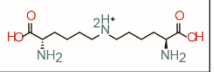
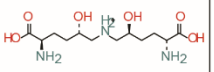
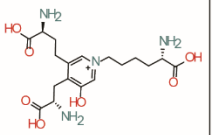
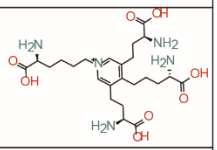
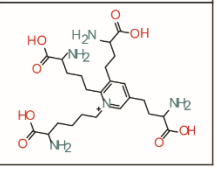
Supplementary Table 1

xAA Standard	Letter Code	Elemental Composition	RT (min)	Mass (Da)	Accuracy (ppm)	LLOD (fmole)	R^2	LLOQ (fmole)	Structure
Lysinonorleucine	LNL	C12H26N3O4+	5.69	276.19	1.5	75	0.96	225	
Dihydroxy lysinonorleucine	DHLNL	C12H26N3O6+	7.16	308.18	2	125	0.95	375	
deoxy Lysyl pyridinoline	dPyr	C18H29N4O7+	8.14	413.2	4.5	250	0.98	750	
Desmosine*	Des	C24H40N5O8+	10.14	526.29	0.9	156	0.92	468	
Isodesmosine*	iDes	C24H40N5O8+	10.14	526.29	0.9	156	0.92	468	

Supplementary Table 1: Summary of crosslinked amino acid standard characterization by mass spectrometry. Serial dilutions of crosslinked amino acid standards were prepared in the background of *E. coli* hydrolysates and the lowest limit of detection (LLOD) and the lowest limit of quantification (LLOQ) were determined on a QExactive mass spectrometer. The LLOD is defined here as the concentration that is required to produce a signal that is three times the noise level. The LLOQ is the analyte concentration that is required to produce a signal that is three times that of the LLOD. *Denotes isomers that were not resolved.

Supplementary Table 2

Distribution of patient and tumor characteristics				
All n=910				
Tumor in tissue microarray, n (%)			Yes 718/910 (78.9)	
Lysyl hydroxylase 2 (LH2) epithelial expression assessable, n (%)		Yes, 468/718 (65.2%)		
LH2 neg/weak/moderate/strong		**LH2 negative**	**LH2 weak**	**LH2 moderate/strong**
		279 (59.6%)	112 (23.9%)	77 (16.5%)
Factor	n (%) or mean (min-max)	n (%) or mean (min-max)	n (%) or mean (min-max)	n (%) or mean (min-max)
Age at diagnosis	65.5 (45.7-87.3)	63.9 (48.4-84.7)	61.7 (46.4-85.6)	63.9 (45.7-87.3)
years (n= 910)				
BMI at baseline (n=910)				
<25	466 (51)	154 (56.8)	64 (57.1)	36 (46.8)
≥25 and <30	310 (34)	77 (28.4)	36 (32.1)	27 (35.1)
>30	134 (15)	40 (14.8)	12 (10.7)	14 (18.2)
Tumor size (n= 887)				
≤20 mm	637 (72)	190 (70.6)	77 (68.8)	43 (55.8)
>20 mm	250 (28)	79 (29.4)	35 (31.3)	34 (44.2)
ALNI (n=859)				
Negative	588 (68.5)	175 (67.6)	67 (61.5)	50 (65.8)
Positive (≥1 metastatic node)	271 (31.5)	84 (32.4)	42 (38.5)	26 (34.2)
Grade, NHG (n=860)				
I	233 (27.1)	84 (31.7)	21 (18.9)	12 (15.8)
II	405 (47.1)	138 (52.1)	52 (46.8)	14 (18.4)
III	222 (25.8)	43 (16.2)	38 (34.2)	50 (65.8)
ER status (n=784)				
Positive (>10%)	690 (88.0)	230 (93.5)	93 (87.7)	44 (63.8)
Negative (<10%)	94 (12.0)	16 (6.5)	13 (12.3)	25 (36.2)
HER2 status (n=609)				
Negative	556 (91.3)	161 (89.9)	71 (86.6)	50 (89.3)
Positive	53 (8.7)	18 (10.1)	11 (13.4)	6 (10.7)
Ki67 (n=655)				
Low (≤10%)	434 (66.3)	143 (71.5)	58 (61.7)	18 (32.7)
High (>10%)	221(33.7)	57 (28.5)	36 (38.3)	37 (67.3)
Molecular subtypes (n=639)				
ER+/HER2-	536 (83.9)	169 (88.0)	76 (80.9)	29 (53.7)
HER2+	53 (8.3)	5 (4.4)	11 (11.7)	6 (11.1)
TNBC	50 (7.8)	3 (2.7)	7 (7.4)	19 (35.2)

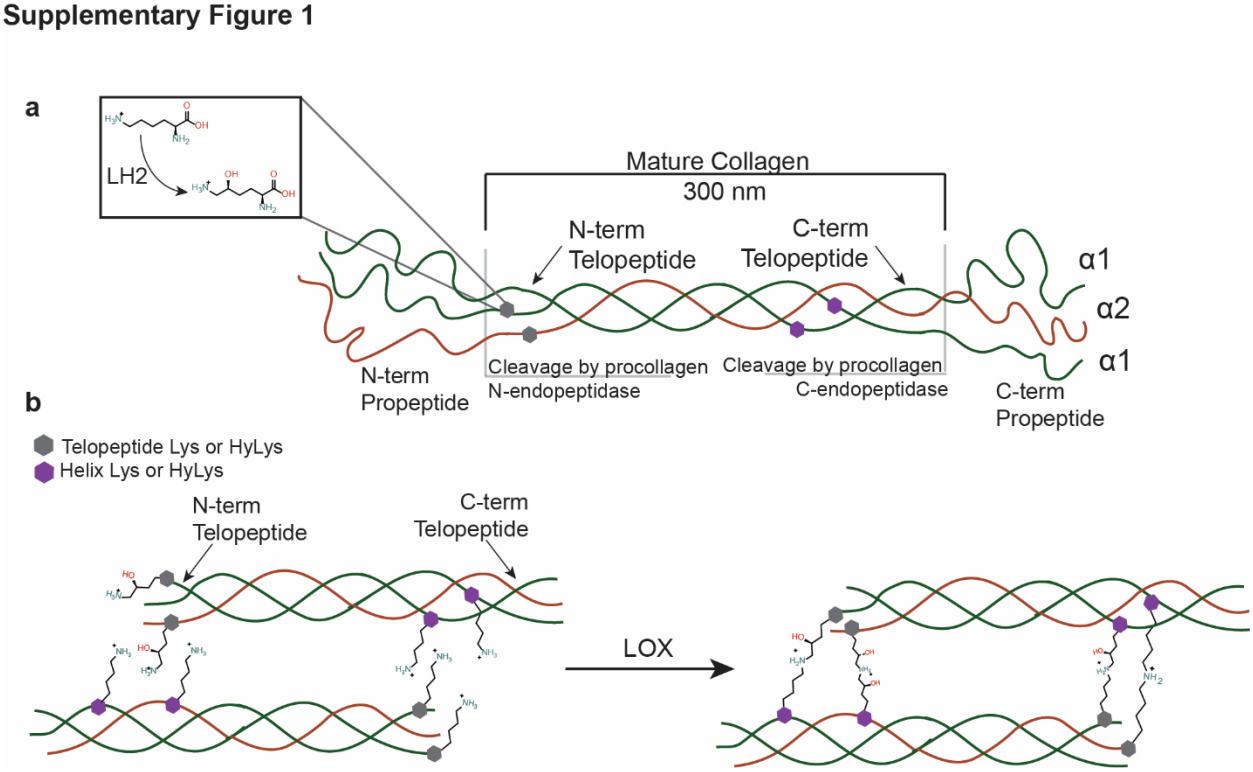
Supplementary Table 2: Characterization of breast cancer patients used to develop neoplastic epithelial LH2 H-score.

Supplementary Table 3

	Distribution of patient and tumor characteristics			
All n=910				
Tumor in tissue microarray, n (%)			Yes 718/910 (78.9)	
Lysyl hydroxylase 2 (LH2) stromal expression assessable, n (%)			Yes, 505/718 (70.3%)	
LH2 low/intermediate/high		**LH2 low**	**LH2 moderate**	**LH2 high**
		171 (33.9%)	188 (37.2%)	146 (28.9%)
Factor	n (%) or mean (min-max)	n (%) or mean (min-max)	n (%) or mean (min-max)	n (%) or mean (min-max)
Age at baseline	56.4 (44.7-73.0)	53.7 (44.9-73.0)	54.2 (44.7-72.7)	53.6 (45.0-72.8)
years (n= 910)				
Age at diagnosis	65.5 (45.7-87.3)	62.6 (48.5-84.7)	63.4 (45.7-87.3)	63.7 (46.4-85.6)
years (n= 910)				
BMI at baseline (n=910)				
<25	466 (51)	98 (57.3)	110 (58.5)	73 (50.0)
≥25 and <30	310 (34)	52 (30.4)	55 (29.3)	47 (32.2)
>30	134 (15)	20 (11.7)	23 (12.2)	26 (17.8)
Tumor size (n= 887)				
≤20 mm	637 (72)	118 (69.8)	131 (70.1)	95 (65.1)
>20 mm	250 (28)	51 (30.2)	56 (29.9)	51 (34.9)
ALNI (n=859)				
Negative	588 (68.5)	116 (72.5)	116 (64.8)	90 (62.5)
Positive (≥1 metastatic node)	271 (31.5)	44 (27.5)	63 (35.2)	54 (37.5)
Grade, NHG (n=860)				
I	233 (27.1)	53 (32.3)	60 (32.3)	21 (14.7)
II	405 (47.1)	86 (52.4)	79 (42.5)	54 (37.8)
III	222 (25.8)	25 (15.2)	47 (25.3)	68 (47.6)
ER status (n=784)				
Positive (>10%)	690 (88.0)	133 (94.3)	155 (89.1)	106 (77.9)
Negative (<10%)	94 (12.0)	8 (5.7)	19 (10.9)	30 (22.1)
HER2 status (n=609)				
Negative	556 (91.3)	104 (95.4)	120 (88.9)	85 (84.2)
Positive	53 (8.7)	5 (4.6)	15 (11.1)	16 (15.8)
Ki67 (n=655)				
Low (≤10%)	434 (66.3)	83 (72.2)	98 (66.7)	54 (48.2)
High (>10%)	221(33.7)	32 (27.8)	49 (26.1)	58 (51.8)
Molecular subtypes (n=639)				
ER+/HER2-	536 (83.9)	105 (92.9)	118 (81.9)	73 (68.2)
HER2+	53 (8.3)	5 (4.4)	15 (10.4)	16 (15.0)
TNBC	50 (7.8)	3 (2.7)	11 (7.6)	18 (16.8)

Supplementary Table 3: Characterization of breast cancer patients used to develop stromal LH2 H-score.

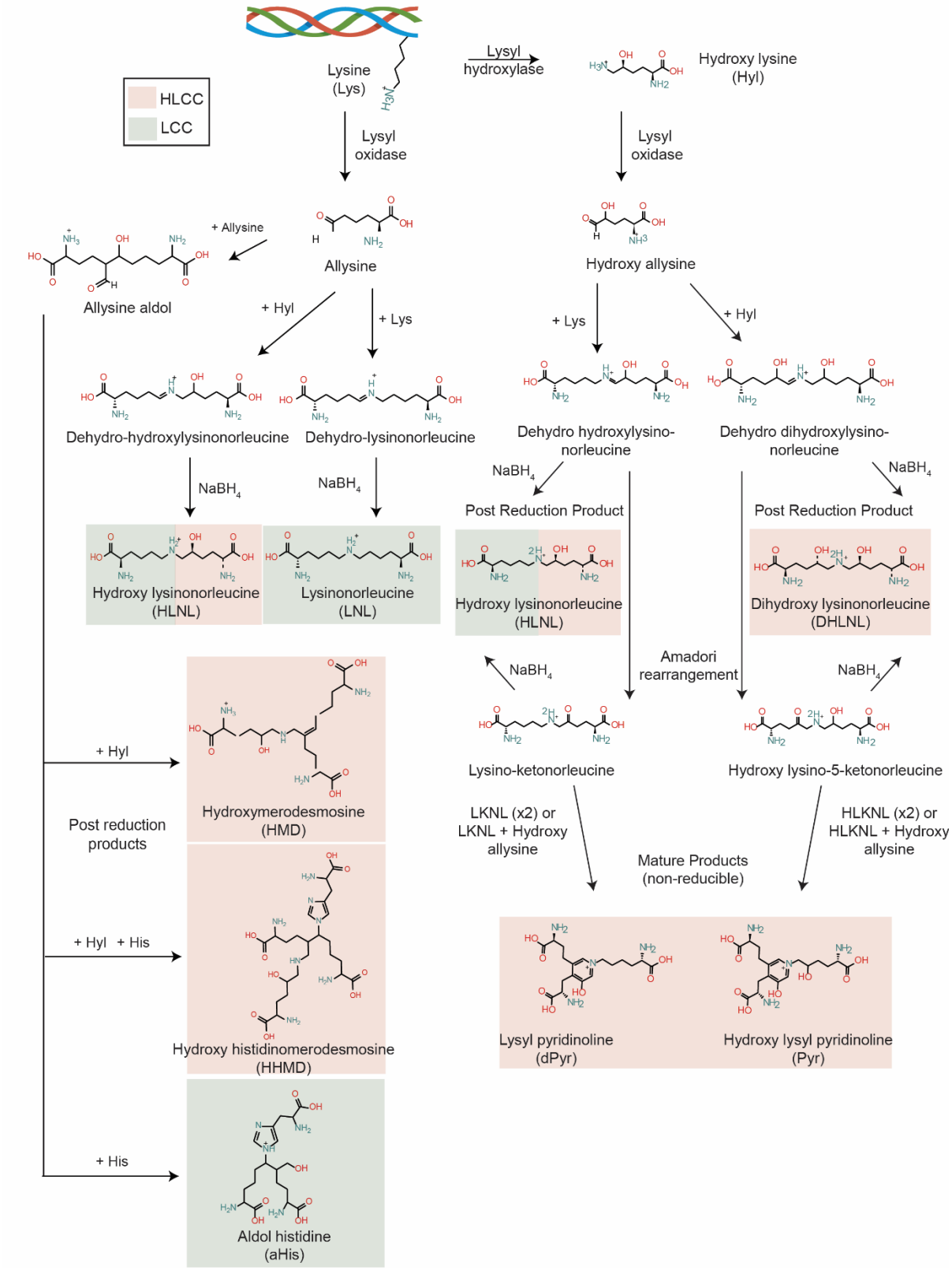
Supplementary Figure 1



Supplementary Figure 1: Schematic diagram of lysine hydroxylation and crosslinking in fibrillar collagen. (**a**) Schematic of a mature fibrillar collagen fiber. N- and C- terminal telopeptides are hydroxylated by lysyl hydroxylase 2 (LH2). N- and C- terminal propeptides are cleaved by procollagen endopeptidases to form the mature collagen fiber (300 nm). (**b**) Lysine (Lys) and the hydroxylysine (Hyl) residues in the telopeptide region of mature collagen are targeted by the crosslinking enzyme lysyl oxidase (LOX), which forms reactive aldehyde groups that spontaneously undergo aldol condensation reactions to form covalent collagen crosslinks.

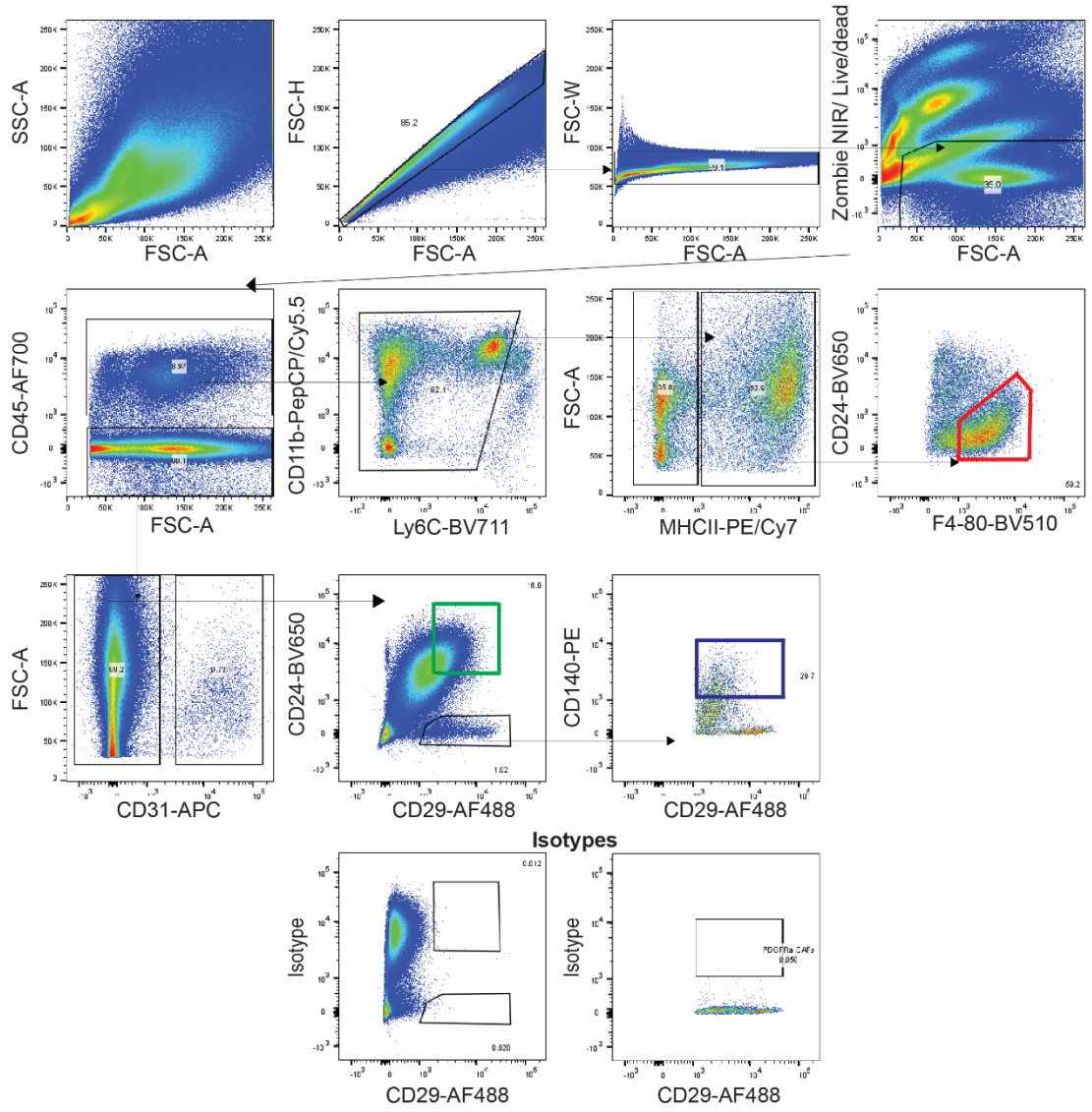
Supplementary Figure 2

Lysine aldehyde (Lys^{ald}) and hydroxylysine aldehyde (Hyl^{ald}) collagen crosslinking pathways



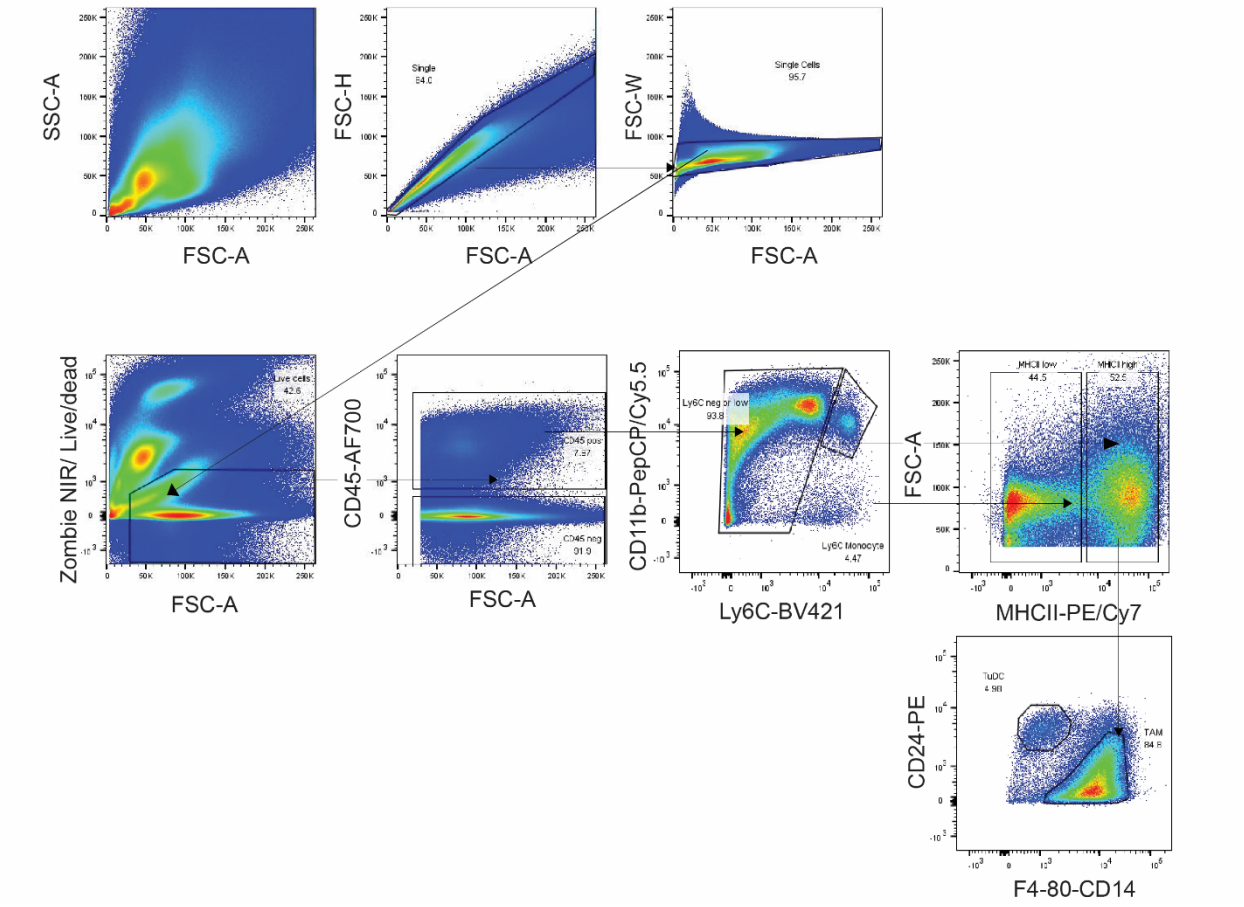
Supplementary Figure 2: Lysine aldehyde (Lys^{ald}) and hydroxylysine aldehyde (Hyl^{ald}) crosslinking pathway. Lysyl oxidase modifies Lys residues to form allysine (Lys^{ald}), which spontaneously reacts with Lys and Hyl residues in the helical region to form the Schiff bases dehydro-hydroxylysinonorleucine (deH-HLNL) dehydro-lysinonorleucine (deh-HLNL). These crosslinks can be reduced with $NaBH_4$ to form LNL and HLNL. The mature products of these crosslinks are currently unknown. Allysine can combine with an additional allysine residue to form allysine aldol. Allysine aldol can form the trivalent crosslink hydroxyl merodesmosine, aldol histidine, or the tetravalent crosslink histidino-hydroxymerodesmosine (HHMD) (only post-reduction products shown) through aldol condensation reactions with Hyl or histidine (His), or a combination of the two. Red and green shading denotes lysine-derived collagen crosslinks (LCC) or hydroxyl lysine-derived collagen crosslinks (HLCC)[20]. Telopeptide lysine residues are modified by lysyl hydroxylase. Lysyl oxidase modifies Hyl residues to hydroxyl allysine (Hyl^{ald}), which spontaneously reacts with Lys and Hyl residues to form the Schiff bases dehydro-dihydroxylysinonorleucine (deH-DHLNL) and dehydro-hydroxylysinonorleucine (deh-HLNL). They then undergo Amadori rearrangements to form hydroxylysino-keto-norleucine (HLKNL) or lysine-keto-norleucine (LKNL), respectively. These crosslinks can be reduced with $NaBH_4$ to form LNL and DHLNL. Mature crosslink products (Pyr and dPyr) are formed from the reaction of lysine ketonorleucine (LKNL) or hydroxyl lysinoketonorleucine (HKLNL) with hydroxyl allysine. Red and green shading denotes lysine-derived collagen crosslinks (LCC) or hydroxyl lysine-derived collagen crosslinks (HLCC)[20].

Supplementary Figure 3



Supplementary Figure 3: Gating strategy for sorting tumor cells, cancer-associated fibroblasts, and tumor-associated macrophages from PyMT mice via flow cytometry. Gates include tumor-associated macrophages (red), tumor cells (green), and cancer-associated fibroblasts (blue). This gating strategy was used for all flow cytometry experiments except Extended Data 7h.

Supplementary Figure 4



Supplementary Figure 4: Gating strategy for sorting tumor-associated macrophages from PyMT mice via flow cytometry for RNA-Seq (Extended Data Figure 7h).