Supplementary Tables

Casasent et al. 2017

Table S1. Related to Subject Details. Synchronous DCIS-IDC Patient Cohort.

Patient	Age	TNBC	ER	PR	HER2	grade	Stage	Cells
P1	57	Y	-/-	-/-	-/-	3/3	IIB	57
P2	36	N	+/+	+/+	+/+	3/3	IIB	114
P3	64	N	+/+	+/+	-/-	1/1	IV	102
P4	66	N	+/+	+/+	-/-	2/2	IIIC	104
P5	47	N	+/+	-/-	-/-	3/3	IIA	148
P6	77	Y	-/-	-/-	-/-	3/3	IIA	204
P7	66	N	-/+	-/-	-/-	3/3	IIIC	112
P8	62	Y	-/-	-/-	-/-	3/3	IIA	235
P9	49	Y	-/-	-/-	-/-	3/3	IIA	96
P10	48	Y	-/-	-/-	-/-	3/3	IIA	122

This table contains clinical information on the 10 patients with synchronous DCIS-IDC tumors that were analyzed by single cell and exome sequencing. Clinical parameters listed include patient age, triple-negative breast cancer status, estrogen, progesterone and HER2 receptor status, tumor grade and tumor stage. The receptor status and grade were scored independently for the DCIS and IDC regions and are displayed on the left (DCIS) and right-hand (IDC) side in these columns. The total number of single cells analyzed by TSCS is also indicated in the last column.

Table S2. Related to Figures 6-7. Exome Sequencing Metrics.

Patient	Normal		In situ		Invasive	
	Depth	Breadth	Depth	Breadth	Depth	Breadth
P1	107	0.9599	116	0.9628	58	0.9626
P2	76	0.9601	142	0.9553	124	0.9536
P3	100	0.9587	187	0.9665	116	0.9597
P4	153	0.9543	136	0.9538	140	0.9524
P5	104	0.9571	46	0.9458	125	0.9526
P6	87	0.9558	180	0.9607	89	0.9539
P7	177	0.9552	211	0.9561	136	0.9501
P8	144	0.9584	315	0.965	105	0.9577
P9	280	0.9638	335	0.9635	298	0.9621
P10	213	0.9608	287	0.9628	110	0.9528
Mean	144.1	0.95841	195.5	0.95923	130.1	0.95575

This table shows the exome sequencing metrics for the 10 DCIS-IDC patients, in which laser-capture-microdissection was used to isolate in situ and invasive regions from frozen tissue sections. Matched normal breast tissue was sequenced in parallel. Coverage depth was calculated for the in situ regions (mean=195.5X, SEM=29.4), invasive regions (mean=130.1X, SEM=20.1) and normal tissues (mean=144X, SEM=20.3). Coverage breadth, or physical coverage, was also calculated from the targeted exon regions for each patient and tissue region. Coverage breadth is defined as the percentage of the targeted capture regions in which at least 1X coverage depth was achieved after sequencing.

Table S3. Related to Figures 6-7. Invasive-Specific Mutations.

Patient	Gene	Chr	Position	Ref	Var	nucleotide	AA	Validated	SITU_FQ	INV_FQ	POLY	SIFT	genbank ID
P3	CRY1	chr12	107395083	C	T	G659A	R220H	Invasive Only	0.01	0.52	0.91	0	NM_004075
P3	DRD1	chr5	174869715	G	A	C388T	R130W	Invasive Only	0.00	0.47	0.98	0	NM_000794
P4	TECRL	chr4	65274856	G	A	C214T	Q72X	Invasive Only	0.00	0.27	0.74	0.66	NM_001010874
P7	FAP	chr2	163070565	C	G	G885C	W295C	Not Tested	0.00	0.22	1	0.09	NM_004460
P7	SCN4A	chr17	62025343	G	C	C3225G	I1075M	Invasive Only	0.00	0.08	NA	0	NM_000334
P7	PCDHA5	chr5	140201856	C	G	C496G	Q166E	Invasive Only	0.00	0.08	0.52	0.02	NM_031501
P8	AKAP6	chr14	33242950	C	G	C3439G	L1147V	Invasive Only	0.01	0.46	0.87	0	NM_004274
P8	SORBS2	chr4	186541289	T	A	A696T	R232S	Invasive Only	0.01	0.71	0.71	0	NM_001145675
P8	FANCA	chr16	89815132	A	T	T3283A	F1095I	Not Tested	0.00	0.38	0.96	0.42	NM_001286167
P8	CEP192	chr18	13071040	T	A	T5177A	I1726N	Not Tested	0.00	0.38	0.92	NA	NM_032142
P8	TRIM47	chr17	73872438	C	A	G907T	D303Y	Not Tested	0.00	0.62	0.98	0.02	NM_033452

This table lists the invasive-specific nonsynonymous mutations that were identified by exome sequencing of laser-capture-microdissected tissue regions. The mutations listed were detected in the invasive regions and were not detected in the in situ region, after filtering by matched normal germline variants. The table lists the following columns in order: patient identifiers, gene names, chromosome and position, reference nucleotide, variant nucleotide, nucleotide position in gene, amino acid positions (AA), amplicon deep-sequencing validation status, in situ mutation frequencies (SITU_FQ) normalized by tumor purity, invasive mutation frequencies (INV_FQ) normalized by tumor purity, polyphen2 damaging impact scores (POLY), SIFT functional impact prediction scores, and GenBank identifiers.sequencing.

Table S4. Related to Figures 6-7. Targeted Amplicon Deep Sequencing.

Patient	Gene	Chr	Position	Ref	Var	Type	A	C	G	T	SITU_MF	INV_MF	NORM_MF	P-value DN	P-value IN	Validated
						in situ	72	9	52208	123						
						Invasive	1666	1	10190	45						
P3	DRD1	chr5	174869715	G	A	Normal	75	4	14579	35	0.00	0.14	0.01	2.7211E-02	1.7892E-03	INV ONLY
						in situ	503	128175	11	245						
						Invasive	59	11798	4	2161						
P3	CRY1	chr12	107395083	C	T	Normal	76	10319	1	38	0.00	0.15	0.00	Not Significant	1.3216E-03	INV ONLY
						in situ	9770	171	4214428	1993						
						Invasive	62	0	464	0						
P4	TECRL	chr4	65274856	G	A	Normal	0	1	556	0	0.00	0.12	0.00	1.9430E-18	3.2333E-03	INV ONLY
						in situ	276279	541	1033	569662						
						Invasive	14452	26	93	47125						
P4	MYO18A	chr17	27423866	T	A	Normal	114	13	99	59998	0.33	0.23	0.00	1.9430E-18	2.0994E-04	Pre-existing
						in situ	2815	19	90	89006						
						Invasive	25	0	4	121						
P4	ZBTB14	chr18	5291971	T	A	Normal	1	0	9	5430	0.03	0.17	0.00	2.1105E-06	5.1415E-12	Pre-existing
						in situ	480489	780	925921	754						
						Invasive	7491	48	81899	148						
P4	MYO18B	chr22	26239717	G	A	Normal	113	43	128978	37	0.34	0.08	0.00	3.3535E-21	4.2214E-04	Pre-existing
						in situ	575	929093	81	191						
						Invasive	111	28932	2373	31						
P7	PCDHA5	chr5	140201856	C	G	Normal	126	86789	28	18	0.00	0.08	0.00	2.0250E-02	1.4055E-04	INV ONLY
						in situ	63	194	404521	567						
						Invasive	37	12095	145897	442						
P7	SCN4A	chr17	62025343	G	C	Normal	12	246	100712	80	0.00	0.08	0.00	2.4231E-05	1.4316E-03	INV ONLY
						in situ	32	14378	16	1244						
						Invasive	7	11738	13	20						
P7	ARHGAP36	chrX	130217875	C	T	Normal	1	2693	3	9	0.08	0.00	0.00	1.0944E-04	Not Significant	DCIS ONLY
						in situ	39	20	81322	3933						
						Invasive	76	24	231953	15359						
P7	ABCC11	chr16	48226526	G	T	Normal	54	23	185477	502	0.05	0.06	0.00	1.1563E-11	1.4828E-03	Pre-existing
						in situ	18229	58	117956	423						
						Invasive	2372	48	212951	538						
P7	MBL2	chr10	54530499	G	A	Normal	191	4	16598	28	0.13	0.01	0.01	3.3084E-08	Not Significant	DCIS ONLY
						in situ	1736	22	17703	49						
						Invasive	405	145	234196	403						
P7	ARID1B	chr6	157431633	G	A	Normal	95	133	246785	240	0.09	0.00	0.00	7.6277E-76	Not Significant	DCIS ONLY
						in situ	11	3258	19190	92						
						Invasive	2	32	14489	49						
P7	C9orf24	chr9	34382807	G	C	Normal	9	334	77171	74	0.14	0.00	0.00	6.8761E-18	5.2801E-02	DCIS ONLY
						in situ	7	413	139	65000						
						Invasive	4	10052	66	39499						
P8	NCOA2	chr8	71075008	T	C	Normal	8	83	44	23264	0.01	0.20	0.00	Not Significant	1.4828E-03	Pre-existing
						in situ	28146	38	6	1012						
						Invasive	139769	544	141	60075						
P8	MMP8	chr11	102589262	A	T	Normal	171278	606	48	536	0.03	0.30	0.00	2.3582E-13	1.8028E-05	Pre-existing
						in situ	43564	104	21	1320						
						Invasive	18089	96	16	7499						
P8	RNF182	chr6	13977826	A	T	Normal	16286	109	17	174	0.03	0.29	0.01	1.5575E-02	3.9687E-04	Pre-existing
						in situ	442	125801	47	88						
						Invasive	38570	108833	42	73						
P8	LAMTOR1	chr11	71809862	C	A	Normal	337	48816	15	28	0.00	0.26	0.01	Not Significant	2.3879E-04	INV ONLY
						in situ	817	549025	111	8784						
						Invasive	2187	1148556	896	468608						
P8	LTBP2	chr14	75019600	C	T	Normal	80	21492	23	3765	0.02	0.29	0.15	2.1247E-28	Not Significant	Pre-existing
						in situ	32	10	99	42162						
						Invasive	2986	7	24	9309						
P8	SORBS2	chr4	186541289	T	A	Normal	54	33	230	21864	0.00	0.24	0.00	Not Significant	6.4445E-44	INV ONLY
						in situ	215	242691	297	30						
						Invasive	483	238533	43940	52						
P8	AKAP6	chr14	33242950	C	G	Normal	72	76655	331	8	0.00	0.16	0.00	2.3346E-02	4.6470E-04	INV ONLY
						in situ	1	335	0	0						
						Invasive	1	469	0	3						
P8	HDAC4	chr2	240078423	C	T	Normal	162	82716	4	30	0.00	0.01	0.00	Not Significant	Not Significant	False Postive

Reference
Variant

This table shows the results of targeted amplicon deep sequencing for a subset of the in situ-specific and invasive-specific mutations detected by exome sequencing. DeepSNV was used to determine the statistical significance of each mutation relative to the site-specific background error rate in matched normal tissues (Methods). The table columns include patient number, gene name, chromosome number, chromosome position, reference base, variant base, region, read counts for A, C, G, and T, in situ mutation frequencies (SITU_MF), invasive mutation frequencies (INV_MF), p-value for DeepSNV of in situ to normal comparison (p-value DN), p-value for DeepSNV invasive to normal comparison (p-value), and finally the validation results.

Table S5. Related to Figures 6-7. Increased Mutations Frequencies in Invasions.

Patient	Gene	Chr	Position	Ref	Var	nucleotide	AA	Validated	SITU_FQ	INV_FQ	POLY	SIFT	genbank ID
P1	MEGF9	chr9	123476336	C	A	G301T	E101X	Not Tested	0.19	1.00	NA	0	NM_001080497
P3	NPY4R	chr10	47087501	C	T	C718T	R240C	Not Tested	0.48	1.00	0.332	0.16	NM_001278795
P5	AHDC1	chr1	27874569	G	A	C4058T	T1353M	Not Tested	0.33	1.00	0.983	0	NM_001029882
P8	KIAA0195	chr17	73489016	C	A	C1919A	T640N	Not Tested	0.07	0.85	0.732	0.02	NM_014738
P8	NCOA2	chr8	71075008	T	C	A914G	Q305R	Pre-existing	0.05	0.60	NA	0.05	NM_006540
P8	RNF182	chr6	13977826	A	T	A476T	D159V	Pre-existing	0.21	0.86	0.996	NA	NM_001165033
P8	LTBP2	chr14	75019600	C	T	G1187A	R396H	Pre-existing	0.11	0.85	0.603	0	NM_000428
P8	LAMTOR1	chr11	71809862	C	A	G231T	E77D	Pre-existing	0.03	0.77	0.334	0.03	NM_017907
P8	MMP8	chr11	102589262	A	T	T667A	L223M	Pre-existing	0.25	0.88	0.732	0	NM_002424

This table lists nonsynonymous mutations with increased frequencies (>0.5) in the exome data of the laser-microdissected in situ and invasive regions. The table columns list patient identifiers, gene names, chromosome and position, reference nucleotide, variant nucleotide, nucleotide position in gene, amino acid (AA) positions, amplicon deep-sequencing validation status, in situ mutation frequency adjusted by tumor purity (SITU_FQ), invasive mutation frequency adjusted by tumor purity (INV_FQ), polyphen2 damaging impact scores (POLY), SIFT functional impact prediction scores, GenBank identifiers, and the full gene names.

Table S6. Related to Oligonucleotides. Primer Sequences.

Patient	Gene	Chr	Position	Forward Primer	Reverse Primer
P3	CRY1	chr12	107395083	CTGTTAATCCTCCCCTTTCAACA	CAGGTGGAGAAACTGAAGCAC
P3	DRD1	chr5	174869715	CTGCACTGGGATGAAGGAGA	TGTAACATCTGGGTGGCCTT
P4	MYO18A	chr17	27423866	CTTGGTGTCTTGCAGCTCAG	TCCACCTATGCTTCTGGGTC
P4	ZBTB14	chr18	5291971	GGACACATCACGCTTCTGAG	AGATGTGTTCTTGCTGCCTG
P4	MYO18B	chr22	26239717	TGTCTGTTTCCAGGGTCCTC	AAACACAGTGCCTCCAATGC
P4	TECRL	chr4	65274856	TTTGGGGCAGAGAAGAAACAC	GGCCCTCTAAGACCAACTCC
P7	MBL2	chr10	54530499	GGGTGAAGTCAGCTCAGACC	AGGGCTCAGAGGCTTACAGG
P7	ABCC11	chr16	48226526	GAAGAGCTTGTTGTGCAGGG	AGACCAATAGCAGCCGAGAG
P7	SCN4A	chr17	62025343	GACTCACATCCACGATGAGG	GAGTCATTCGCACCATCCTA
P7	PCDHA5	chr5	140201856	AATGCCAGATTCGCGGTTTC	GGGGCGTTATCATTAGCATCC
P7	ARID1B	chr6	157431633	TCTTGGTTTTGCATGACTGC	CCTTATCGCCTGAAAGCACT
P7	C9orf24	chr9	34382807	GTGCTGGGTGCAACAGATAC	CCACCCCTATTATCCAGCAG
P7	ARHGAP36	chrX	130217875	GTGGTGCGAAGGGTGTTT	GGCCTCACATTCCTGCATAC
P8	MMP8	chr11	102589262	GCCTGAATGCCATCGATGTC	GCTGTTTACCACCCGTGTTT
P8	LAMTOR1	chr11	71809862	TGAGAAGCTTGGACTAGGGG	GCACTGAGTCCTGTTTGTCC
P8	LTBP2	chr14	75019600	CTCTCTGGCCATCTACCCTC	GATCGTCTTCACTCCCACCA
P8	AKAP6	chr14	33242950	TCTCTCTATGTTGCAGTCCCTC	GGCTTCCCACCTTCTTTCAA
P8	HDAC4	chr2	240078423	CTGGGCACTGACCTGTTTTC	TTGACCAGAGTTCTCCACCC
P8	SORBS2	chr4	186541289	TGGCTGGTAAGTGCTGTATCA	ACTGGCCCTGTTCTTTCCAA
P8	RNF182	chr6	13977826	CAGTCCTTCTCACACGTCCT	GTTCCACACAGTCCAGTTGT
P8	NCOA2	chr8	71075008	ACAGCTCACGAGAACAGTCA	GAGAGCAGCCATGAAACCAG

This table lists the forward and reverse oligonucleotide primer sequences used for PCR amplification of the targeted in situ-specific or invasive-specific mutations for targeted deep-sequencing.