

# The Role of Aldh1a1 in Contributing to Breast Tumour Aggressiveness: a Study Conducted in an African Population

Eric Gyan (✉ [erickgyan@gmail.com](mailto:erickgyan@gmail.com))

University of Nottingham School of Medicine <https://orcid.org/0000-0001-5215-4421>

**Andrew Green**

University of Nottingham School of Medicine

**Linda Ahenkorah-Fondjo**

Kwame Nkrumah University of Science and Technology School of Medical Sciences

**Andrew Jackson**

University of Nottingham School of Medicine

**Micheal Shawky Toss**

University of Nottingham School of Medicine

**Patrick Kafui Akakpo**

University of Cape Coast School of Medical Sciences

**Leonard Derkyi-Kwarteng**

University of Cape Coast School of Medical Sciences

**Ganiyu A. Rahman**

University of Cape Coast School of Medical Sciences

**William K. B. A Owiredu**

Kwame Nkrumah University of Science and Technology, Department of Molecular Medicine, School of Medical Sciences, Kumasi, Ghana

---

## Research article

**Keywords:** Breast cancer, CD44, CD24, ALDH1, Africans, prognosis, Immunohistochemistry

**Posted Date:** August 17th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-55483/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

---

**Version of Record:** A version of this preprint was published at Annals of Diagnostic Pathology on April 1st, 2021. See the published version at <https://doi.org/10.1016/j.anndiagpath.2020.151696>.

## Abstract

Aldehyde dehydrogenase 1 member A1 (ALDH1A1) is one of the most well studied breast cancer stem cells. Its expression has been associated with poor clinicopathological features and clinical outcomes in several studies. This paper studies the expression of ALDH1A1 and its combination with CD44<sup>+</sup>/CD24<sup>-/low</sup> breast cancer stem cell and their association with clinicopathological parameters and molecular subtypes.

**METHOD:** Tissue Microarray was constructed from 222 Formalin Fixed Paraffin Embedded (FFPE) breast cancer tissues. The expression of ALDH1A1, CD44 and CD24 were assessed by Immunohistochemistry (IHC). The association of ALDH1A1 and its association with clinicopathological parameters, molecular subtypes, CD44 and CD24 were studied in an African population. The association between CD44<sup>+</sup>/CD24<sup>-/low</sup>/ALDH1<sup>+</sup> and the clinicopathological phenotypes were also studied.

**RESULTS:** A high ALDH1A1 expression of 90% was recorded in this study. No association was found between ALDH1A1 and clinicopathological parameters. ALDH1A1 was positively associated with CD24 ( $r = 0.228$ , OR-4.599 95% CI- 1.751–12.076,  $p = 0.001$ ) and CD44 ( $r = 0.228$ , OR-5.538 95%CI- 1.841–16.662,  $p = 0.001$ ) but not associated with CD44<sup>+</sup>/CD24<sup>-/low</sup> ( $r = 0.134$ , OR- 2.720 95%CI- 0.959–7.710,  $p = 0.052$ ). CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> however had significant associations with Age ( $p = 0.020$ ,  $r = 0.161$ , OR- 2.771, 95%CI 1.147–6.697), Gender ( $p = 0.004$ , OR- 15.333 95%CI 1.339–175.54), Tumour grade ( $p = 0.005$ ,  $r = 0.197$ , OR-3.913 95%CI 1.421–10.776) and clinical prognostic staging ( $p = 0.014$ ,  $r = 0.182$ , OR-3.028 95%CI- 1.217–7.536). There was no association between CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> and the molecular subtypes.

**CONCLUSION:** The high expression of ALDH1A1 in breast cancer makes it an important target for targeted therapy. This study further confirms the increased tumourigenicity of CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> combination phenotype and its association with increased tumour grade and clinical prognostic stage. Survival studies of ALDH1A1 and other breast cancer stem cells in African populations is strongly recommended to help further understand their effect on tumour aggressiveness.

## Introduction

Despite advances in molecular techniques in cancer research, the role of cancer stem cells (CSC) in tumour initiation, differentiation, progression, therapy resistance and recurrence remain puzzling. This is partly because of their representation as a smaller but key subpopulation in the highly heterogenous tumour microenvironment with their ability to enter quiescence to evade regular therapeutic agents that target proliferating cells. The involvement of CSC in the aforementioned sequence of events in cancer development has been established(1–3) but the molecular interactions among themselves and with other intratumoral components is still under intense research. Targeting and eradicating CSC is crucial in the development of novel cancer therapies.

One of the well-studied CSC marker is Aldehyde dehydrogenase 1 (ALDH1), a cytosolic ubiquitous detoxifying enzyme expressed in several tumours such as breast, laryngeal, ovarian, gastric and non-small cell lung cancer(4–7). This enzyme converts aliphatic aldehydes and retinol into carboxylic acid and retinoic acid respectively in an NAD(P)<sup>+</sup> dependent oxidation. Its increased expression has largely been associated with breast cancer aggressiveness and poor prognosis(1, 8–11). On the contrary, other studies such as that of Liu *et al* reported that ALDH1 associated with better outcome in triple negative breast cancer(12) and still others did not find any association with clinicopathological features or clinical outcomes(13–15). The role of ALDH1 in breast cancer outcomes therefore remains controversial and further investigation is warranted in the quest to unravel its 'for now' confusing role(16).

ALDH1 proteins consist of 3 main isozymes namely ALDH1A1, ALDH1A2 and ALDH1A3 with ALDH1A1 being the most specific which is largely reported to be related to poor prognosis(17). The chemoresistance role of ALDH1 stems from its ability to detoxify anticancer drugs such as oxazaphosphorine via oxidation of aldophosphamide by activation of ALDH1 expression. The main stay of cancer treatment in Africa involves chemotherapy and radiotherapy. Such treatment regimens are limited in their inability to target cancer stem cells resulting in drug resistance and tumour recurrence via the aforementioned mechanism(18–20).

The role of ALDH1 related to cancer stem cell and ALDH1's association with poor prognosis in most tumours has been widely determined. CD44<sup>+</sup>/CD24<sup>-/low</sup> /ALDH1<sup>+</sup> cancer stem cell forms a phenotype with a higher tumourigenicity relative to the individual stem cell contributions. Evidence exist for this combined phenotype (CD44<sup>+</sup>/CD24<sup>-/low</sup>/ALDH1<sup>+</sup>) having the propensity to achieve tumour initiation from as low as 20 cells as opposed to 100 cells by CD44<sup>+</sup>CD24<sup>-/low</sup> alone in *in-vivo* studies(1, 21). With such

enhanced tumorigenic potential, the role of the CD44<sup>+</sup>/CD24<sup>-/low</sup>/ALDH1<sup>+</sup> phenotype in contributing to the aggressiveness of breast cancers of African populations is not well understood. A study of breast cancer stem cells is crucial in improving breast cancer management and clinical outcomes in Africans. The aim of this study is to analyse the expression of ALDH1, and its combination phenotype CD44<sup>+</sup>/CD24<sup>-/low</sup>/ALDH1<sup>+</sup> and their associations with molecular and clinicopathological features of breast cancer in an African population.

## Materials And Methods

A retrospective study was carried out of some breast biopsies and reports of patients (n = 222) presenting with breast cancer at the departments of pathology, Korle-Bu Teaching Hospital, Accra and the Cape Coast Teaching Hospital, Cape Coast, two of the five teaching hospitals in Ghana between 2012 and 2018. The Korle-Bu Teaching Hospital department of pathology is the largest in Ghana, receiving specimens from the Korle-Bu Teaching Hospital, the largest referral hospital in Ghana and from other health facilities within the Greater Accra Region(22). The department also receives specimen from all other regions of Ghana. The Cape Coast Teaching Hospital's pathology department receives specimen from mainly the Central and Western regions of Ghana. Clinical data was obtained from histopathology request forms. Two pathologist (PKA and LDK) reviewed histopathology slides of selected cases within the study period (2012–2018). Selection was done according to the quality of the FFPE blocks. Archival blocks of primary breast carcinoma from the two pathology departments were retrieved. Additional clinical information and histopathological features were obtained from the histopathology reports of patients. The information included the mean age of presentation, duration of symptoms, tumour grade, (based on mitotic count, nuclear grade, tubule formation). All cases were reviewed histopathologically and classified according to the recent WHO classification for breast tumours and histopathological grading done in accordance with the Nottingham criteria.

## Tissue Microarray (TMA) Construction

Areas of tumour were selected after preparation of histopathology slides from archival formalin fixed paraffin embedded (FFPE) blocks' and staining with Hematoxylin and Eosin (H&E). Areas of normal tissue, necrosis, and haemorrhage were ignored. Using the TMA Grand Master® (3D HISTECH®, Budapest, Hungary), three cores 1 mm each (2 from peripheral tumour and 1 central tumour) were punched out from the representative selected areas and arrayed into a new recipient paraffin block. Four micrometer thickness of TMA sections were cut and mounted on Superfrost slides.

## Immunohistochemistry (IHC)

TMA were stained using ALDH1A1 Rabbit Polyclonal antibody, Sigma Life Science (Prestige Antibodies) HPA002123, at dilution of 1:50. CD24 and CD44 staining was done with CD24 Monoclonal antibody (SN3), Thermofisher and CD44 monoclonal antibodies (156-3C11), Thermofisher in dilutions of 1:200 and 1:750 respectively and incubated in a black box for 1hour at room temperature. Immunohistochemical antibody labelling was done using the NOVOLINK polymer detection system (Leica, Newcastle, UK). Adhesion of tissue to the slide was done by pre-heating tissue microarrays at 60°C on a hot plate for 20 minutes and cooled. Tissue sections were deparaffinised in xylene and rehydrated through a series of graded alcohols and rinsed in distilled water. Antigen retrieval was achieved by boiling slides in citrate buffer (27 ml of citrate in 123 ml disodium citrate and made up to 1.5L with ddH2O) at pH-6.0 and microwaved (Whirlpool JT359 Jet Chef 1000W) at full power for 20 minutes. Peroxidase blocking reagent from the NOVOLINK® kit was used to block the endogenous peroxidase activity for 5 minutes and rinsed with PBS for 15 minutes. Protein blocking was done for 5 minutes to minimize nonspecific binding and rinsed thoroughly with PBS for 15 minutes. Primary antibody was added in the following dilutions: ALDH1- 1:50 and incubated in a black box for 1hour at room temperature. Positive control was Liver tissue, and negative control was obtained by omitting the primary antibody in the staining protocol. A thorough rinse was done for 15 minutes with PBS tween and then incubated with Post Primary Novolink reagent for 30 minutes in a black box. After a 15-minute thorough rinse, a polymer was added and incubated for 30 minutes. The reaction was then developed by incubating a 3,3'-diaminobenzidine chromogen solution (DAB) made up to 1:20 in dilution with DAB substrate buffer and incubated for 5 minutes. Counter staining with hematoxylin was done and incubated for 6 minutes. Dehydration and clearing were done using the Leica auto Stainer. Sections were then mounted with DPX. Evaluation of staining was done.

The semi-quantitative H scoring system was employed in scoring. The intensity of ALDH1 expression was scored as 0 (no expression), 1 (weak), 2 (moderate) and 3 (strong). The total score was calculated as the percentage of positive cells multiplied by the intensity giving a range of 0-300. A cut point of  $\leq 40\%$  score was designated as negative and  $> 40\%$  as positive.

For oestrogen and progesterone receptor staining, positive expression was considered as nuclear immunoreactivity in  $\geq 1\%$  of neoplastic cells. HER2 was analyzed at the time of diagnosis according to the American Society of Clinical Oncology (ASCO) / College of American Pathologists (CAP) protocols.

## Statistical Analysis

IBM SPSS version 24.0 package program (SPSS inc., Chicago, IL, USA) was used in the statistical analysis. The association between the markers and clinicopathological features were done with cross tables using chi-square test and odd ratios. Correlations were done with Pearson's correlation test. Statistical significance was set at 95% confidence interval.

## Results

A high Aldehyde dehydrogenase 1A1 (ALDH1A1) expression of 90.2% was recorded in our cohort. There was no association of ALDH1A1 with any of the clinicopathological characteristics. Table 1 shows the association between ALDH1A1 and clinicopathological features. However ALDH1A1 was positively correlated with CD24 ( $r = 0.228$ , OR-4.599 95% CI- 1.751–12.076,  $p = 0.001$ ) and CD44 ( $r = 0.228$ , OR-5.538 95%CI- 1.841–16.662,  $p = 0.001$ ) but not associated with CD44+/CD24- ( $r = 0.134$ , OR- 2.720 95%CI- 0.959–7.710,  $p = 0.052$ ). There was an inverse association with ER expression ( $r = -0.145$ , OR = 0.209, 95% CI- 0.047–0.928,  $p = 0.025$ ) but no significant association existed between PR and Her2 (Table 2).

Malignant Phyllodes was the only histologic types that had association with ALDH1A1. A 10-fold decreased expression of ALDH1A1 was recorded among other histologic types ( $p = 0.047$ ,  $r = 0.136$  OR- 10.15 95%CI 0.611-168.989) (Table 3).

## CD44/CD24/ALDH1 combination phenotypes

The predominant combination phenotype was CD44+/CD24+/ALDH1+ representing 72% of the total. CD44+/CD24-/ALDH1+ was the second most occurring combination phenotype (12.3%). The least occurring combination phenotype was CD44-/CD24+/ALDH1- representing 0.9%. Table 4 shows the frequency of the combination phenotypes and Table 5 also shows the distribution of CD44/CD24/ALDH1 combination phenotypes across molecular subtypes.

From Table 6, CD44+/CD24-/ALDH1+ had significant associations with age ( $p = 0.020$ ,  $r = 0.161$ , OR- 2.771, 95%CI 1.147–6.697), gender ( $p = 0.004$ , OR- 15.333 95%CI 1.339–175.54), tumour grade ( $p = 0.005$ ,  $r = 0.197$ , OR-3.913 95%CI 1.421–10.776) and clinical prognostic staging ( $p = 0.014$ ,  $r = 0.182$ , OR-3.028 95%CI- 1.217–7.536). There was no association between CD44+/CD24-/ALDH1+ and the molecular subtypes (Table 2).

There were significant inverse associations between CD44+/CD24+/ALDH1+ and grade ( $p = 0.003$ ,  $r = -0.206$ , OR-0.281, 95%CI 0.116–0.683) and mitotic count ( $p = 0.023$ ,  $r = -0.182$ , OR-0.438 95%CI-0.212-0.909) (Table 6)

CD44-/CD24-/ALDH1+ combination phenotype was only associated with mitosis ( $p = 0.017$ ,  $r = 0.192$ ) and NPI ( $p = 0.028$ ,  $r = 0.212$ ). There was however no association between CD44-/CD24+/ALDH1+ and any of the clinicopathological parameters (Table 6).

This study recorded no association between the combination phenotypes and the molecular subtypes ( $p = 0.555$ ).

Table 1  
ALDH1A1 Cytoplasmic expression and its relationship with clinicopathological features

Parameters	ALDH1A1 Cytoplasmic expression		Significance	
	Negative(%)	Positive (%)	OR r	p Value (95% CI)
<b>Patient's age (n = 216)</b>	11(10.6)	93(89.4)	1.171	0.024 0.731
< 50	10 (9.2)	99(90.8)		
≥ 50				
<b>Grade (n = 206)</b>	3(12.5)	21(87.5)	1.714	0.056 0.421
1	14(7.7)	168(92.3)		
2&3				
<b>Tumour size (n = 190)</b>	1(6.7)	14(93.3)	0.664	-0.028 0.699
≤ 2	17(9.7)	158(90.3)		
> 2				
<b>Gender (n = 215)</b>	0(0.0)	3 (100.0)	NA	0.566
Male	21(9.9)	191(90.1)		
Female				
<b>Vascular Invasion (n = 170)</b>	14(12.5)	98(87.5)	1.514	0.058 0.447
Yes	5(8.6)	53(91.4)		
No				
<b>LN stage (n = 108)</b>	5(19.2)	21(80.8)	2.551	0.145 0.131
1 (Negative)	7(8.5)	75(91.5)		
2 (Positive)				
<b>Mitosis (n = 159)</b>	7 (12.5)	49 (87.5)	1.492	0.060 0.451
≤ 10	9 (8.7)	94 (91.3)		
> 10				
<b>NPI (n = 109)</b>	4(12.1)	29(87.9)	1.172	0.023 0.807
Moderate to Good NPI(< 3.4–5.4)	8(10.5)	68(89.5)		
Poor NPI (>/=5.41)				
<b>Clinical Prognostic stage(n = 184)</b>	5(7.5)	60(92.3)	0.818	-0.026 0.721
I&II	11(9.2)	108(90.8)		
III				

Table 2  
Associations between ALDH1A1 expression, Hormone receptor status, Her2, CD24 and CD44 status

Marker	ALDH1A1 cytoplasmic expression			Significance	
	Negative (%)	Positive(%)	OR (95%CI)	r	p value
ER	2(10.0)	67(34.7)	1.0	- 0.145	0.025
Positive	18(90.0)	126(65.3)	0.209 (0.047–0.928)		
Negative					
PR	4(19.0)	67(34.7)	1.0	- 0.099	0.148
Positive	17(81.0)	126(65.3)	0.442 (0.143–1.368)		
Negative					
Her2	3(15.8)	35(19.0)	1.0	-0.024	0.731
Positive	16(84.2)	149(81.0)	0.798 (0.220–2.891)		
Negative					
CD24	9(45.0)	11(55.0)	1.0	0.228	0.001
Negative	29(15.1)	163(84.9)	4.599 (1.751–12.075)		
Positive					
CD44	6(28.6)	15(71.4)	1.0	0.228	0.001
Negative	13(6.7)	180(93.3)	5.538 (1.841–16.662)		
Positive					
CD44 <sup>+</sup> /CD24 <sup>-</sup>	6(30.0)	14(70.0)	1.0	0.134	0.052
Negative	26(13.6)	165(86.4)	2.720 (0.959–7.710)		
Positive					
Molecular Subtypes	4(20.0)	64(36.7)	1	0.058	0.080
Luminal A	0(0.0)	21(11.2)	NA	-0.165	
Luminal B	4(20.0)	18(9.6)	1.481(0.428–5.128)		
Her2+	12(60.0)	80(42.6)	0.0386(0.119–1.253)		
Triple Neg					

Table 3  
Association between ALDH1A1 expression and Histologic types

Histologic type	ALDH1A1 H score set		Total	p-value
	Negative (</=40)	Positive > 40		
Invasive Ductal carcinoma (NOS)	17(80.0)	179(92.3)	195	0.066
invasive lobular carcinoma	2(10.0)	6(3.1)	8	0.121
Intracystic papillary carcinoma	0(0.0)	1(0.5)	1	0.748
DCIS	0(0.0)	2(1.0)	2	0.648
Malignant phyllodes	1(5.0)	1(0.5)	2	0.047*
Metaplastic carcinosarcoma	0(0)	1(0.5)	1	0.748
Invasive colloid carcinoma	1(5.0)	2(0.9)	3	0.151
Lymphocytic Lymphoma	0(0.0)	1(0.5)	1	0.748
Tubular Carcinoma	0(0.0)	1(0.5)	1	0.748
<b>Total</b>	<b>20</b>	<b>194</b>	<b>214</b>	

Table 4  
Frequency distribution of combination phenotypes

Phenotypes	Frequency	Percent
CD44+/CD24-/ALDH+	26	12.3
CD44-/CD24-/ALDH1+	3	1.4
CD44-/CD24+/ALDH1+	10	4.7
CD44+/CD24+/ALDH1+	152	72.0
CD44+/CD24-/ALDH1-	7	3.3
CD44-/CD24-/ALDH1-	3	1.4
CD44-/CD24+/ALDH1-	2	.9
CD44+/CD24+/ALDH1-	8	3.8
<b>Total</b>	<b>211</b>	<b>100.0</b>

Table 5: The distribution of CD24/CD44/ALDH1 combination phenotypes across molecular subtypes

Hormonal status	CD44+/CD24-/ALDH+	CD44-/CD24-/ALDH1+	CD44-/CD24+/ALDH1+	CD44+/CD24+/ALDH1+	CD44+/CD24-/ALDH1-	CD44-/CD24-/ALDH1-	CD44-/CD24+/ALDH1-	CD44+/CD24+/ALDH1-	Total
Luminal A	11(45.8)	2(66.7)	3(33.3)	52(34.7)	1(16.7)	1(33.3)	1(50.0)	1(12.5)	72
Luminal B	2(8.3)	0(0.0)	3(33.3)	16(10.7)	0(0.0)	0(0.0)	0(0.0)	0	21
Her2+	2(8.3)	0(0.0)	0	16(10.7)	0(0.0)	0(0.0)	1(50.0)	3(37.5)	22
Triple Negative	9(37.5)	1(33.3)	3(33.3)	66(44.0)	5(83.3)	2(66.7)	0(0.0)	4(50.0)	90
<b>Total</b>	<b>24</b>	<b>3</b>	<b>9</b>	<b>150</b>	<b>6</b>	<b>3</b>	<b>2</b>	<b>8</b>	<b>205</b>

Table 6  
Association between CD44/CD24/ALDH1 combination phenotypes and clinicopathological features

Parameter		%CD44 <sup>+</sup> CD24 <sup>-</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>+</sup> CD24 <sup>+</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>-</sup> CD24 <sup>-</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>-</sup> CD24 <sup>+</sup> ALDH1 <sup>+</sup>
Age	< 50(n = 100)	69.2(n = 18)	45.3(n = 68)	33.3(n = 1)	30.0(n = 3)
	≥ 50 (n = 109)	30.8(n = 8)	54.7(n = 82)	66.7(n = 2)	70.0(n = 7)
	OR	2.771(1.147–6.697)	0.700(0.338–1.281)	0.540(0.048–6.053)	0.451(0.113–1.793)
	r	0.161	-0.080	-0.035	-0.080
	p	0.020*	0.246	0.612	0.247
Grade	1(n = 23)	28.0(n = 7)	7.4(n = 11)	33.3(n = 1)	10.0(n = 1)
	2&3(n = 179)	72.0 (n = 18)	92.6(n = 137)	66.7(n = 2)	90.0(n = 9)
	OR	3.913(1.421–10.776)	0.281(0.116–0.683)	4.023(0.350-46.199)	0.859(0.104–7.104)
	r	0.197	-0.206	0.085	-0.010
	p	0.005*	0.003*	0.228	0.887
Tumour size	≤ 2(n = 15)	16.7(n = 4)	6.7(n = 9)	0.0(n = 0)	10.0(n = 1)
	> 2(n = 173)	83.3(n = 20)	93.3(n = 125)	100.0(n = 3)	90.0 (n = 9)
	OR	2.782(0.809–9.571)	0.576(0.195–1.705)	-0.037	1.302(0.154–11.028)
	r	0.123	-0.073	0.607	0.018
	P	0.093	0.314		0.808
Gender	Male(n = 3)	7.7(n = 2)	0.7(n = 1)	0.0(n = 0)	0.0(n = 0)
	Female(n = 208)	92.3(n = 24)	99.3(n = 151)	100.0(n = 3)	100(n = 10)
	OR	15.333(1.339–175.54)	0.189(0.017–2.122)	0.834	0.670
	p	0.004*	0.130		
Vascular Invasion	Present (n = 103)	68.8 (n = 11)	63.6(n = 77)	66.7(n = 2)	87.5(n = 7)
	Absent (n = 55)	31.3 (n = 5)	36.4(n = 44)	33.3(n = 1)	12.5 (n = 1)
	OR	1.133(0.374–3.438)	0.636(0.298–1.357)	1.019(0.090-11.481)	3.738(0.448–31.163)
	r	0.017	-0.091	0.001	0.101
	p	0.825	0.240	0.988	0.193
Tumour weight	< 1000 (n = 43)	40.0(n = 2)	45.7(n = 32)	100.0(n = 2)	60.0(n = 3)
	≥ 1000 (n = 49)	60.0(n = 3)	54.3(n = 38)	0.0(n = 0)	40.0(n = 2)
	OR	0.748 (0.119-4.70)	0.842(0.323–2.197)	0.159	1.763(0.280-11.078)
	r	-0.032	-0.307	0.063	0.064
	P	0.756	0.725		0.541

Parameter		%CD44 <sup>+</sup> CD24 <sup>-</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>+</sup> CD24 <sup>+</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>-</sup> CD24 <sup>-</sup> ALDH1 <sup>+</sup>	%ofCD44 <sup>-</sup> CD24 <sup>+</sup> ALDH1 <sup>+</sup>
Lymph Node stage	Negative(n = 26)	22.2(n = 2)	21.3(n = 17)	50.0(n = 1)	25.0(n = 1)
	Positive (n = 81)	77.8(n = 7)	78.8(n = 63)	50.0(n = 1)	75.0(n = 3)
	OR	0.881 (0.171–4.530)	0.540(0.206–1.414)	3.200(0.193–53.043)	1.040(0.103–10.452)
	r	-0.015	-0.122	0.083	0.003
	P	0.879	0.206	0.392	0.973
Mitosis	< 10(n = 54)	44.4(n = 8)	29.5(n = 33)	100(n = 3)	42.9(n = 3)
	≥ 10(n = 101)	55.6(n = 10)	70.5(n = 79)	0.0(n = 0)	57.1(n = 4)
	OR	1.583 (0.585–4.281)	0.438(0.212–0.902)	0.192	1.426(0.307–6.620)
	r	0.073	-0.182	0.017*	0.037
	p	0.363	0.023*		0.649
Nottingham Prognostic Index	< 3.4–5.4 (n = 32)	30.0 (n = 3)	27.5(n = 22)	100.0(n = 2)	25.0(n = 1)
	> 5.4 (n = 76)	70.0 (n = 7)	72.5(n = 58)	0.0(n = 0)	75.0(n = 3)
	OR	1.020(0.246–4.220)	0.683(0.273–1.706)	0.212	0.785(0.079–7.844)
	r	0.003	-0.079	0.028*	-0.020
	P	0.485	0.413		0.863
Triple Negative status	Present (n = 91)	73.3(n = 66)	73.3(n = 66)	1.3(n = 1)	3.3(n = 3)
	Absent (n = 116)	26.7(n = 24)	26.7(n = 24)	98.7(n = 89)	96.7(n = 87)
	OR	0.741(0.308–1.780)	1.015(0.544–1.892)	0.635(0.057–7.114)	0.626(0.152–2.577)
	r	-0.047	0.003	-0.026	-0.046
	p value	0.501	0.963	0.710	0.513
Clinical prognostic staging	I & II (n = 65)	59.1(n = 13)	32.8 (n = 44)	0.0(n = 0)	33.3(n = 3)
	III (n = 118)	40.9(n = 9)	67.2 (n = 90)	100(n = 3)	66.7(n = 62)
	OR	3.028 (1.217–7.536)	0.652(0.333–1.275)	-0.096	0.903(0.218–3.738)
	r	0.182	-0.093	0.195	-0.010
	p value	0.014*	0.210		0.888

Histologic type * CD44 + CD24-/ALDH1 + Crosstabulation						
Histologic type	CD44 + CD24-/ALDH1+		Total	OR	r	P value
	Yes	No				
Invasive carcinoma (NOS)	23(85.5)	169(91.8)	192	0.680 (0.183–2.532)	-0.040	0.564
Invasive lobular carcinoma	0(0.0)	8(4.3)	8		-0.075	0.278
Intracystic papillary carcinoma	0(0.0)	1(0.5)	1		-0.026	0.706
DCIS	1(3.8)	0(0.0)	1		0.184	0.008*
Malignant phyllodes	1(3.8)	1(0.5)	2	7.320(0.444-120.753)	0.112	0.105
Metaplastic carcinosarcoma	0(0.0)	1(0.5)	1		-0.026	0.706
invasive colloid carcinoma	0(0.0)	3(1.6)	3		-0.045	0.512
Lymphocytic Lymphoma	1(3.8)	0(0.0)	1		0.184	0.008*
Tubular Carcinoma	0(0.0)	1(0.5)	1		-0.026	0.706
Total	26	184	210			

## Discussion

There is considerable evidence of ALDH1 as a breast cancer stem cell marker marker(5, 23, 24). Breast cancer in Africans and those of African descent are known to be very aggressive with poor prognosis. This study explored the expression of ALDH1A1 and its association with CD44<sup>+</sup>/CD24<sup>-/low</sup> and the effects of the combined phenotype (CD44<sup>+</sup>/CD24<sup>-/low</sup>/ALDH1<sup>+</sup>) on the molecular and clinicopathological features of breast cancers in an African population.

A high expression of ALDH1A1 of about 90% was recorded in this current study. This percentage is higher than what has been reported in some earlier studies(8, 25, 26) but comparable to Althobiti *et al* and Pan *et al*'s studies which reported 71% and 93% ALDH1 expression respectively(10, 27). This high expression is however not consistent with the general assertion that CSC represent a minute subpopulation of cells in the tumour microenvironment(28, 29). Higher expression of ALDH1 has been associated with Triple negative breast cancer(16, 30) and hence might explain the high expression in our cohort with a relatively high triple negative prevalence of 44.3%. This disparity may also be attributable to the varied range of cut off points and scoring systems employed by various investigators. This study used the semi quantitative H scoring method widely accepted both in research and in clinical practice.

It was observed from our African cohort that ALDH1A1 as a stand-alone breast cancer stem cell marker was not associated with any of the clinicopathological parameters. This finding was contrary to what has been reported some literature. For instance, Yao *et al*, analysed 137 paraffin embedded breast tissues and found an association with tumour grade, size and node metastasis in an Asian cohort(16). A relatively larger Caucasian study (n = 930) recently published in early stage invasive breast cancer also revealed ALDH1A1 associated with high grade, high mitotic count, increased nuclear pleomorphism, poor NPI, advanced nodal stage ( $\geq 4$  positive nodes) and lympho-vascular invasion. However at the protein level of this same paper, there was no association transcriptionally(27). Furthermore, our finding also differed from an African study by Nalwoga and colleagues, who analysed 192 breast carcinomas in Uganda and had associations with tumour grade, high mitotic count, and high nuclear grade(31). In their study however, there was no evidence of the use of the more specific ALDH1A1 isozyme used in this study which may account for the disparity.

Despite the non-association of ALDH1A1 with the clinicopathological parameters, this study recorded an association with ER negativity in keeping with Nalwoga *et al*'s but not with PR and Her2 status as the latter concluded. There was no association between ALDH1 and triple negative breast cancer, a non-consistent finding to Nalwoga *et al*'s study(31). The difference in findings might also be attributable to the different scoring systems and the cut off points for the IHC.

Although ALDH1 as a stand-alone BCSC marker was not associated with clinicopathological parameters the combination phenotype CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> was associated with increased age, higher histological grade, and higher clinical prognostic staging. This goes to affirm the assertion of CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> phenotype having a higher tumourigenic potential as evidenced by high tumour grade(32) and poor prognosis(30). It was interesting to note that CD44<sup>-</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> phenotype also had features of aggression as an association was found with higher mitotic count and higher NPI (Table 6); a finding not yet reported in any literature to the best of our knowledge. This implies that in the absence of CD44 and CD24 expression, ALDH1A1 shows features of aggression and poor prognosis. To speculate, the expression of CD44 and CD24 may hinder or neutralize the aggressiveness conferred by ALDH1A1.

CD44<sup>+</sup>/CD24<sup>+</sup>/ALDH1<sup>+</sup> was the only phenotype among all combination phenotypes that did not relate to any feature of aggressiveness as it was associated with lower histological grade and less mitotic count (Table 6).

## Conclusion

This study has established that a high expression of ALDH1A1 in breast cancers in Africans and this is associated with adverse clinicopathological parameters. It further confirms the increased tumourigenic potential of the CD44<sup>+</sup>/CD24<sup>-</sup>/ALDH1<sup>+</sup> phenotype. The high expression of ALDH1A1 in Africans and its association with poor clinicopathological features makes it an important marker for targeted therapies among this race which is notable for more aggressive breast cancer. It is therefore highly recommended that survival studies are conducted in African populations to help further understand the role of ALDH1 in contributing to breast tumour aggressiveness.

## Declarations

### ETHICAL CONSIDERATION

This article does not contain any studies with human participants or animals by any of the authors. Archival blocks were used with no contact with human participants.

### CONSENT FOR PUBLICATION

'Not applicable'

### AVAILABILITY OF DATA

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### COMPETING INTEREST

The authors declare that they have no competing interests

### FUNDING

Commonwealth Scholarship Commission offered a split-site PhD scholarship to corresponding author to undertake this research. There has been no other financial support for other authors for this work that could have influenced its outcome.

### ACKNOWLEDGEMENT

We thank the University of Nottingham, Division of Cancer Stem Cell, for provision of laboratory and equipments for this work. We thank Commonwealth Scholarship Commission for the offer of Split-site commonwealth scholarship to undertake the project. Special thanks to Chris Nolan and Holly Nicholls for their immense technical support in the project.

### AUTHOR CONTRIBUTIONS

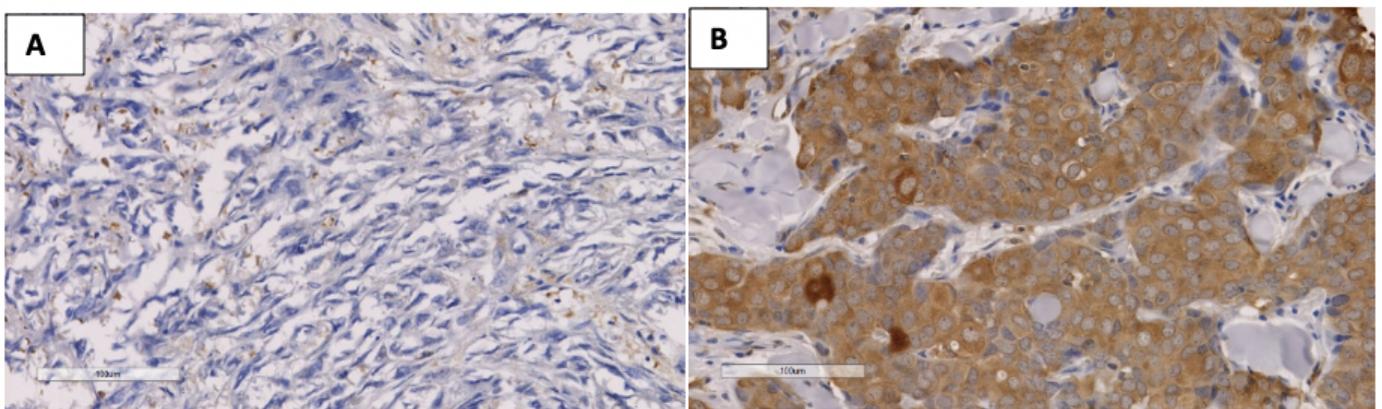
**EG:** laboratory work, writing the manuscript, data analysis and interpretation. **LA:** reviewing the manuscript. **PKA and LDK:** contributed by double scoring and reviewing the manuscript. **MST:** Laboratory work and reviewing the manuscript. **GAR:** reviewing the manuscript. **AJ:** reviewing the manuscript. **WO:** reviewing the manuscript **AG:** reviewing the manuscript.

## References

1. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. *Cell Stem Cell*. 2007;1(5):555–67.
2. Li W, Ma H, Zhang J, Zhu L, Wang C, Yang Y. Unraveling the roles of CD44/CD24 and ALDH1 as cancer stem cell markers in tumorigenesis and metastasis. *Scientific reports*. 2017;7(1):13856.
3. Chute JP, Muramoto GG, Whitesides J, Colvin M, Safi R, Chao NJ, et al. Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells. *Proc Natl Acad Sci USA*. 2006;103(31):11707–12.
4. Martín M, Hinojar A, Cerezo L, García J, Lopez M, Prada J, et al. Aldehyde dehydrogenase isoform 1 (ALDH1) expression as a predictor of radiosensitivity in laryngeal cancer. *Clinical translational oncology: official publication of the Federation of Spanish Oncology Societies of the National Cancer Institute of Mexico*. 2016;18(8):825–30.
5. Kim SJ, Kim YS, Jang ED, Seo KJ, Kim JS. Prognostic Impact and Clinicopathological Correlation of CD133 and ALDH1 Expression in Invasive Breast Cancer. *Journal of breast cancer*. 2015;18(4):347–55.
6. Barr MP, MacDonagh L, Gray SG, O'Byrne K, Cuffe S, Finn S. 75P Inhibition and exploitation of aldehyde dehydrogenase 1 (ALDH1) as a cancer stem cell marker in cisplatin resistant NSCLC. *Journal of Thoracic Oncology*. 2016;11(4):87.
7. Zhao Y, Jin X, Li N, Zhu B, Qian J. Expression and Clinical Significance of ALDH1 and Twist in Gastric Adenocarcinoma. *Sichuan Da Xue Xue Bao Yi Xue Ban*. 2016;47(1):54–9.
8. Kida K, Ishikawa T, Yamada A, Shimada K, Narui K, Sugae S, et al. Effect of ALDH1 on prognosis and chemoresistance by breast cancer subtype. *Breast cancer research treatment*. 2016;156(2):261–9.
9. Yoshioka T, Umekita Y, Ohi Y, Souda M, Sagara Y, Sagara Y, et al. Aldehyde dehydrogenase 1 expression is a predictor of poor prognosis in node-positive breast cancers: a long-term follow-up study. *Histopathology*. 2011;58(4):608–16.
10. Pan H, Wu N, Huang Y, Li Q, Liu C, Liang M, et al. Aldehyde dehydrogenase 1 expression correlates with the invasion of breast cancer. *Diagnostic pathology*. 2015;10:66.
11. Dong Y, Bi LR, Xu N, Yang HM, Zhang HT, Ding Y, et al. The expression of aldehyde dehydrogenase 1 in invasive primary breast tumors and axillary lymph node metastases is associated with poor clinical prognosis. *Pathol Res Pract*. 2013;209(9):555–61.
12. Liu Y, Baglia M, Zheng Y, Blot W, Bao PP, Cai H, et al. ALDH1A1 mRNA expression in association with prognosis of triple-negative breast cancer. *Oncotarget*. 2015;6(38):41360–9.
13. Seo AN, Lee HJ, Kim EJ, Jang MH, Kim YJ, Kim JH, et al. Expression of breast cancer stem cell markers as predictors of prognosis and response to trastuzumab in HER2-positive breast cancer. *Br J Cancer*. 2016;114(10):1109–16.
14. Currie MJ, Beardsley BE, Harris GC, Gunningham SP, Dachs GU, Dijkstra B, et al. Immunohistochemical analysis of cancer stem cell markers in invasive breast carcinoma and associated ductal carcinoma in situ: relationships with markers of tumor hypoxia and microvasculature. *Hum Pathol*. 2013;44(3):402–11.
15. Collina F, Di Bonito M, Li Bergolis V, De Laurentiis M, Vitagliano C, Cerrone M, et al. Prognostic Value of Cancer Stem Cells Markers in Triple-Negative Breast Cancer. *BioMed research international*. 2015;2015:158682.
16. Yao J, Jin Q, Wang XD, Zhu HJ, Ni QC. Aldehyde dehydrogenase 1 expression is correlated with poor prognosis in breast cancer. *Medicine*. 2017;96(25):e7171.
17. Marcato P, Dean CA, Giacomantonio CA, Lee PW. Aldehyde dehydrogenase: its role as a cancer stem cell marker comes down to the specific isoform. *Cell cycle (Georgetown Tex)*. 2011;10(9):1378–84.
18. Phillips TM, McBride WH, Pajonk F. The response of CD24(-/low)/CD44 + breast cancer-initiating cells to radiation. *J Natl Cancer Inst*. 2006;98(24):1777–85.
19. Li X, Lewis MT, Huang J, Gutierrez C, Osborne CK, Wu MF, et al. Intrinsic resistance of tumorigenic breast cancer cells to chemotherapy. *J Natl Cancer Inst*. 2008;100(9):672–9.
20. Fillmore CM, Kuperwasser C. Human breast cancer cell lines contain stem-like cells that self-renew, give rise to phenotypically diverse progeny and survive chemotherapy. *Breast Cancer Res*. 2008;10(2):R25.
21. Ginestier C, Hur MH, Charafe-Jauffret E, Monville F, Dutcher J, Brown M, et al. ALDH1 Is a Marker of Normal and Malignant Human Mammary Stem Cells and a Predictor of Poor Clinical Outcome. *Cell Stem Cell*. 2007;1(5):555–67.

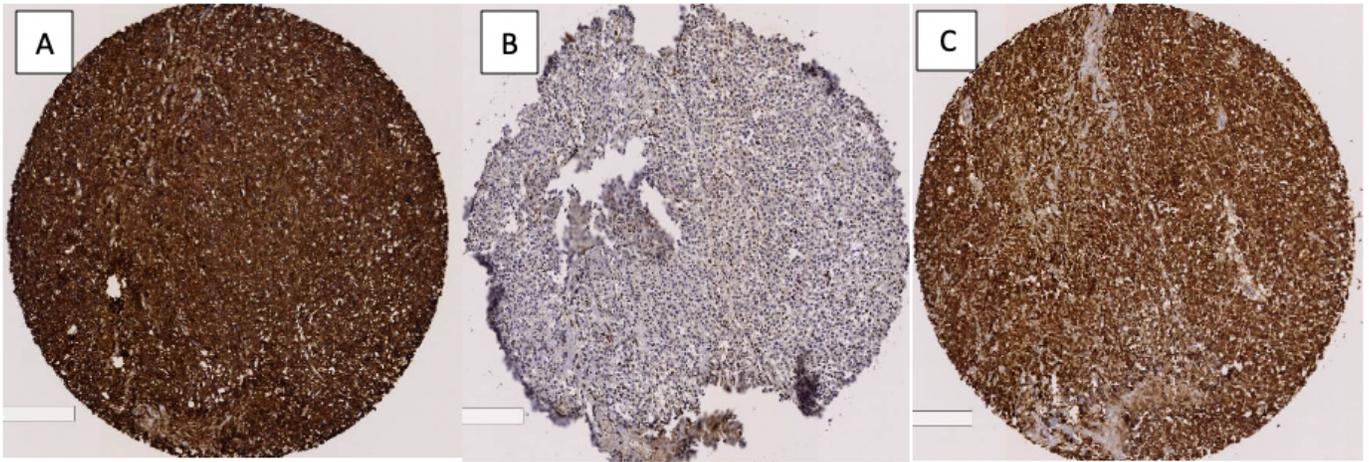
22. Gyan E, Derkyi-Kwarteng L, Brown AA, Derkyi-Kwarteng A, Darkwa AA, Quayson S, et al. Benign breast conditions: An eight-year single-centre histopathological review of women presenting with mass lesions at the Korle-Bu Teaching Hospital, Ghana. *Annals of Diagnostic Pathology*. 2019;42:33–8.
23. Lv X, Wang Y, Song Y, Pang X, Li H. Association between ALDH1+/CD133 + stem-like cells and tumor angiogenesis in invasive ductal breast carcinoma. *Oncology letters*. 2016;11(3):1750–6.
24. Tomita H, Tanaka K, Tanaka T, Hara A. Aldehyde dehydrogenase 1A1 in stem cells and cancer. *Oncotarget*. 2016;7(10):11018–32.
25. Miyoshi Y, Shien T, Ogiya A, Ishida N, Yamazaki K, Horii R, et al. Differences in expression of the cancer stem cell marker aldehyde dehydrogenase 1 among estrogen receptor-positive/human epidermal growth factor receptor type 2-negative breast cancer cases with early, late, and no recurrence. *Breast Cancer Res*. 2016;18(1):73.
26. Li H, Ma F, Wang H, Lin C, Fan Y, Zhang X, et al. Stem cell marker aldehyde dehydrogenase 1 (ALDH1)-expressing cells are enriched in triple-negative breast cancer. *Int J Biol Mark*. 2013;28(4):e357-64.
27. Althobiti M, El Ansari R, Aleskandarany M, Joseph C, Toss MS, Green AR, et al. The prognostic significance of ALDH1A1 expression in early invasive breast cancer. 2020;n/a(n/a).
28. Allan AL, Vantyghem SA, Tuck AB, Chambers AF. Tumor dormancy and cancer stem cells: implications for the biology and treatment of breast cancer metastasis. *Breast Dis*. 2006;26:87–98.
29. Wang W, Quan Y, Fu Q, Liu Y, Liang Y, Wu J, et al. Dynamics between cancer cell subpopulations reveals a model coordinating with both hierarchical and stochastic concepts. *PLoS One*. 2014;9(1):e84654.
30. Ricardo S, Vieira AF, Gerhard R, Leitao D, Pinto R, Cameselle-Teijeiro JF, et al. Breast cancer stem cell markers CD44, CD24 and ALDH1: expression distribution within intrinsic molecular subtype. *J Clin Pathol*. 2011;64(11):937–46.
31. Nalwoga H, Arnes JB, Wabinga H, Akslen LA. Expression of aldehyde dehydrogenase 1 (ALDH1) is associated with basal-like markers and features of aggressive tumours in African breast cancer. *Br J Cancer*. 2010;102(2):369–75.
32. Iris R, Ana Paula Martins S, Rubens Silveira L, Cícero de Andrade U, Eduardo SJ, Karina FA, et al. Cancer stem cell markers ALDH1 and CD44+/CD24– phenotype and their prognosis impact in invasive ductal carcinoma. *Eur J Histochem*. 2018;62:2943.

## Figures



**Figure 1**

A- Negative staining for ALDH1A1, B- Positive staining for ALDH1A1



**Figure 2**

CD44, CD24 and ALDH1A1 expression in breast cancer using IHC. A, B and C Same case stained positive for CD44 in A, negative for CD24 and positive for ALDH1A1