

Supplementary Table S1: Specific analysis roles and blinding status of contributing teams

Order of analysis	Role	Team members	Location/Affiliation	Blinding status
1	Pathological analyses, tissue banking, sample provision	CD, KE, BS, CSo, IS, TF	Local trial units, central pathology	blinded to molecular and clinical study data
2	RNA preparation and sequencing	BMY, BLJ	Avera Cancer Institute	fully blinded
3	Primary RNA-Seq raw data analysis and QC	TM, JB	Avera Cancer Institute	fully blinded
4	Sample coding, dataset assembly and distribution	KW, VN	GBG-Statistics Dept.	unblinded
5	Blinded gene expression analysis, development of statistical analysis plan and SPSS script	TK, UH	Goethe University	fully blinded
6	Correlation with patient data according to predeveloped SPSS script	KW, VN	GBG-Statistics Dept.	unblinded
7	Blinded interpretation of summary results	TK, UH	Goethe University	no patient level data
8	Monitoring and review of results	SL, MU, PAF, FM, VM, BG, CSch, CH, ES, JH, MvM	GBG-Boards	no patient level data

Supplementary Table S2: Pre-defined analytical aims of the study:

Pre-defined analytical aims of the study:	
1.	Concordance of RNA-Seq-derived genomic ER-/PR-status, proliferation, immune signature expression, and molecular subtype with pathology-derived IHC-based ER-/PR-status, histological grading, and tumor-infiltrating-lymphocyte (TIL)-scoring, respectively.
2.	Robustness of the above concordances with regard to sample quality (QC class).
3.	Univariate predictive value for pCR of RNA-Seq-derived molecular subtypes, and signatures for proliferation, stroma, T-cell signature, and hypoxia signature.
4.	Multivariate logistic regression of pCR including the following predictor variables: a) Hormone receptor status, treatment arm (+/- Bev), hypoxia signature, and the interaction between hypoxia signature and treatment arm. b) All predictor variables from (a), with additional clinical variables (age, cT, cN, histological grade) as predictors.

Supplementary Table S3: Comparison of clinical parameters of the complete trial cohort and the RNA-Seq cohorts

Parameter	Category	Total cohort 1540 [#] (100.0%)	RNA-Seq data 289 (100.0%)	<i>P</i> -Value [§]	High Quality RNA-Seq data 221 (100.0%)	<i>P</i> -Value [§]
Age	median	48	46	<0.001	46	0.009
clin. tumor status	T1	261 (17.0%)	45 (15.6%)	0.028	37 (16.8%)	0.035
	T2	883 (57.5%)	188 (65.3%)		145 (65.9%)	
	T3	224 (14.6%)	33 (11.5%)		23 (10.5%)	
	T4a-c	79 (5.1%)	8 (2.8%)		8 (3.6%)	
	T4d	89 (5.8%)	14 (4.9%)		7 (3.2%)	
	missing	4	1		1	
clin. lymph node status	negative	772 (51.0%)	148 (51.6%)	n.s.	115 (52.5%)	n.s.
	LN1-3	668 (44.1%)	124 (43.2%)		93 (42.5%)	
	LN4-9	57 (3.8%)	11 (3.8%)		7 (3.2%)	
	LN>=10	17 (1.1%)	4 (1.4%)		4 (1.8%)	
	missing	26	2		2	
HER2 status	negative	1540 (100.0%)	289 (100.0%)		221 (100.0%)	
Hormone receptor status	negative	558 (36.2%)	133 (46.0%)	<0.001	102 (46.2%)	0.001
	positive	982 (63.8%)	156 (54.0%)		119 (53.8%)	
Histological grade	G1	53 (3.5%)	9 (3.1%)	<0.001	7 (3.2%)	<0.001
	G2	781 (51.0%)	119 (41.5%)		86 (39.3%)	
	G3	697 (45.5%)	159 (55.4%)		126 (57.5%)	
	missing	9	2		2	
Histological subtype	Ductal invasive	1241 (80.8%)	238 (82.4%)	n.s.	179 (81.0%)	n.s.
	Lobular invasive	162 (10.5%)	26 (9.0%)		23 (10.4%)	
	other	133 (8.7%)	25 (8.7%)		19 (8.6%)	
	missing	4	0		0	
Treatment arm	EC-T	743 (48.2%)	150 (51.9%)	n.s.	117 (52.9%)	n.s.
	ECB-TB	797 (51.8%)	139 (48.1%)		104 (47.1%)	
pCR	no	1156 (75.1%)	149 (51.6%)	<0.001	112 (50.7%)	<0.001
	yes	384 (24.9%)	140 (48.4%)		109 (49.3%)	

[§] *P*-values are the result of Fisher's exact tests for binary variables, of chi-square tests for variables with three or more levels, and of Wilcoxon test for continuous variables, respectively.

[#] 1540 total patients with HER2 negative disease with response data from treatment arms out of the ITT population of 2572 patients from the GeparQuinto trial.

Supplementary Table S4: Gene lists of signatures

Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID	Signature	Gene	EntrezID			
Adipocyte	AASS	10157	Basal	ABC13		ER-related	ABCC8	6833	Hypoxia	CSTB	1476	IFN	CMPK2		MHC1	GBP1	2633	ProLiferation	ANLN		Ribosomal	CSDE1	7812	Stroma	ADAM12	8038	T-Cell	B2M	567
	ABCA10		Basal	COL27A1		ER-related	AFF3	3899	Hypoxia	NDRG1	10397	IFN	DDX58	23586	MHC1	GBP2	2634	ProLiferation	ARHGAP11A	9824	Ribosomal	EEF1A1	1915	Stroma	ADAMTS12	81792	T-Cell	CD74	972
	ABCA6	23460	Basal	CRYAB	1410	ER-related	AGR2	10551	Hypoxia	VEGFA	7422	IFN	DDX80	55601	MHC1	HLA-A	3105	ProLiferation	ASPM	259266	Ribosomal	EEF1G	1937	Stroma	BNC2	54796	T-Cell	CIT2	4261
	ABCA8	10351	Basal	ELF5	2001	ER-related	AR	367				IFN	DDX60L		MHC1	HLA-B	3106	ProLiferation	AURKA	6790	Ribosomal	HSP01	3329	Stroma	CACNA1C	775	T-Cell	CYBB	1536
	ABCA9		Basal	FOXK1	2296	ER-related	CAN12	771				IFN	EIF2AK2	5610	MHC1	HLA-C	3107	ProLiferation	BRP1	83990	Ribosomal	LRPPRC	10128	Stroma	COL10A1	1300	T-Cell	DOCK10	55619
Adipocyte	ADH1B	125	Basal	GABRP	2568	ER-related	CAPN8				IFN	EPSTI1		MHC1	HLA-E	3137	ProLiferation	BRB1	699	Ribosomal	NACA	4666	Stroma	COL11A1	1301	T-Cell	DOCK2	1794	
Adipocyte	ANK2	287	Basal	KRT23	25984	ER-related	CCDC170	80129				IFN	HERC5	51191	MHC1	HLA-F	3134	ProLiferation	BRB18	701	Ribosomal	PABPC1	26986	Stroma	COL5A3	50509	T-Cell	DOCK8	
Adipocyte	APOD	347	Basal	PROM1	8842	ER-related	CYP2B7P1				IFN	HERC6	55008	MHC1	NLRCS		ProLiferation	CNBN1	891	Stroma	PRKDC	5591	Stroma	COL8A1		T-Cell	HLA-DPA1	3113	
Adipocyte	COL14A1	7373	Basal	RGMA		ER-related	ERBB4	2066				IFN	IFI44	10561	MHC1	PSMB8	5696	ProLiferation	CDC20	991	Ribosomal	RPL10	6134	Stroma	COMP	1311	T-Cell	HLA-DRA	3122
Adipocyte	FAT4	79633	Basal	SFRP1	6422	ER-related	ESR1	2099				IFN	IFI44L	10964	MHC1	TAP1	6890	ProLiferation	CDC42		Ribosomal	RPL11	6135	Stroma	CERS1PLD2	83716	T-Cell	HLA-DRB1	3123
Adipocyte	FREM1		Basal	SLC34A2	10568	ER-related	FOXA1	3169				IFN	IFIH1	64135	MHC1	TAP2	6891	ProLiferation	CENPE	1062	Ribosomal	RPL12	6136	Stroma	CTGF	1490	T-Cell	IL10RA	3587
Adipocyte	IGF1	3479	Basal	SLC6A14	11254	ER-related	FSIP1				IFN	IFIT1	3434	ProLiferation	CENPF	1063	MHC1	WAR2	7453	ProLiferation	CIT	11113	Ribosomal	RPL19	6143	Stroma	NCKAP1L	3071	
Adipocyte	LAMA2	3908	Basal	SOX10	6663	ER-related	GATA3	2625				IFN	IFIH2	3433	ProLiferation	CIT				ProLiferation	DEPDC1	55635	Ribosomal	RPL23A		Stroma	DPVSL3	1809	
Adipocyte	PLEKHH2		Basal	TCF7L1	83439	ER-related	GFRA1	2674				IFN	MX1	4599	ProLiferation	DIAPH3	81624			ProLiferation	DEPDC1	55635	Ribosomal	RPL27A	6157	Stroma	FAP	2191	
Adipocyte	SPARCL1	8404	Basal	ZNF462		ER-related	GREB1	9687				IFN	MX2	4600	ProLiferation	DIAPH3	81624			ProLiferation	DIAPH3	81624	Ribosomal	RPL3	6122	Stroma	FBLN2	2199	
Adipocyte	SVEP1	79987	Basal	DSC3	1825	ER-related	GRPR	2925				IFN	OAS1	4938	ProLiferation	ECT2	1894			ProLiferation	ECT2	1894	Ribosomal	RPL32	6161	Stroma	FLNC	2318	
Adipocyte	TNXB	7148	Basal	DSG3	1830	ER-related	KDM4B	23030				IFN	OAS2	4939	ProLiferation	ESPL1	9700			ProLiferation	ESPL1	9700	Ribosomal	RPL4	6124	Stroma	FNDC1		
			Basal	FAT2	2196	ER-related	MAPT	4137				IFN	OAS3	4940	ProLiferation	EXO1	9156			ProLiferation	EXO1	9156	Ribosomal	RPL5	6125	Stroma	GAS6	2621	
			Basal	KRT14	3861	ER-related	MLPH	79083				IFN	PARP12	64761	ProLiferation	FANCA	2175			ProLiferation	FANCA	2175	Ribosomal	RPL9	6133	Stroma	GAS7	8522	
			Basal	KRT16		ER-related	NAT1	9				IFN	PLSCR1	5359	ProLiferation	FANCI	55215			ProLiferation	FANCI	55215	Ribosomal	RPLP0	6175	Stroma	GLIS2		
			Basal	KRT17	3872	ER-related	PGR	5241				IFN	RSAD2	91543	ProLiferation	FOXM1	2305			ProLiferation	FOXM1	2305	Ribosomal	RPS11	6205	Stroma	IGFBP4	3487	
			Basal	KRT5	3852	ER-related	SCUBE2	57758				IFN	SP100	6672	ProLiferation	GMPS	8833			ProLiferation	GMPS	8833	Ribosomal	RPS18	6222	Stroma	ITGA11		
			Basal	KRT6A		ER-related	SLC44A4	80736				IFN	SP110	3431	ProLiferation	GTSF1	51512			ProLiferation	GTSF1	51512	Ribosomal	RPS20	6224	Stroma	ITGA5	3678	
			Basal	KRT6B	3854	ER-related	TBC1D9	23158				IFN	UBA7	7318	ProLiferation	HIST1H2BO	8348			ProLiferation	HIST1H2BO	8348	Ribosomal	RPS24	6229	Stroma	ITGB5	3693	
			Basal	TRIM29	23650	ER-related	THSD4	79875				IFN	XAF1	54739	ProLiferation	HJURP	55355			ProLiferation	HJURP	55355	Ribosomal	RPS27A	6233	Stroma	ITGBL1	9358	
												IFN	ZNFX1		ProLiferation	IQGAP3				ProLiferation	IQGAP3		Ribosomal	RPS4X	6191	Stroma	KANK2	25959	
															ProLiferation	KIF14	9928			ProLiferation	KIF14	9928	Ribosomal	RPS6	6194	Stroma	KIAA1462	57608	
															ProLiferation	KIF23	9463			ProLiferation	KIF23	9463	Ribosomal	RPS7	6201	Stroma	KIF26B	55083	
															ProLiferation	KIF2C	11004			ProLiferation	KIF2C	11004	Ribosomal	RPSA		Stroma	LMOD1	25802	
															ProLiferation	KIFC1	3833			ProLiferation	KIFC1	3833	Ribosomal	TPT1	7178	Stroma	LRRC15	131578	
															ProLiferation	KPNA2	3838			ProLiferation	KPNA2	3838			Stroma	MAP1A	4130		
															ProLiferation	LMNB1	4001			ProLiferation	LMNB1	4001			Stroma	MICAL2	9645		
															ProLiferation	MCM10	55388			ProLiferation	MCM10	55388			Stroma	MMP11	4320		
															ProLiferation	MELK	9833			ProLiferation	MELK	9833			Stroma	MMP14	4323		
															ProLiferation	MKI67	4288			ProLiferation	MKI67	4288			Stroma	MRC2	9902		
															ProLiferation	MYBL2	4605			ProLiferation	MYBL2	4605			Stroma	MXRA8	54587		
															ProLiferation	NCAPO2	9918			ProLiferation	NCAPO2	9918			Stroma	MYH1	4619		
															ProLiferation	ORCS	23594			ProLiferation	ORCS	23594			Stroma	MYL9	10398		
															ProLiferation	POIQ	10721			ProLiferation	POIQ	10721			Stroma	NID1	4811		
															ProLiferation	PRC1	9055			ProLiferation	PRC1	9055			Stroma	NID2	22795		
															ProLiferation	PRR11				ProLiferation	PRR11			Stroma	PCOLCE	5118			
															ProLiferation	RACGAP1	29127			ProLiferation	RACGAP1	29127			Stroma	PDLIM7	9260		
															ProLiferation	RRM2	6241			ProLiferation	RRM2	6241			Stroma	PHLD81	23187		
															ProLiferation	SMC4	10051			ProLiferation	SMC4	10051			Stroma	PMEPA1	56937		
															ProLiferation	SPAG5	10615			ProLiferation	SPAG5	10615			Stroma	PRNP22	5376		
															ProLiferation	STIL	6491			ProLiferation	STIL	6491			Stroma	PODN			
															ProLiferation	TICRR				ProLiferation	TICRR			Stroma	RIN2	54453			
															ProLiferation	TOP2A	7153			ProLiferation	TOP2A	7153			Stroma	SFRP2			
															ProLiferation	TPX2	22974			ProLiferation	TPX2	22974			Stroma	SFRP4	6424		
															ProLiferation	TTK	7272			ProLiferation	TTK	7272			Stroma	SPON1	10418		
																									Stroma	SSCSO			
																									Stroma	TAGLN	6876		
																									Stroma	TENM4	26011		
																									Stroma	THY1	7070		
																									Stroma	TIMP2	7077		
																									Stroma	TIMP3	7078		
																									Stroma	UNC5B	219699		
																									Stroma	ZFHX4	79776		

Supplementary Table S5: Core genes of the hypoxia signature cluster from different datasets with correlated expression

Gene symbol	Gene name	Category	Details	FFPE-RNA-Seq data
VEGFA	Vascular endothelial growth factor	Angiogenesis	Growth factor active in angiogenesis, vasculogenesis and endothelial cell growth. Induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis and induces permeabilization of blood vessels. Binds to the FLT1/VEGFR1 and KDR/VEGFR2 receptors, heparan sulfate and heparin.	yes
NDRG1	N-myc downstream regulated gene 1	Stress response	Involved in stress responses, hormone responses, cell growth, and differentiation. Necessary for p53-mediated caspase activation and apoptosis.	yes
ANGPTL4	Angiopoietin-like 4	Angiogenesis, hypoxia	Hypoxia-induced expression in endothelial cells. May act as a regulator of angiogenesis and modulate tumorigenesis. In response to hypoxia, the unprocessed form of the protein accumulates in the subendothelial extracellular matrix.	no
ADM	Adrenomedullin	Angiogenesis	Adrenomedullin functions include vasodilation, regulation of hormone secretion, and promotion of angiogenesis.	no
DDIT4	DNA damage induced transcript 4	Stress response	Regulates cell growth, proliferation and survival via inhibition of mTORC1. Important role in responses to cellular energy levels and cellular stress, including responses to hypoxia and DNA damage.	no
CSTB	Cystatin-B	Proteinase inhibitor	Intracellular thiol proteinase inhibitor thought to play a role in protecting against proteases leaking from lysosomes.	yes

Supplementary Table S6: Accuracy of IHC detection of NDRG1 as marker of the hypoxia signature

Cohort	Finding (N=23)	Finding (N=23)	Finding (N=23)	Full (N=193)
Source	Whole slide	Whole slide	TMA	TMA
Method	pathological scoring	digital image analysis	digital image analysis	digital image analysis
Cutoff	pos/neg	>10% positive cells	>10% positive cells	>10% positive cells
Positive by RNA-Seq*	19	19	19	20
Negative by RNA-Seq*	4	4	4	173
Accuracy	91.3 %	73.9 %	47.8 %	81.9 %
Sensitivity	89.5 %	68.4 %	36.8 %	40.0 %
Specificity	100.0 %	100.0 %	100.0 %	86.7 %
PPV	100.0 %	100.0 %	100.0 %	25.8 %
NPV	66.7 %	40.0 %	25.0 %	92.6 %

* based on cutoff z-score 1.5 from RNA-Seq

Supplementary Table S7: Comparison of hormone receptor status from RNA-Seq and IHC

High Quality samples	Sensitivity	Specificity	PPV	NPV	Accuracy
ER _{RNA-Seq} vs. ER _{IHC} (N=221)	75.7 %	93.4 %	92.6 %	78.0 %	84.2 %
PR _{RNA-Seq} vs. PR _{IHC} (N=221)	76.6 %	83.5 %	77.4 %	82.8 %	80.5 %
Low Quality samples	Sensitivity	Specificity	PPV	NPV	Accuracy
ER _{RNA-Seq} vs. ER _{IHC} (N=68)	72.7 %	88.6 %	85.7 %	77.5 %	80.9 %
PR _{RNA-Seq} vs. PR _{IHC} (N=67)	61.3 %	86.1 %	79.2 %	72.1 %	74.6 %

Supplementary Table S8: Comparison of Molecular Subtyping between High Quality and Low Quality Samples

Group	Basal-like	Her2-enrich.	LumA	LumB	Normal-like
Total (N=289)	119 (41.2%)	60 (20.8%)	33 (11.4%)	46 (15.9%)	31 (10.7%)
HQ no dupl. (N=221, %)	103 (46.6%)	33 (14.9%)	28 (12.7%)	42 (19.0%)	15 (6.8%)
LQ no dupl. (N=68, %)	16 (23.5%)	27 (39.7%)	5 (7.4%)	4 (5.9%)	16 (23.5%)

The distributions of molecular subtypes according to AIMS differ significantly ($P = 8.6 \times 10^{-9}$, Fisher's Exact Test) between samples with high and low quality.

Supplementary Table S9: Univariate logistic regression of pCR by molecular markers (N=221 High Quality samples)

Molecular marker	OR	95% CI	P-value
Basal-like*	8.88	2.34-33.6	0.001
HER2-enriched*	3.33	0.79-14.1	0.10
Lum-A*	0.87	0.18-4.28	0.86
Lum-B*	2.22	0.54-9.14	0.27
T-cell signature [#]	1.60	1.21-2.12	0.001
Proliferation signature [#]	2.88	2.00-4.16	<0.001
Hypoxia signature [#]	1.92	1.41-2.60	<0.001

* vs. Normal-like subtype, [#] z-score

Supplementary Table S10: Multivariate logistic regression of pCR with NDRG1 from TMA analysis

	OR	95% CI	P-value
Hormone receptor (neg. vs. pos.)	4.35	2.34-8.07	<0.001
NDRG1-TMA-IHC (z-score >1.5)	3.79	0.85-16.9	0.080
Bevacizumab therapy	0.92	0.53-1.60	0.766
Interaction NDRG1-TMA-IHC * bevacizumab	1.31	0.78-2.21	0.309
cN (≥10 vs 4-9 vs 1-3 vs 0 positive nodes)	0.80	0.52-1.24	0.323
cT (T4d vs T4a-c vs T3 vs T2 vs T1)	0.80	0.59-1.0	0.173
Grading (G3 vs G2 vs G1)	1.43	0.83-2.48	0.197