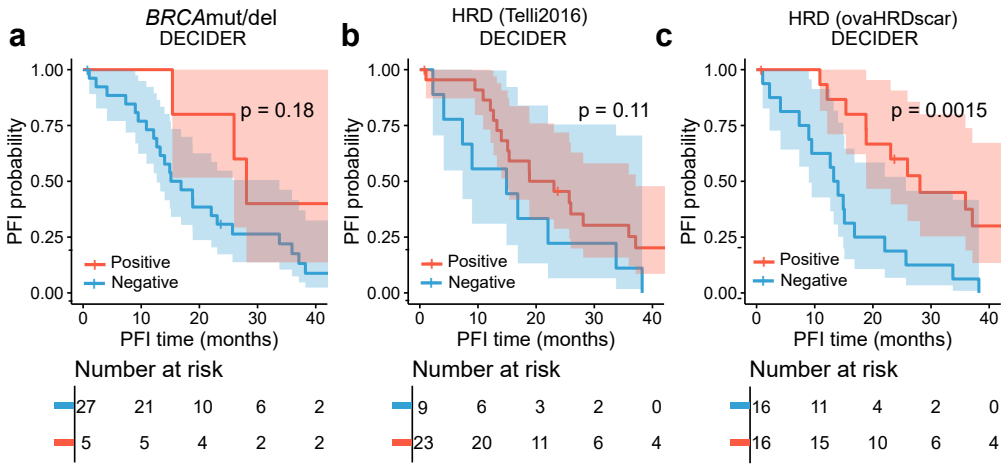
Supplementary Figure 2



	ovaHRDscar +	ovaHRDscar -	Telli2016-54 +	Telli2016-54 -
SigMA +	116	26	121	21
SigMA -	29	83	37	75

Supplementary Figure 2. Descriptive statistics of HRD and HRP in HGSC. a Box plot showing the number of allelic imbalances in HRD (red) and HRP (blue) samples in OVA-TCGA. The black horizontal lines represent the sample medians, the boxes extend from first to third quartile and whiskers indicate the values at 1.5 times the interquartile range. **b** The average proportion of segments (AIs) equal or greater than the given length in HRD (red) and HRP (blue) samples, the blue and red lines correspond to smoothing using cubic splines, confidence intervals are shown in shaded colors. **c** Density distribution of LOH events in HRD (red) and HRP (blue) samples. **d** The accuracy of the new LOH criteria (blue boxes) and those utilized in Telli et al. (black box); the size of dots represents the decision tree balanced accuracy (BA) when using the corresponding cut-off, colors correspond to the statistical difference in abundance of AIs between HRD versus HRP samples (U test p value). **e** Accuracy of HR classification using three tandem allelic imbalances LSTs of a given minimum length (x axis) and distance between them smaller than 1 to 4 Mb (y axis). Dot sizes and colors are presented similarly as in panel d. **f** Upper panel: Visualization of the statistical difference (U test p-values) in the abundance of TAIs between HRD versus HRP samples for selected TAIs length. Lower panel: The dot sizes and colors in the lower panel correspond to the description in panel d and e. **g** Density distribution of HRD-AIs according to Telli2016 algorithm, the red dashed line represents a cut-off vale of 54 to define the HR-status. The balanced accuracy (BA) of classifying the annotated HRD and HRP is shown, density distribution colors correspond to the samples annotated as in Fig. 2a. **h** OVA-TCGA samples stratified by genomic alterations and their corresponding ovaHRDscar levels. U test p values are shown for the comparison of ovaHRDscar levels between the corresponding alterations as compared to the samples with CCNE1 amplification. Horizontal lines represent the median values per group. **i** Bland-Altman plot that shows the concordance (Concordance correlation coefficient, CCC = 0.90) between the number of ovaHRDscars detected using SNP-arrays and WGS in the intersecting samples from OVA-TCGA and PCAWG. Blue line represent the average difference and the red dotted lines the 95% confidence intervals of the mean. **j** Correlation (Pearson, r=0.38) between the SBS3 proportion in WGS data from PCAWG versus the number of scars using the Telli2016 approach. Blue line shows the regression line, and the 95% confidence intervals are shown in grey. **k** The SBS3 status inferred using SigMA16 showing a higher agreement with ovaHRDscar (agreement=78.3%, Cohen's kappa = 0.56) than with the Telli2016 algorithm using an HRP/HRP cut-off value of 54 (agreement=77.2%, Cohen's kappa = 0.53). In the pie charts and table + and - correspond to the number of HRD positive and HRD negative samples identified under each criterion, respectively. On the bottom is shown the number of samples and the level of agreement between the corresponding criteria.

Supplementary Figure 3

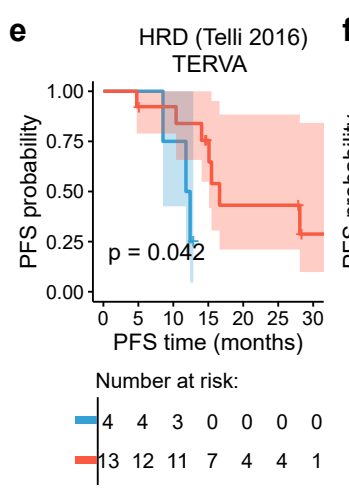


d *Cox model of PFS*

	OVA-TCGA n=353				OVA-TCGA* n=260				DECIDER n=32			
	N	**Prop**	**Hazard ratio**	**Pval**	**N**	**Prop**	**Hazard ratio**	**Pval**	**N**	**Prop**	**Hazard ratio**	**Pval**
*BRCA*mut/del	31	9%	0.94(CI:0.62−1.4)	0.78000	0	0%	NA(CI:NA−NA)	NA	5	16%	0.49(CI:0.17−1.4)	0.1900
Telli2016	289	82%	0.69(CI:0.51−0.94)	0.01700	208	80%	0.64(CI:0.46−0.9)	0.01100	23	72%	0.52(CI:0.23−1.2)	0.1200
Telli2016−54	215	61%	0.7(CI:0.55−0.89)	0.00420	146	56%	0.68(CI:0.51−0.89)	0.00600	15	47%	0.34(CI:0.15−0.76)	0.0090
Takaya2020	166	47%	0.67(CI:0.53−0.85)	0.00093	106	41%	0.63(CI:0.47−0.84)	0.00150	14	44%	0.37(CI:0.16−0.85)	0.0190
ovaHRDscar	196	56%	0.66(CI:0.52−0.84)	0.00068	132	51%	0.62(CI:0.47−0.82)	0.00069	16	50%	0.29(CI:0.13−0.65)	0.0027
*BRCA*mut/del excluded:			n=322				n=260				n=27	
Telli2016	258	80%	0.69(CI:0.51−0.94)	0.01900	208	80%	0.64(CI:0.46−0.9)	0.01100	18	67%	0.63(CI:0.27−1.5)	0.2800
Telli2016−54	187	58%	0.69(CI:0.53−0.88)	0.00340	146	56%	0.68(CI:0.51−0.89)	0.00600	10	37%	0.41(CI:0.16−1)	0.0560
Takaya2020	141	44%	0.66(CI:0.51−0.85)	0.00110	106	41%	0.63(CI:0.47−0.84)	0.00150	9	33%	0.45(CI:0.17−1.2)	0.0980
ovaHRDscar	169	52%	0.65(CI:0.51−0.83)	0.00068	132	51%	0.62(CI:0.47−0.82)	0.00069	11	41%	0.33(CI:0.13−0.83)	0.0190

Hazard ratio 0.8 0.7 0.6 0.5 0.4 · -log10(pval) 1 2 3 · -log10(pval) 2.0 2.5 3.0 · -log10(pval) 1.0 1.5 2.0 2.5

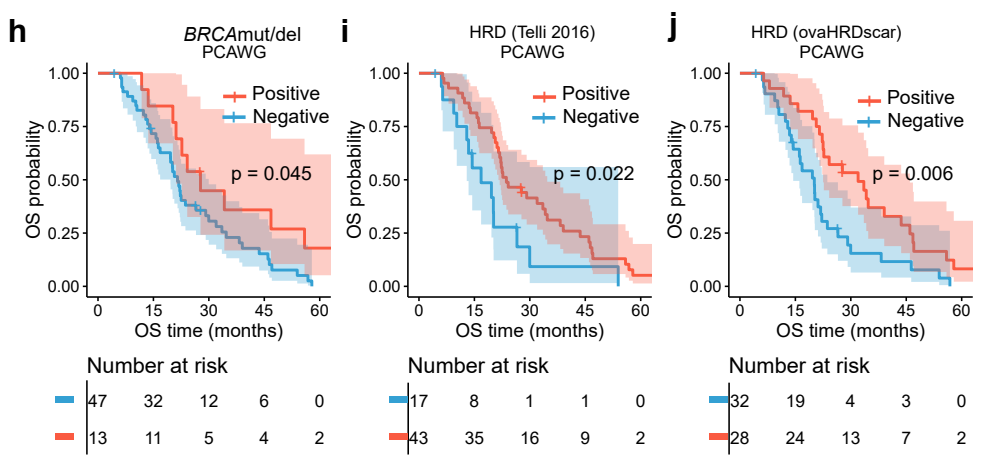
**Excluding patients used for ovaHRDscar optimization*



g *Cox model of PFS adjusted for residual tumor*

	TERVA n=17			
	N	**Prop**	**HR**	**Pval**
Telli2016	13	76%	0.213	0.110
Tacaya2020	6	37%	0.445	0.250
ovaHRDscar	8	47%	0.085	0.046

Hazard ratio (HR) 0.2 0.3 0.4 0.5 · -log10(pval) 0.75 1.00 1.25

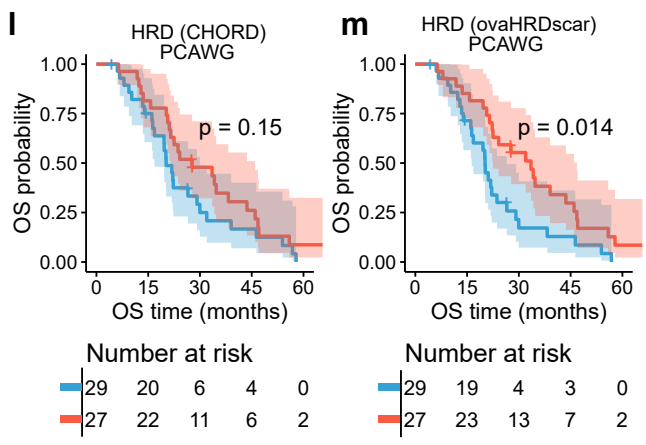


k *Cox model of OS*

	OVA-TCGA n=353				OVA-TCGA* n=260				PCAWG n=60			
Names	**N**	**Prop**	**Hazard ratio**	**Pval**	**N**	**Prop**	**Hazard ratio**	**Pval**	**N**	**Prop**	**Hazard ratio**	**Pval**
*BRCA*mut/del	31	9%	0.5(CI:0.28−0.9)	2.0e−02	0	0%	NA(CI:NA−NA)	NA	13	22%	0.5(CI:0.25−1)	0.0490
Telli2016	289	82%	0.51(CI:0.37−0.71)	6.5e−05	208	80%	0.49(CI:0.34−0.7)	9.7e−05	43	72%	0.48(CI:0.26−0.91)	0.0240
Telli2016−54	215	61%	0.56(CI:0.42−0.73)	2.2e−05	146	56%	0.52(CI:0.38−0.7)	2.8e−05	30	50%	0.51(CI:0.3−0.89)	0.0170
Takaya2020	166	47%	0.54(CI:0.41−0.71)	1.1e−05	106	41%	0.5(CI:0.36−0.7)	3.9e−05	20	33%	0.58(CI:0.32−1)	0.0710
ovaHRDscar	196	56%	0.52(CI:0.4−0.69)	2.6e−06	132	51%	0.49(CI:0.36−0.67)	8.2e−06	28	47%	0.46(CI:0.27−0.81)	0.0072
*BRCA*mut/del excluded:			n=322				n=260				n=47	
Telli2016	258	80%	0.54(CI:0.39−0.76)	3.0e−04	208	80%	0.49(CI:0.34−0.7)	9.7e−05	30	64%	0.57(CI:0.3−1.1)	0.0970
Telli2016−54	187	58%	0.56(CI:0.43−0.74)	5.0e−05	146	56%	0.52(CI:0.38−0.7)	2.8e−05	19	40%	0.7(CI:0.38−1.3)	0.2600
Takaya2020	141	44%	0.55(CI:0.42−0.73)	4.4e−05	106	41%	0.5(CI:0.36−0.7)	3.9e−05	11	23%	0.98(CI:0.47−2)	0.9500
ovaHRDscar	169	52%	0.52(CI:0.4−0.69)	5.5e−06	132	51%	0.49(CI:0.36−0.67)	8.2e−06	17	36%	0.58(CI:0.31−1.1)	0.1000

Hazard ratio 0.8 0.7 0.6 0.5 0.4 · -log10(pval) 2 3 4 5 · -log10(pval) 4.5 5.0 · -log10(pval) 0.5 1.0 1.5 2.0

**Excluding patients used for ovaHRDscar optimization*



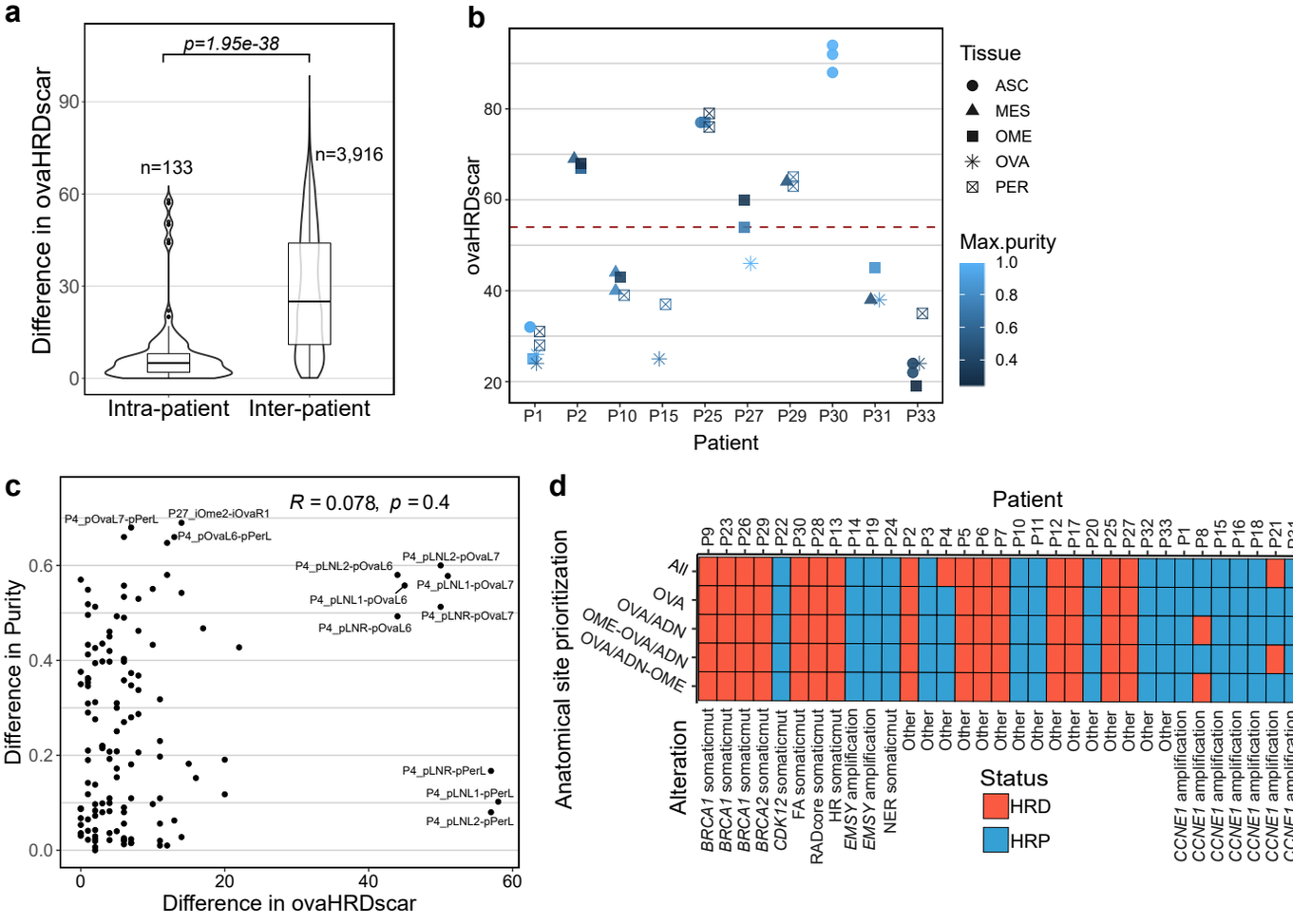
n *Cox model of OS*

	PCAWG n=56			
	N	**Prop**	**HR**	**Pval**
CHORD	27	48%	0.66	0.15
ovaHRDscar	27	48%	0.49	0.02

Hazard ratio (HR) 0.7 0.6 0.5 0.4 · -log10(pval) 1.0 1.2 1.4 1.6

Supplementary Figure 3. ovaHRDscar shows an improved prediction of PFS and OS in HGSC patients. **a to c** Kaplan-Meier plots for PFS in the DECIDER cohort. The patients were stratified using different criteria: the *BRCAmut*/del status in the left panel (**a**), the Telli2016 algorithm in the middle panel (**b**) and the ovaHRDscar algorithm (**c**). **d** Cox regression models for PFS in HGSC patients using different selection criteria. The colors, rows and columns descriptions are the same as in Figure 3d. **e to f** Kaplan-Meier plots of PFS in the prospective TERVA cohort. The patients were stratified using: The Telli2016 algorithm on the left (**e**), the ovaHRDscar on the right (**f**). **g** Cox regression models for PFS in the TERVA cohort. The patients were stratified using different criteria. The colors, rows and columns descriptions are the same as in Figure 3d. **h to j** Kaplan-Meier plots for OS in the DECIDER cohort stratified using the different criteria: *BRCA*mut/del status on the left (**h**), the Telli2016 algorithm (**i**) and the ovaHRDscar algorithm (**j**). **k** Cox regression models for OS adjusted for patient age at diagnosis in HGSC patients stratified using different criteria similarly as in Figure 3i. **l to m** Kaplan-Meier plots for OS in the PCAWG cohort stratified using: the CHORD signature on the left (**l**), the ovaHRDscar on the right (**m**). **n** Cox regression models for OS in the PCAWG cohort. Patients were stratified using CHORD and ovaHRDscar criteria; the dot colors, rows and columns descriptions are the same as in Figure 3d.

Supplementary Figure 4



e Cox regression for PFS adjusted for residual tumor

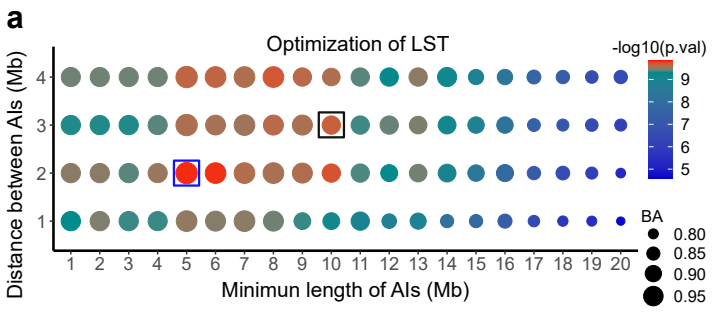
	All				OVA				OVA/ADN				OME-OVA/ADN				OVA/ADN-OME			
	N	Prop	HR	Pval	N	Prop	HR	Pval	N	Prop	HR	Pval	N	Prop	HR	Pval	N	Prop	HR	Pval
Telli2016	23	72%	0.73	0.092	22	69%	0.65	0.071	22	69%	0.65	0.071	23	72%	0.74	0.093	23	72%	0.74	0.093
Takaya2020	15	47%	0.57	0.056	14	44%	0.51	0.039	14	44%	0.51	0.039	15	47%	0.56	0.049	14	44%	0.51	0.039
ovaHRDscar	17	53%	0.54	0.048	15	47%	0.44	0.026	16	50%	0.38	0.019	16	50%	0.49	0.033	16	50%	0.38	0.019

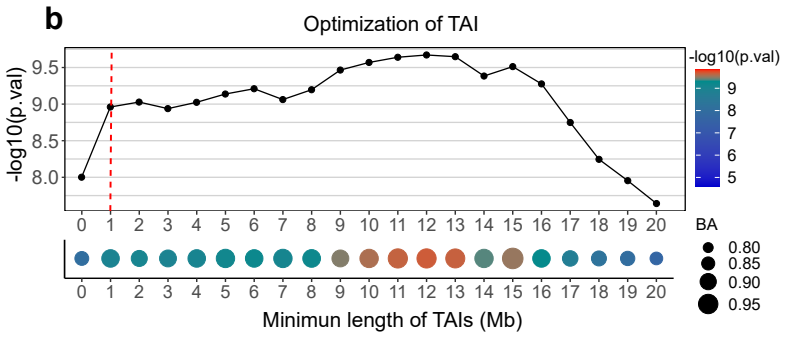
Hazard ratio (HR)
0.6 0.5 0.4 0.3

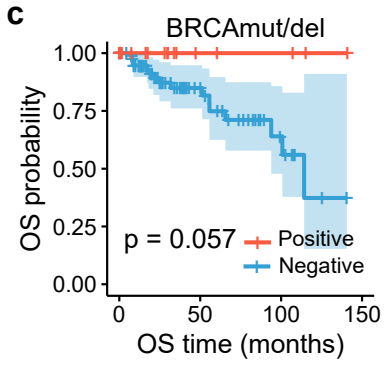
-log10(p.val)
0.9 1.2 1.5 1.8

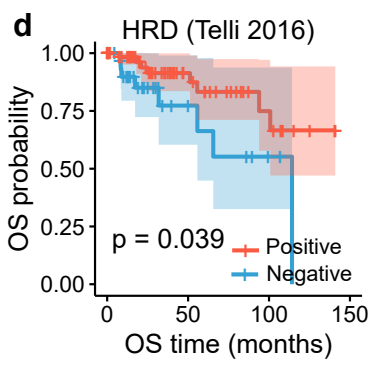
Supplementary Figure 4. Inter and intra-patient variability of ovaHRDscar levels. a Left box plot for the difference of ovaHRDscar values between all possible intra-patients' samples pairs. Right box plot for the corresponding difference of ovaHRDscar values in all possible inter-patients' samples pairs in the DECIDER prospective cohort. The black horizontal lines represent the sample medians, the boxes extend from first to third quartile and whiskers indicate the values at 1.5 times the interquartile range. U-test p-value is shown. **b** Samples from different anatomical sites with tumor purity and ovaHRDscar levels indicated. **c** The correlation between the difference of ovaHRDscar values between all intra-patient sample pairs and the difference in tumor purity of the corresponding sample pairs. Patient P4 was ignored as an outlier. **d** Color table of HR-status classification and HR-related genomic alterations using five different approaches to prioritize anatomical sites for ovaHRDscar calculations. **e** Cox regression models for PFS adjusted for the residual tumor after surgery in the DECIDER cohort using different algorithms and five different anatomical site prioritizations. Colors, rows and column descriptions are the same as in Figure 3d.

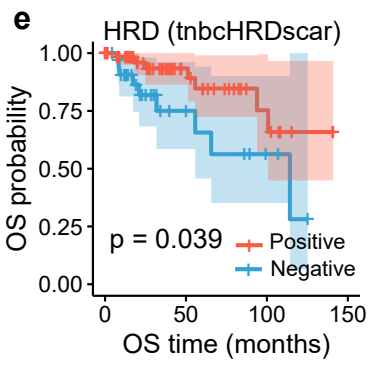
Supplementary Figure 5











Number at risk:

80	27	8	0
19	4	3	0

Number at risk:

31	8	2	0
68	23	9	0

Number at risk:

34	9	3	0
65	22	8	0

f *Cox model of PFI* TCGA, TNBC n=99

	N	Prop	Hazard ratio	Pval
*BRCA*mut/del	19	19%	0.6(CI:0.14−2.6)	0.490
Telli2016	68	69%	0.46(CI:0.19−1.1)	0.090
tnbcHRDscar	65	66%	0.34(CI:0.14−0.83)	0.018

Hazard ratio 0.7 0.6 0.5 0.4

-log10(Pval) 0.6 0.9 1.2 1.5

g *Cox model of OS* TCGA, TNBC n=99

	N	Prop	Hazard ratio	Pval
BRCAmut/del	19	19%	3.8e−09(CI:0−Inf)	1.000
Telli2016	68	69%	0.37(CI:0.14−0.99)	0.048
tnbcHRDscar	65	66%	0.37(CI:0.14−0.99)	0.047

Hazard ratio 0.7 0.6 0.5 0.4

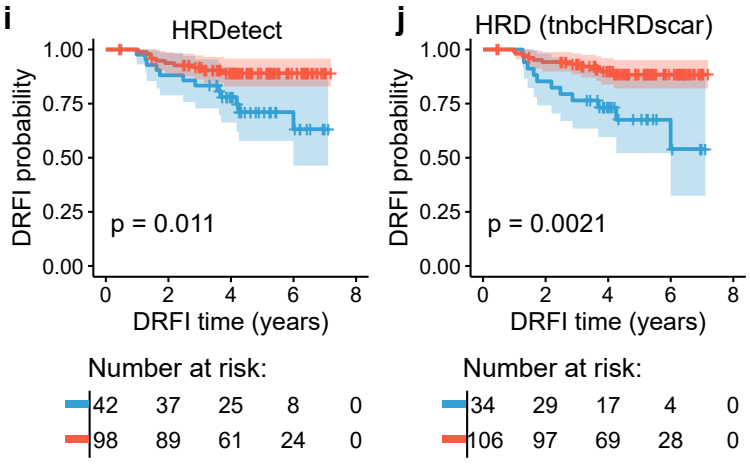
-log10(Pval) 0.25 0.75 1.25

h *Cox model of DFRI* n=140

	N	Prop	Hazard ratio	Pval
*BRCA*mut/del	21	15%	0.92(CI:0.27−3.1)	0.9000
Telli2016	117	84%	0.51(CI:0.2−1.3)	0.1700
HRDetect	98	70%	0.35(CI:0.15−0.82)	0.0150
tnbcHRDscar	106	76%	0.29(CI:0.13−0.67)	0.0039

Hazard ratio 0.7 0.6 0.5 0.4

-log10(Pval) 0.5 1.0 1.5 2.0

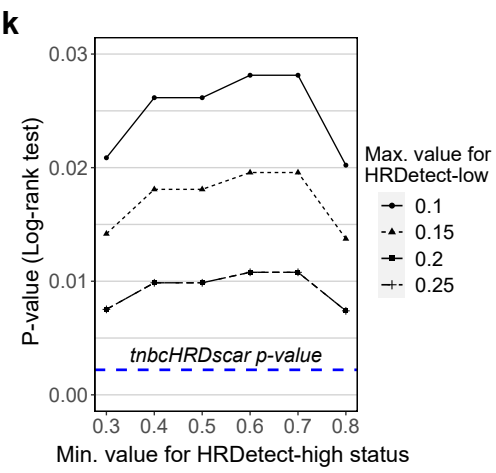


Number at risk:

42	37	25	8	0
98	89	61	24	0

Number at risk:

34	29	17	4	0
106	97	69	28	0



Supplementary Figure 5. Machine learning-aided detection of HRD-AI in TNBC improves the prediction of clinical outcomes. a Generation of HRD-LST events. The size of the dots represents the decision tree balanced accuracy (BA) of the classification of HRD and HRP when selected LSTs with the corresponding criteria. The dot colors correspond to the statistical difference in abundance of the selected LSTs between HRD versus HRP samples (U test, p-value). The black box corresponds to the selection criteria proposed by Telli2016, blue box corresponds to the tnbcHRDscar BA and U test value. **b** Upper panel: Visualization of the change in p-values (U test) when selecting TAIs >1Mb (red dashed line). Lower panel: the difference in abundance of TAIs of selected length between HRD versus HRP samples. The dot sizes and colors in the lower panel correspond to the description in panel a. **c to e** Kaplan-Meier plots for OS in the TCGA's TNBC patients stratified using the different criteria; **c** BRCAmut/del status on the left, **d** the Telli2016 algorithm, and **e** the tnbcHRDscar algorithm. **f** Cox regression models for PFS in HGSC patients stratified using different criteria, the dot colors descriptions are the same as in Figure 3d. **g** Cox regression models for OS in in the TCGA's TNBC patients stratified using the different criteria. **h** Cox regression models for DRFI in an independent TNBC patient cohort[21] stratified using different criteria. **i to j** Kaplan-Meier plots for DRFI in an independent TNBC patient cohort[21] stratified using: HRDetect based on whole genome sequencing data (**i**), the tnbcHRDscar based on SNP-array data (**j**). **k** Selection of different values to define the HRDetect-high/low status for patient stratification in the TNBC patient cohort[21] and its association with DRFI (Log-rank test, p-value). Patients with intermediate HRDetect values were ignored. In blue line, the Log-rank p-value when using tnbcHRDscar in panel j.