

## REVIEW ARTICLE OPEN



# TGFBR1\*6A as a modifier of breast cancer risk and progression: advances and future prospects

Kojo Agyemang <sup>1</sup>, Allan M. Johansen <sup>1</sup>, Grayson W. Barker <sup>1</sup>, Michael J. Pennison <sup>1</sup>, Kimberly Sheffield <sup>1</sup>, Hugo Jimenez <sup>1</sup>, Carl Blackman <sup>1</sup>, Sambad Sharma <sup>1</sup>, Patrick A. Fordjour<sup>2</sup>, Ravi Singh <sup>1,3</sup>, Katherine L. Cook <sup>1,4</sup>, Hui-Kuan Lin <sup>1,3</sup>, Wei Zhang <sup>1,3,5</sup>, Hui-Wen Lo <sup>1,3</sup>, Kounosuke Watabe<sup>1,3</sup>, Peiqing Sun<sup>1,3</sup>, Carl D. Langefeld <sup>3,6,7</sup> and Boris Pasche <sup>1,3</sup>✉

There is growing evidence that germline mutations in certain genes influence cancer susceptibility, tumor evolution, as well as clinical outcomes. Identification of a disease-causing genetic variant enables testing and diagnosis of at-risk individuals. For breast cancer, several genes such as *BRCA1*, *BRCA2*, *PALB2*, *ATM*, and *CHEK2* act as high- to moderate-penetrance cancer susceptibility genes. Genotyping of these genes informs genetic risk assessment and counseling, as well as treatment and management decisions in the case of high-penetrance genes. *TGFBR1\*6A* (rs11466445) is a common variant of the TGF- $\beta$  receptor type I (*TGFBR1*) that has a global minor allelic frequency (MAF) of 0.051 according to the 1000 Genomes Project Consortium. It is emerging as a high frequency, low penetrance tumor susceptibility allele associated with increased cancer risk among several cancer types. The *TGFBR1\*6A* allele has been associated with increased breast cancer risk in women, OR1.15 (95% CI 1.01–1.31). Functionally, *TGFBR1\*6A* promotes breast cancer cell proliferation, migration, and invasion through the regulation of the ERK pathway and Rho-GTP activation. This review discusses current findings on the genetic, functional, and mechanistic associations between *TGFBR1\*6A* and breast cancer risk and proposes future directions as it relates to genetic association studies and mechanisms of action for tumor growth, metastasis, and immune suppression.

npj Breast Cancer (2022)8:84; <https://doi.org/10.1038/s41523-022-00446-6>

## HERITABLE BREAST CANCER GENES

Heritable predisposition genes are important risk factors for breast cancer susceptibility, accounting for 5.03% of all breast cancer cases<sup>1–4</sup>. Pathogenic variants of high-risk predisposition genes such as *BRCA1* and *BRCA2* are the most widely known and are used in genetic testing and counseling to predict breast cancer risk and clinical outcomes<sup>4,5</sup>. However, these variants are uncommon (<1.3% allelic frequency), and account for less than 1.5% of all breast cancer cases<sup>4,6</sup>. Like other complex genetic traits such as diabetes, obesity, and autoimmune diseases, recent advances in genome-wide association studies reveal that the vast majority of hereditary breast cancer cases are genetically multifactorial<sup>7–10</sup>, involving numerous other polymorphisms of varying penetrance acting as tumor modifier genes<sup>11–13</sup>. Thus, several lower penetrance risk variants are now included in breast cancer susceptibility gene screening panels for testing and counseling. Inclusion of these polymorphisms into screening panels results in a 40 to 50% increase in breast cancer risk detection among women, and a 5 to 15% increase in detection among *BRCA1/2*-negative females<sup>6,11,12,14</sup>.

The Transforming Growth Factor-Beta (TGF- $\beta$ ) signaling pathway plays a critical role during cancer development and progression<sup>15–17</sup>. Variants of the TGF- $\beta$  pathway genes, particularly *TGFBR1\*6A*, have been studied in a large number of females, and have been associated with low penetrance risk for breast cancer<sup>18</sup> (Table 1). This review details our current clinical and pre-clinical knowledge on *TGFBR1\*6A* as a high frequency, low penetrance tumor susceptibility allele, and provides further

rationale to assess its role as a modifier gene for breast cancer predisposition and tumor progression.

## TGFBR1\*6A AND BREAST CANCER

### Identification and characterization

The TGF- $\beta$  pathway functions as a tumor suppressor during cancer development but enhances tumor growth, immune evasion, and metastasis in advanced cancers<sup>15–17</sup>. During TGF- $\beta$  signaling, the TGF- $\beta$  receptor 1 (*TGFBR1*) plays a critical role of binding and activating downstream receptor-regulated (R-) SMADs (*SMAD2* and *SMAD3*), co-SMADs (*SMAD4*)<sup>19,20</sup>, and non-SMAD (*MAPK-ERK*, *RAS*, *AKT*, *JNK*, and *RHOA*)<sup>17,21</sup> intermediary proteins. In 1998, Pasche et al. identified *TGFBR1\*6A* (rs11466445) as a polymorphic variant of *TGFBR1* with an in-frame deletion of three GCG codons encoding alanine within exon 1 of the human *TGFBR1* signal peptide sequence<sup>22,23</sup> (Fig. 1). Signal peptides are responsible for intracellular transport, targeting of nascent proteins (secretory and membrane proteins) to the endoplasmic reticulum, and integration of newly translated proteins into their respective compartments. N-terminal signal peptides are usually cleaved off and degraded after insertion and targeting of the nascent protein to the endoplasmic reticulum<sup>24</sup>. Investigations into *TGFBR1\*6A* signal peptide cleavage showed that the 3-alanine deletion from the 9-alanine repeat within the hydrophobic core of the signal peptide does not affect the posttranslational cleavage of a signal peptide. Studies showed that *TGFBR1\*6A* signal peptide is cleaved between Ala30 and Leu31, whereas the wild-type *TGFBR1* is cleaved

<sup>1</sup>Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC 27157-1082, USA. <sup>2</sup>Department of Thoracic Medicine, School of Medicine, University of Adelaide, Adelaide, SA 5000, Australia. <sup>3</sup>Comprehensive Cancer Center, Wake Forest School of Medicine, Winston Salem, NC 27157-1082, USA. <sup>4</sup>Department of Surgery, Wake Forest School of Medicine, Winston-Salem, NC 27157-1082, USA. <sup>5</sup>Center for Cancer Genomics and Precision Oncology, Wake Forest School of Medicine, Winston Salem, NC 27157-1082, USA. <sup>6</sup>Department of Biostatistics and Data Science, Wake Forest School of Medicine, Winston-Salem, NC 27157-1082, USA. <sup>7</sup>Center for Precision Medicine, Wake Forest School of Medicine, Winston-Salem, NC 27157-1082, USA. ✉email: bpasche@wakehealth.edu

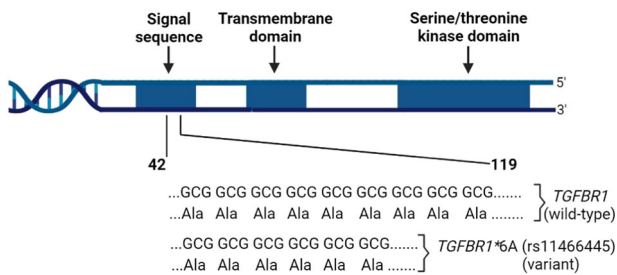
**Table 1.** Case-control studies showing *TGFBR1\*6A* genotypic and allelic distribution and frequency.

Study	Country of study participants	Self-reported race/ethnicity	Genotype distribution and frequency (%)						Allelic frequency	
			Cases			Controls			Cases	Controls
			9/9A	9A/6A	6A/6A	9A/9A	9A/6A	6A/6A	<i>TGFBR1*6A</i>	<i>TGFBR1</i>
Pasche et al. 1999 (a)	US	Mixed	128 (84.2)	24 (15.8)	0 (0)	654 (89.3)	78 (10.7)	0 (0)	0.079	0.053
Pasche et al. 1999 (b)	Northern Italy	Caucasian	39 (81.3)	8 (16.7)	1 (2.1)	38 (76)	12 (24)	0 (0)	0.104	0.120
Baxter et al. 2002	United Kingdom	Caucasian	268 (75.5)	83 (23.4)	4 (1.1)	207 (83.5)	39 (15.7)	2 (0.8)	0.128	0.087
Reiss, 2004	US	Mixed	87 (88.8)	11 (11.2)	0 (0)	77 (84.6)	14 (15.4)	0 (0)	0.056	0.077
Caldes, 2004	Spain	Caucasian	214 (79)	56 (20.7)	1 (0.4)	250 (85.6)	42 (14.4)	0 (0)	0.107	0.072
Offit, 2004	US	NS	391 (84.6)	67 (14.5)	4 (0.9)	291 (88.2)	38 (11.5)	1 (0.3)	0.081	0.061
Northwestern, 2004	US	NS	74 (86.1)	12 (13.9)	0	105 (85.4)	17 (13.8)	1 (0.8)	0.070	0.077
Jin et al. 2004 (a)	Finland	Caucasian	177 (80.1)	38 (17.2)	6 (2.7)	171 (73.1)	60 (25.6)	3 (1.3)	0.113	0.141
Jin et al. 2004 (b)	Poland	Caucasian	140 (82.4)	28 (16.5)	2 (1.2)	176 (87.1)	26 (12.9)	0 (0)	0.094	0.064
Kaklamani et al. 2005	US	Mixed	515 (84.3)	92 (15.1)	4 (0.7)	612 (88.7)	77 (11.2)	1 (0.1)	0.082	0.057
Chen et al. 2006	US	Mixed	92 (80)	23 (20)	0 (0)	111 (85.4)	18 (13.8)	1 (0.8)	0.100	0.077
Feigelson et al. 2006 <sup>#</sup>	US	Mixed	387 (80.5)	94 (19.5) <sup>a</sup>	NS	384 (74)	100 (26) <sup>a</sup>	NS	0.098	0.130
Cox et al. 2007	US	NS	968 (81.6)	207 (17.4)	12 (1)	1352 (80.8)	302 (18.1)	19 (1.1)	0.097	0.102
Song et al. 2007	Sweden	Caucasian	598 (78.4)	152 (19.9)	13 (1.7)	682 (80)	160 (18.8)	10 (1.2)	0.117	0.106
Jakubowska et al. 2009	Poland	Caucasian	282 (88.7)	33 (10.4)	3 (0.9)	252 (86.9)	38 (13.1)	0 (0)	0.061	0.066
Colleran et al. 2009	Ireland	Caucasian	796 (82.9)	154 (16)	10 (1)	785 (81.9)	160 (16.7)	13 (1.4)	0.091	0.097
Joshi 2011 (a)	India	Asian	163 (97.6)	4 (2.4)	0 (0)	213 (95.9)	9 (4.1)	0 (0)	0.012	0.020
Joshi 2011 (b)	India	Asian	33 (78.6)	8 (19)	1 (2.4)	148 (87.6)	19 (11.2)	2 (1.2)	0.119	0.068
Kamali et al. 2015 <sup>#</sup>	Iran	Middle Eastern	251 (89.6)	25 (8.9)	4 (1.4)	241 (86.1)	27 (9.6)	12 (4.3)	0.059	0.091

Case-control studies from 1999 to date have included a total of 14,837 participants (6787 cases /8050 controls); of which 13,312 (6026 cases/7286 controls) were in Hardy–Weinberg (HW) equilibrium. The data shows genotype distribution and allelic frequency (%) in the order *TGFBR1\*9A/9A* > *TGFBR1\*9A/6A* > *TGFBR1\*6A/6A*.

NS not stated.

<sup>#</sup>Study population not in Hardy–Weinberg equilibrium.



**Fig. 1** *TGFBR1* and *TGFBR1\*6A* gene and protein sequences. Sequence analyses reveal nine (9) GCG/alanine repeats within nucleotides 42–119 of the *TGFBR1* signal sequence. *TGFBR1\*6A* variant has six (6) GCG/alanine repeats in its signal sequence (Pasche, Luo et al. 1998, Pasche, Kolachana et al. 1999).

between Ala33 and Leu34. Importantly, the mature forms of the *TGFBR1\*6A* and *TGFBR1* receptors are identical and the TGF- $\beta$  ligand binds to each receptor with the same affinity<sup>23,25</sup>. Also, the differences in cleavage sites do not influence *TGFBR1\*6A* protein targeting and translocation functions<sup>25</sup> nor its sensitivity, and half-life<sup>23</sup>. Despite *TGFBR1\*6A* and *TGFBR1* physicochemical similarities, *TGFBR1\*6A* is intriguingly associated with breast cancer risk<sup>26,27</sup>, and promotion of cell growth, migration, and invasion<sup>25,28</sup>. Only the released signal peptide separates *TGFBR1\*6A* from *TGFBR1*, which strongly suggests that *TGFBR1\*6A* signal peptide contributes to oncogenesis.

## GENETIC ASSOCIATION WITH BREAST CANCER RISK

### Case-control studies

**Allelic frequency.** *TGFBR1\*6A* (rs11466445) is a common variant of the TGF- $\beta$  receptor type I (*TGFBR1*) with a study-wide minor allelic frequency (MAF) of 0.051 in the 1000 Genomes Project Consortium, ranging from 0.0079 among East Asians to 0.0975 among Europeans<sup>29–31</sup>. Nineteen case-control studies that included 14,837 participants (6787 cases/8,050 controls) have investigated the association of *TGFBR1\*6A* with breast cancer risk (Table 1). Seventeen of these studies, which included 13,312 individuals (6026 cases/7286 controls), show the *TGFBR1\*6A* genotype frequencies were in Hardy–Weinberg equilibrium (HWE). HWE is a useful genotyping quality control metric that relates allele frequencies to genotype frequencies with an expected assumption that genotype frequencies will remain constant in a randomly mating population. The afore-mentioned *TGFBR1\*6A* case/control studies consisted predominantly of Caucasian individuals of European ancestry (>80%), but also included modest numbers of individuals from other self-reported races/ethnicities (e.g., African American, Hispanic, Asian, and Middle Eastern). Early studies examined a mixed sample of Caucasians, African Americans, Hispanics, and Asians from the US, and Northern Italy, and found an allelic frequency of 7.9% in breast cancer patients compared to 5.3% in healthy controls<sup>23</sup>. Several other studies from the US consisting of >80% Caucasians and <20% non-Caucasians reported similar *TGFBR1\*6A* allelic frequency of 5.6–10.0% among breast cancer patients<sup>26,32–35</sup>. The *TGFBR1\*6A* allelic frequency also ranged from 6.1 to 12.8% among breast cancer patients from northern and southwestern Europe,

including samples from the UK<sup>36</sup>, Spain<sup>32</sup>, Sweden<sup>37</sup>, Finland<sup>38</sup>, Poland<sup>38,39</sup>, and Ireland<sup>40</sup> (Table 1). In all, Asian and Middle Eastern women, particularly those from Western India had the lowest *TGFBR1\*6A* allelic frequency, 1.2–11.9% among cases<sup>41</sup>. In all, *TGFBR1\*6A* genotype frequency is in the order 9A/9A > 9A/6A > 6A/6A in all populations studied to date (Table 1)<sup>23,38,41</sup>.

**Risk association.** Published case-control studies examining the association of *TGFBR1\*6A* with breast cancer risk among individuals from different geographical locations and ethnicities show both significant and non-significant risk associations<sup>26,41,42</sup>. It remains unclear if the differences in *TGFBR1\*6A* risk association correlate with geographical location, ethnicity, age, tumor stage, and other confounding factors such as other polymorphism frequencies, lifestyle, and environment. In 2002, Baxter et al. reported the first association between *TGFBR1\*6A* and breast cancer risk (OR 1.6, 95% CI 1.1–2.5)<sup>36</sup>. Cases and controls were residents of Southampton, UK. Controls were healthy females. Breast cancer cases were selected based on age at onset under 40 years, family history of breast cancer irrespective of age at onset, or bilateral breast cancer irrespective of family history or age at onset. The study noted that *TGFBR1\*6A* allelic frequency among Caucasians from the UK does not differ by age of onset (<40 years), bilateral breast cancer, family history, and germline mutations of *BRCA1* and *BRCA2*<sup>36</sup>. In 2004, Caldes investigated breast cancer patients of Caucasian descent only from Madrid, Spain. The study found a significant association between *TGFBR1\*6A* and breast cancer risk (OR 1.55, 95% CI 1.02–2.34)<sup>32</sup>. In the Kaklamani 2005<sup>26</sup> study that included >80% Caucasians from New York, NY, *TGFBR1\*6A* association with breast cancer risk was significant under both dominant (OR 1.50, 95% CI 1.07–2.11) and additive (OR 1.46, 95% CI 1.04–2.06) models. Breast cancer risk for women aged >50 years was higher (OR 2.20, 95% CI 1.25–3.87), than for women aged ≤50 years (OR 1.18, 95% CI 0.75–1.84). However, further subgroup analyses show no association between *TGFBR1\*6A* and ER/PR status or cancer stage at diagnosis. In this study, both breast cancer patients and healthy controls were matched for age, gender, and location<sup>26</sup>. Other studies that included participants of similarly mixed ethnicities but from other states in the United States did not confirm an association between *TGFBR1\*6A* and breast cancer risk<sup>32,33,35</sup> (Fig. 2a). This indicates that *TGFBR1\*6A* association with breast cancer risk may have a more modest effect size than in the original reports (regression toward the mean) and these studies were not powered to account for the effects of other confounding/modifying factors apart from ethnicity and geographical location that modify the magnitude of the risk attributable to the *TGFBR1\*6A* polymorphism.

Intriguingly, other studies have investigated participants of similar Caucasian backgrounds from other European countries (Sweden, Ireland, Finland, and Poland) and found no significant associations (Fig. 2a)<sup>37–40</sup>. In a case-control study among Swedish Caucasians<sup>37</sup>, the study further enriched its population sampling for genetic susceptibility by recruiting patients with family history of bilateral breast cancer cases. Overall, *TGFBR1\*6A* carriership was not associated with breast cancer risk (OR 1.12, 95% CI 0.90–1.39). However, subgroup analyses showed that *TGFBR1\*6A* was associated with increased risk among low-risk familial breast cancer patients with one first- or second-degree breast cancer relative (OR 1.3, 95% CI 1.0–1.9). In that study, *TGFBR1\*6A* also correlated with higher tumor grade (OR 2.27 (95% CI 1.01–5.11) but had no association with tumor stage or ER/PR status<sup>37</sup>. For women of Asian and Middle Eastern descent, *TGFBR1\*6A* association with breast cancer risk was neither significant nor protective<sup>41,42</sup>. Among Asians from India, Joshi 2011<sup>41</sup> found no association among Western Indians (OR 0.69, 95% CI 0.18–1.92) and those from the Parsi community (OR 1.85, 95% CI 0.84–4.05)<sup>41</sup>. Although these are important studies, the sample sizes are modest, and the lower allelic frequency in this population<sup>43</sup> results

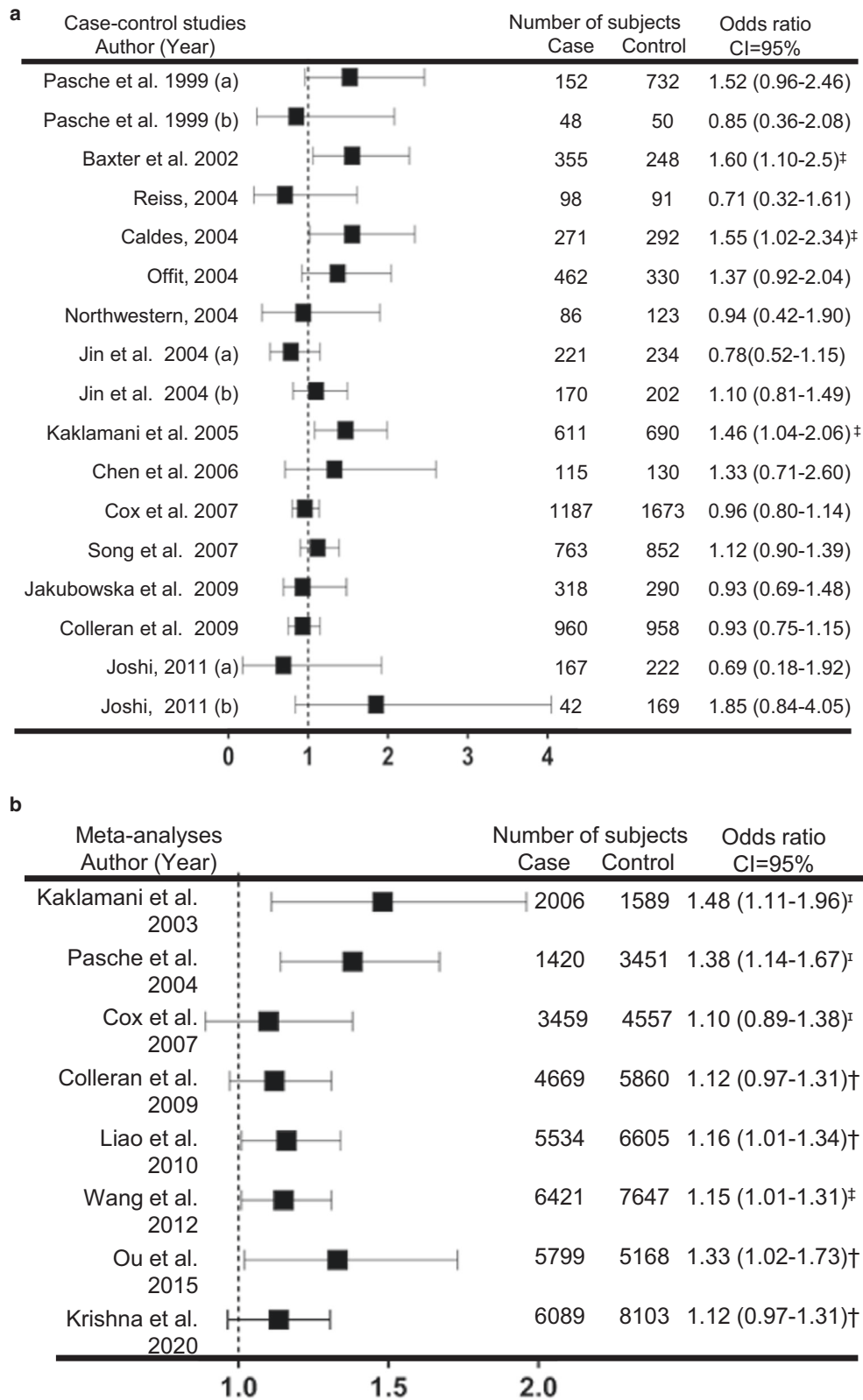
in very low a priori statistical power to detect an association of the size previously reported. To date, *TGFBR1\*6A* appears to be associated with decreased breast cancer risk only in Middle Eastern women, particularly in Iran (OR 0.62, 95% CI 0.39–0.98)<sup>42</sup>. However, it is worth noting that the genotype frequencies in this sample of 280 cases and 280 controls show strong deviations from Hardy–Weinberg equilibrium expectations ( $\chi^2$   $p$  value = 2.81E-12), an important measure of genotyping quality. Specifically, there were an excess of both *TGFBR1\*6A* and *TGFBR1* homozygotes, with a corresponding deficit of heterozygotes, when compared to HWE expectations.

In summary, it appears that *TGFBR1\*6A* is a common breast cancer susceptibility allele in diverse ethnic backgrounds and its influence may vary by ancestry, geographical location, and other risk modifying factors. Large case/control studies, with important related phenotypes, paralleling the magnitude of other complex genetic traits, and meta-analyses that incorporate strict genotype quality control assessment will be necessary to clarify its population-specific subgroup risk associations.

### META-ANALYSES OF *TGFBR1\*6A* ASSOCIATION WITH BREAST CANCER RISK

To derive a more precise estimate of *TGFBR1\*6A* association with breast cancer risk, several meta-analyses have investigated its risk association in up to 14,837 participants (6787 cases/8050 controls) from 19 case-control studies. In all, eight meta-analyses have studied the association of *TGFBR1\*6A* with breast cancer risk (Fig. 2b). Five of the studies found evidence that supports the association between *TGFBR1\*6A* and breast cancer risk. Early meta-analyses by ref. 44 and ref. 32 that analyzed up to 12 case-control studies (4871 subjects) found an association of *TGFBR1\*6A* with breast cancer risk with odds ratios of 1.48 (95% CI 1.11–1.96) and 1.38 (95% CI 1.14–1.67), respectively. Subsequent studies examined ten case/control studies and did not find a significant risk association (OR 1.10, 95% CI 0.89–1.38), albeit in a concordant direction. The heterogeneity in the reported results may be partly due to meaningful differences in the genotyping methods employed across these studies for this GCG repeat polymorphism, and increased variation in estimates due to lower sample sizes included in the case/control studies<sup>35</sup>.

Three of the most recent meta-analyses that investigated up to 17 case/control studies (14,068 subjects) found an association between *TGFBR1\*6A* and breast cancer risk. Liao et al. 2010 reported an odds ratio of 1.16 (95% CI 1.01–1.34) among 11,220 case/controls<sup>45</sup>. Wang et al. 2012 investigated 14,068 participants and showed a nearly identical OR of 1.15 (95% CI 1.01–1.31)<sup>27</sup>. Additionally, Ou et al. 2015 found 14.8 and 9.6% *TGFBR1\*6A* allelic distribution among 6275 cases and controls, respectively. There was a significant association of the *TGFBR1\*6A* allele with breast cancer risk (OR 1.33 95% CI 1.02–1.73)<sup>46</sup>. The study also noted that the homozygous *TGFBR1\*6A/6A* is not significantly associated with breast cancer risk. Collieran 2010<sup>40</sup> and Krishna 2020<sup>47</sup> found no significant associations between *TGFBR1\*6A* and breast cancer risk (Fig. 2b). It is interesting to note that the latter two meta-analyses selectively excluded some data from the Pasche 2004, and other case/control studies in their analyses. For example, case/control studies by refs. 32,39, all of which are in Hardy–Weinberg equilibrium were selectively omitted from the ref. 35, ref. 40 and ref. 47 meta-analyses. Collieran et al., 2010 explained that some study populations from Jin et al., Reiss, and Offit had been reported in Kaklamani et al. 2005, and as such omitted them in their analyses. However, these study populations were different and there was no overlap with the population reported in Kaklamani et al. 2005 and prior reports. Additionally, Collieran et al., measured heterogeneity in the case/control studies used in their meta-analyses, and, noted that studies with sample size less than 1000 had the most extreme odds ratios, indicating odds ratio



**Fig. 2 Studies investigating *TGFBR1\*6A* association with breast cancer risk.** Forest plot showing the number of subjects and odds ratios of **a** Case-control studies, and **b** Meta-analyses associating *TGFBR1\*6A* to breast cancer risk. Plot **a** includes only case/control studies that are in Hardy-Weinberg equilibrium. CI confidence interval, † dominant association ( $p \leq 0.01$ ), ‡ additive association ( $p \leq 0.05$ ), † allelic association ( $p \leq 0.05$ ).

heterogeneity among different case/controls as another contributing factor<sup>40</sup>. Overall, most meta-analyses published to date show an association between *TGFBR1\*6A* and breast cancer risk. Further studies are warranted to establish the risk association and its relatedness to different population sub-groups.

## ***TGFBR1\*6A* FUNCTIONAL EFFECTS**

### **Promotion of cell proliferation, migration, and invasion**

Physicochemical studies using R1B/L17 and HEK 293 cells revealed that the mature *TGFBR1\*6A* and the wild-type *TGFBR1* receptors have similar half-life, receptor turnover, and binding affinity to TGF- $\beta$  ligand<sup>23,25</sup>. In MCF-7 breast cancer cells stably transfected with either *TGFBR1* or *TGFBR1\*6A*, TGF- $\beta$ /SMAD signaling was comparable<sup>28</sup>. However, the stably transfected *TGFBR1\*6A* MCF-7 breast cancer cells exhibited enhanced cell growth, migration, and invasion<sup>25,28</sup>. Importantly, *TGFBR1\*6A* switched TGF- $\beta$  anti-proliferative effects into growth stimulatory effects in the MCF-7 breast cancer cells. The *TGFBR1\*6A*-mediated switch to growth stimulation was independent of *TGFBR1\*6A* kinase domain, indicating a mechanism likely due to *TGFBR1\*6A* signal peptide<sup>25</sup>. A similar conclusion was also reached following the investigation of *TGFBR1\*6A* induction of migration and invasion. There was a 1.2-fold and 1.7-fold increase in migration and invasion, respectively, in *TGFBR1\*6A* MCF-7 cells when compared to *TGFBR1* transfected cells<sup>28</sup>. In the MCF-7 cells expressing low *TGFBR1\*6A* levels, there were 1.3 and 1.9 times increase in migration and invasion respectively, when compared to the vector controls. Cells expressing higher and intermediate *TGFBR1\*6A* levels showed 1.8 and 2.2 times increase in migration and invasion, respectively. The induction of migration and invasion observed in the *TGFBR1\*6A* cells were not affected by TGF- $\beta$  stimulation, suggesting an underlying migration mechanism that is TGF- $\beta$  independent<sup>28</sup>. Similar observations of increased growth and invasion were reported in colorectal cancer cells that endogenously harbor *TGFBR1\*6A*. Using SW48 (*TGFBR1/TGFBR1*) and DLD1 (*TGFBR1/TGFBR1\*6A*) cells, TGF- $\beta$  treatment resulted in growth inhibition in the *TGFBR1/TGFBR1* SW48 cells while it resulted in growth stimulation in the *TGFBR1/TGFBR1\*6A* DLD1 cells<sup>25</sup>. Stable transfection of the colorectal cancer cells with *TGFBR1\*6A* also exhibited enhanced proliferation when compared with the vector-transfected cells<sup>48</sup>, indicating that *TGFBR1\*6A* has similar functional outcomes in breast and colorectal cancer.

## **MECHANISM OF ACTION**

The functional and mechanistic understanding of tumor susceptibility alleles are important in establishing their clinical significance. *TGFBR1\*6A* function in breast cancer cells is postulated to be mediated by its cleaved signal peptide, which is 3-alanine shorter than the wild-type *TGFBR1* signal peptide (Figs. 1, 3a)<sup>25</sup>. Thus far, available evidence using SBE4-lux and 3TP-lux luciferase reporters suggest a mechanism that is not mediated by the mature *TGFBR1\*6A* receptor. The SBE4-lux luciferase reporter assesses the binding of activated SMAD2/3/4 with SMAD binding elements in the nucleus as a measure of TGF- $\beta$ /SMAD activation<sup>28,49</sup>. Likewise, 3TP-lux transcription reporter measures the binding of activated SMAD2/3/4 to three consecutive 12-O-tetradecanoylphorbol-13-acetate (TPA) response elements (TREs) and a portion of the plasminogen inhibitor-1 (PAI-1) promoter region<sup>28,50</sup>. In a study that assessed the proliferation of MCF-7 cells expressing intermediate and high levels of *TGFBR1\*6A* with kinase-inactivated domains, *TGFBR1\*6A* enhanced cell proliferation in both clones similar to the kinase-activated *TGFBR1\*6A* MCF-7 clones. Inactivation of the kinase domain did not influence *TGFBR1\*6A* growth stimulation, indicating a cell proliferation enhancing mechanism that is not affected by the *TGFBR1\*6A*

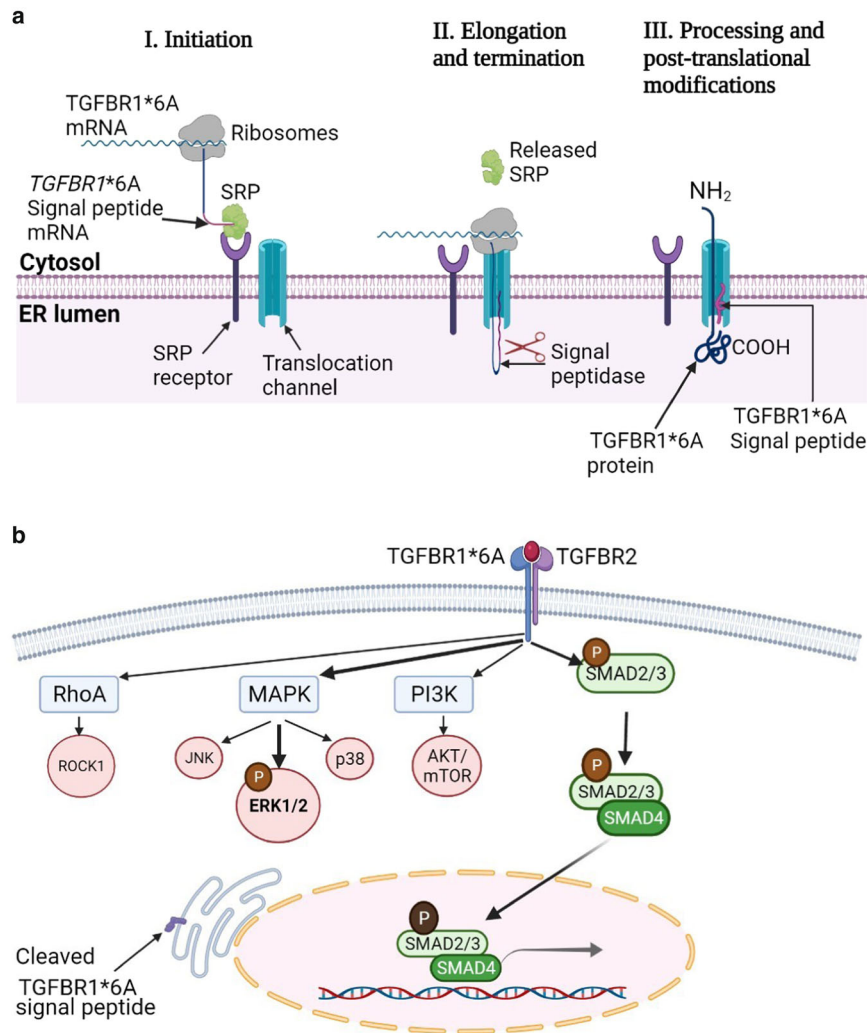
kinase domain. Analyses of the TGF- $\beta$ /SMAD signal transduction system using SBE4-lux and 3TP-lux transcription reporters, and phosphorylated-SMAD2 and SMAD3 levels showed comparable TGF- $\beta$ /SMAD signaling levels for *TGFBR1\*6A* and *TGFBR1* cells<sup>25,28</sup>. The absence of *TGFBR1\*6A* influence on TGF- $\beta$  signaling is consistent with the observed TGF- $\beta$  independent effect on migration and invasion in MCF-7 cells. Investigations into TGF- $\beta$ /non-SMAD pathways revealed that *TGFBR1\*6A* increases ERK1/2 phosphorylation but shows no significant influence on p38 and JNK activation<sup>28</sup>. The upregulation of ERK1/2 phosphorylation is suggestive of MAP-kinase-mediated migration and invasion (Fig. 3b). A similar increase in MAP-kinase activation has also been correlated with increased invasion of *TGFBR1\*6A*-transfected colorectal cancer cells. Zhou et al. showed that *TGFBR1\*6A* expression enhances proliferation and invasion in stably transfected SW48 and DLD1 cells accompanied by increased ERK1/2 phosphorylation<sup>48</sup>. In contrast to breast cancer cells, the study noted an increase in p38 activation in the colorectal cancer cells. In all, the identified functional and mechanistic responses in colorectal cancer cells appears to be TGF- $\beta$ -dependent<sup>48</sup>, indicating cancer-type differences in response mechanisms.

Further investigations into the signaling network of *TGFBR1\*6A*-induced MCF-7 cell migration and invasion using Affymetrix GeneChip Human Genome U133 Plus 2.0 Array identified *ARHGAP5* and *FN1* as the key differentially expressed genes. The levels of *ARHGAP5* and *FN1* were downregulated in *TGFBR1\*6A* MCF-7 cells when assessed with RT-qPCR and western blot assays<sup>28</sup>. *ARHGAP5* encodes the Rho GTPase activating protein 5 that negatively regulates the Rho-family of small GTPases. The Rho-family of small GTPases are small guanosine triphosphatases (GTPases) of the rat sarcoma (Ras) superfamily that function as molecular switches for cytoskeletal remodeling during cell division, cell-cell adhesion, cell contractility, migration, and invasion. *ARHGAP5* (P190B) and *ARHGAP35* (P190A) are the main regulators of the Ras homolog (Rho) family of actin-based regulators and are implicated in cellular adhesion, migration, and invasion<sup>51,52</sup>. *ARHGAP5* silencing inhibits migration and invasion of AGS and MGC-803 gastric adenocarcinoma cells<sup>53</sup>. In colorectal cancer, however, *ARHGAP5* is markedly overexpressed in the liver of metastatic tissues compared to matched primary tumor tissues<sup>54</sup>. *ARHGAP5* suppression leads to decreased wound healing, migration, and invasion in DLD1 and SW480 colorectal cancer cells<sup>54</sup>. Similarly, amplification of *ARHGAP5* on the 14q12 chromosomal locus promotes the spreading and migration of Huh-7 hepatocellular carcinoma cells<sup>55</sup>. *FN1* encodes fibronectin, a soluble glycoprotein that binds to the cell surface and extracellular matrix. It is involved in cell-cell adhesion, cell motility, and maintenance of cell shape. Positive stromal fibronectin expression is significantly associated with low metastatic spread among patients with invasive breast carcinoma<sup>56</sup>. On the other hand, relatively low expression of stromal fibronectin correlates with lymph node metastasis, TNM stage, recurrence, and mortality<sup>57</sup>. Studies have also shown that tamoxifen-induced TGF- $\beta$  expression regulates fibronectin levels in a feedback loop. TGF- $\beta$  treatment reduced the levels of tamoxifen-induced *FN1* in MCF-7 cells<sup>58</sup>. Overall, the unbiased gene expression analysis revealed that *TGFBR1\*6A* alters the expression of known pro- and anti-metastatic effectors, which may contribute to its effect on breast cancer progression.

## **PERSPECTIVES AND FUTURE DIRECTIONS**

### **Genetic association with breast cancer risk and progression**

The case/control studies and meta-analyses assessing *TGFBR1\*6A* association with breast cancer risk suggest an association between *TGFBR1\*6A* and risk for breast cancer, albeit with demonstrable heterogeneity in odds ratios<sup>26,27,32,41</sup>. They also suggest that



**Fig. 3 Schematic showing the role of TGFBR1\*6A signal peptide.** **a** TGFBR1\*6A protein translation and processing, the TGFBR1\*6A signal peptide is cleaved between Ala30 and Leu31, whereas the wild-type TGFBR1 is cleaved between Ala33 and Leu34. Both TGFBR1\*6A and TGFBR1 wild-type exhibit similar binding affinity to TGF $\beta$  ligand and stability (half-life). The TGFBR1\*6A signal peptide also demonstrates similar protein targeting and translocation functions as the wild-type. **b** TGFBR1\*6A intracellular signaling, TGFBR1\*6A maintains intact TGF- $\beta$  signaling to induce growth and migration in breast cancer cells. It shows similar TGF- $\beta$  signaling as wild-type TGFBR1 but enhances phosphorylation of ERK1/2 to induce its tumor-promoting effects.

TGFBR1\*6A-associated breast cancer risk may depend on ancestry, geographical location, and other confounding factors. In genetic association studies, there are primarily four possible explanations that account for a statistical association between a genetic polymorphism (e.g., TGFBR1\*6A) and risk for a disease (e.g., breast cancer). They include (i) the allele directly affects the actual disease phenotype/expression, (ii) the allele is in linkage disequilibrium (correlated) with the true disease allelic locus, (iii) there is a spurious association due to population stratification or other confounders<sup>59,60</sup>, and (iv) type II error. To assert a direct allelic effect on disease phenotype/expression, the study design must control for confounding factors such as population stratification that could lead to false-positive/spurious associations. In general, ancestral population substructure is the main confounding concern contributing to spurious genetic associations<sup>61,62</sup>. Other sources include age, other mutations, and disease-specific contributory factors<sup>13,59,63–65</sup>. By design, population-based case/control studies such as those used in most of the TGFBR1\*6A case/control studies are susceptible to spurious risk associations due to their random or convenient selection of cases and controls from participants that may not belong to the

same ancestral population<sup>66–69</sup>. On the other hand, family-based case/control studies can help control for confounding population stratification because of shared family pedigrees,<sup>59,70</sup> albeit these designs can be more difficult to recruit. Thus, in a bid to conclusively establish the risk association between TGFBR1\*6A and breast cancer, future case/control studies should consider multicenter family-based population studies.

For multicenter population-based case/control studies that draw subjects from different self-reported races/ethnicities, careful consideration should be given to methods such as (i) structured association tests, (ii) principal components analyses (PCA), and (iii) multidimensional scaling (MDS) from GWAS ancestry to control genetic admixture and confounding effects<sup>63,71–73</sup>. Additionally, the effect sizes reported for nearly all complex genetic traits require large sample sizes for robust power and precise estimation. Thus, future studies should be designed with the range of magnitude of the effect, and allele frequencies observed in the published studies, and recruit cases and controls accordingly to enable well-designed and well-powered subgroup analyses of the contribution of the homozygous and heterozygous TGFBR1\*6A genotypes. This will allow assessment of TGFBR1\*6A

risk association by age, ethnicity, family history, tumor subtypes, tumor grade/histology, other mutation (*BRCA1/2*) status, response to cancer treatment, and survival/mortality outcomes. When compared to high (*BRCA1* and *BRCA2*: odds ratios ranging from 5.0 to 10.6) and moderate (*CHEK2* and *ATM*: odds ratios ranging from 2.1 to 2.5) penetrance gene variants<sup>4,74,75</sup>, *TGFBR1\*6A* has been described as low penetrance (<odds ratios <2.0) tumor susceptibility allele<sup>27,36</sup>. Although low penetrance gene variants and SNPs are common, their genetic risk to complex diseases such as cancer are comparatively small; carriers of low penetrance breast cancer gene variants are estimated to have less than 20% lifetime risk of developing breast cancers<sup>75</sup>. As such, low penetrance gene variants are postulated to exert their oncogenic influence through additive or multiplicative covariance interactions with other residual gene variants, whose presence or absence can determine disease trait<sup>76,77</sup>. Thus, the coupling of the *TGFBR1\*6A* family-based case/control studies with the participants' GWAS data will allow an additional estimation of the genome-wide composite association of *TGFBR1\*6A* with breast cancer risk, as it relates to other gene variants and gene loci that are in linkage disequilibrium, polygenic association or epistatic interaction. Altogether, the estimation of the genetic risk and polygenic risk score of *TGFBR1\*6A* from a robust family-based study will inform clinical guidelines for early detection and adoption of preventive measures.

#### ***TGFBR1\*6A* MECHANISM OF ACTION AND INFLUENCE ON TGF- $\beta$ BIOMARKER AND DRUG DEVELOPMENT**

Over the last decade, efforts aimed at developing anti-TGF- $\beta$  drugs has for the most part led to inconclusive pre-clinical and clinical results and serious adverse events. This is at least partly due to imprecise biomarkers and drug targets identified so far. For example, in a recent phase I study of LY3022859<sup>78</sup> (an anti-TGFBR2 kinase inactivating monoclonal antibody) in patients with advanced solid tumors including breast carcinomas, there were associated infusion-related reactions such as cytokine release syndrome. Also, the TGF- $\beta$  small molecule inhibitor galunisertib (LY2157299) and several others acting as neutralizing monoclonal antibodies<sup>79,80</sup> for TGFBR1 and TGFBR2, and protein traps<sup>81</sup> for TGFBR1 and TGFBR2<sup>82</sup> invariably yielded promising but minimal clinical benefits. In all, a key weakness is a difficulty to identify precise biomarkers that would tailor TGF- $\beta$  inhibitors to a specific patient population. Recent studies have proposed and developed a TGF- $\beta$  gene response signature (TBRS) to assess TGF- $\beta$  signaling response as a biomarker for cancer predisposition, clinical outcome, and therapeutic response. It utilizes up to 153 genes as probes, which include mainly the TGF- $\beta$  superfamily ligands (*TGFBR1*, *BMP2*), TGF- $\beta$  receptors (*TGFBR1* and *TGFBR2*), transcription factors (*BACH1*, *TXNIP*, and *CREB1*), and TGF- $\beta$  responsive genes (*ID1*, *HMOX1*, *MMP2*, and *ZEB1*)<sup>83–85</sup>. Padua, 2008<sup>83</sup> and Wahdan-Alaswad 2016<sup>85</sup>, used the TGF- $\beta$  gene response signature (TBRS) to classify breast tumors as TGF- $\beta$  gene responsive (TBRS+) and TGF- $\beta$  gene unresponsive (TBRS-) breast tumor subtypes and found a higher correlation between TBRS+, and ER-<sup>83</sup> and TNBC<sup>85</sup> breast tumor subtypes. Additionally, recent studies relate TGF- $\beta$  signaling with several cross-talk pathways including microRNA synthesis, stromal fibrosis, and endoplasmic reticulum (ER) stress that regulate immune checkpoint inhibitors, stem cell formation, and metastasis<sup>86–90</sup>.

The question of whether *TGFBR1\*6A* functions singularly or in combination with other pro-oncogenic pathways to signal modifications in oncogenic traits needs to be answered to provide clues for tailored drug and biomarker development for the affected *TGFBR1\*6A* individuals<sup>91–95</sup>. First, it is imperative to firmly establish whether *TGFBR1\*6A* signaling is a signal peptide or receptor-mediated in various breast cancer subtypes and cell lines. Subsequent interactions and networking with identified cross-talk

pathways and mediators can be traced using recent investigational tools such as Whole Transcriptome RNA Sequencing (RNA-seq) and proteomic profiling techniques coupled with computational network analyses such as comparative gene ontology (GO) enrichment analyses, ingenuity pathway analyses (IPA), and protein-protein interaction (PPI) network analyses<sup>96–98</sup>. In all, the development of a unique TGF- $\beta$  gene response signature (TBRS) for *TGFBR1\*6A* as well as the elucidation of a druggable target from the *TGFBR1\*6A* transcriptomic and proteomic studies will be valuable for the prognosis of at-risk individuals and clinical evaluation for precision medicine.

#### **CONCLUDING REMARKS**

Genetic association studies have identified *TGFBR1\*6A* as a high frequency, low penetrance breast cancer susceptibility allele in breast cancer patients of varied ethnic backgrounds and geographical locations. Its oncogenic influence is attributed to the promotion of cell proliferation, migration, and invasion. Further investigations into the risk associated with individual *TGFBR1\*6A* genotypes as well as their correlation with breast cancer subtypes, disease progression (tumor grade), metastasis, and survival are needed to clarify the population-specific susceptibilities of *TGFBR1\*6A* carriers.

#### **DATA AVAILABILITY**

Any additional dataset other than cited published data were available upon request to the corresponding author.

Received: 18 January 2022; Accepted: 13 June 2022;

Published online: 19 July 2022

#### **REFERENCES**

- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2019. *CA Cancer J. Clin.* **69**, 7–34 (2019).
- Siegel, R. L., Miller, K. D. & Jemal, A. Cancer statistics, 2020. *CA Cancer J. Clin.* **70**, 7–30 (2020).
- DeSantis, C. E., Miller, K. D., Goding, S. A., Jemal, A. & Siegel, R. L. Cancer statistics for African Americans, 2019. *CA Cancer J. Clin.* **69**, 211–233 (2019).
- Hu, C. et al. A population-based study of genes previously implicated in breast cancer. *N. Engl. J. Med.* **384**, 440–451 (2021).
- Metcalfe, K. et al. International trends in the uptake of cancer risk reduction strategies in women with a *BRCA1* or *BRCA2* mutation. *Br. J. Cancer* **121**, 15–21 (2019).
- Couch, F. J. et al. Associations between cancer predisposition testing panel genes and breast cancer. *JAMA Oncol.* **3**, 1190–1196 (2017).
- Zhang, H., Ahearn, T. U., Lecarpentier, J. & Barnes, D. Genome-wide association study identifies 32 novel breast cancer susceptibility loci from overall and subtype-specific analyses. *Nat. Genet.* **52**, 572–581 (2020).
- Ferreira, M. A. et al. Genome-wide association and transcriptome studies identify target genes and risk loci for breast cancer. *Nat. Commun.* **10**, 1741 (2019).
- Barnes, D. R. et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of *BRCA1* and *BRCA2* pathogenic variants. *Genet. Med.* **22**, 1653–1666 (2020).
- Ho, W. K. & Tan, M. M. European polygenic risk score for prediction of breast cancer shows similar performance in Asian women. *Nat. Commun.* **11**, 3833 (2020).
- Desmond, A. et al. Clinical actionability of multigene panel testing for hereditary breast and ovarian cancer risk assessment. *JAMA Oncol.* **1**, 943–951 (2015).
- Hauke, J. et al. Gene panel testing of 5589 *BRCA1/2*-negative index patients with breast cancer in a routine diagnostic setting: results of the German Consortium for Hereditary Breast and Ovarian. *Cancer* **7**, 1349–1358 (2018).
- Narod, S. A. Modifiers of risk of hereditary breast cancer. *Oncogene* **25**, 5832–5836 (2006).
- Buys, S. S. et al. A study of over 35,000 women with breast cancer tested with a 25-gene panel of hereditary cancer genes. *Cancer* **123**, 1721–1730 (2017).
- Massague, J. TGFbeta in cancer. *Cell* **134**, 215–230 (2008).
- Shi, M. et al. Latent TGF- $\beta$  structure and activation. *Nature* **474**, 343–349 (2011).
- Principe, D. R. et al. TGF- $\beta$ : duality of function between tumor prevention and carcinogenesis. *J. Natl Cancer Inst.* **106**, djt369 (2014).

18. Moore-Smith, L. & Pasche, B. TGFBR1 signaling and breast cancer. *J. Mammary Gland Biol. Neoplasia* **16**, 89–95 (2011).
19. Feng, X. H. & Derynck, R. Specificity and versatility in tgf-beta signaling through Smads. *Annu. Rev. Cell Dev. Biol.* **21**, 659–693 (2005).
20. Tzavlaki, K. & Moustakas, A. TGF- $\beta$  Signaling. *Biomolecules* **10** (2020).
21. Moustakas, A. & Heldin, C. H. Non-Smad TGF-beta signals. *J. Cell Sci.* **118**, 3573–3584 (2005).
22. Pasche, B. et al. Type I transforming growth factor beta receptor maps to 9q22 and exhibits a polymorphism and a rare variant within a polyalanine tract. *Cancer Res.* **58**, 2727–2732 (1998).
23. Pasche, B. et al. TbetaR-I(6A) is a candidate tumor susceptibility allele. *Cancer Res.* **59**, 5678–5682 (1999).
24. Ruddon, R. W. Progress in molecular biology and translational science. Preface. *Prog. Mol. Biol. Transl. Sci.* **95**, xi (2010).
25. Pasche, B. et al. Somatic acquisition and signaling of TGFBR1\*6A in cancer. *JAMA* **294**, 1634–1646 (2005).
26. Kaklamani, V. G. et al. Combined genetic assessment of transforming growth factor-beta signaling pathway variants may predict breast cancer risk. *Cancer Res.* **65**, 3454–3461 (2005).
27. Wang, Y. Q., Qi, X. W., Wang, F., Jiang, J. & Guo, Q. N. Association between TGFBR1 polymorphisms and cancer risk: a meta-analysis of 35 case-control studies. *PLoS ONE* **7**, e42899 (2012).
28. Rosman, D. S., Phukan, S., Huang, C. C. & Pasche, B. TGFBR1\*6A enhances the migration and invasion of MCF-7 breast cancer cells through RhoA activation. *Cancer Res.* **68**, 1319–1328 (2008).
29. Abecasis, G. R. et al. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
30. Abecasis, G. R. et al. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
31. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
32. Pasche, B. et al. TGFBR1\*6A and cancer: a meta-analysis of 12 case-control studies. *J. Clin. Oncol.* **22**, 756–758 (2004).
33. Chen, T. et al. Int7G24A variant of transforming growth factor-beta receptor type I is associated with invasive breast cancer. *Clin. Cancer Res.* **12**, 392–397 (2006).
34. Feigelson, H. S. et al. Transforming growth factor beta receptor type I and transforming growth factor beta1 polymorphisms are not associated with postmenopausal breast cancer. *Cancer Epidemiol. Biomark. Prev.* **15**, 1236–1237 (2006).
35. Cox, D. G., Penney, K., Guo, Q., Hankinson, S. E. & Hunter, D. J. TGFBR1 and TGFBR1 polymorphisms and breast cancer risk in the Nurses' Health Study. *BMC Cancer* **7**, 175 (2007).
36. Baxter, S. W., Choong, D. Y., Eccles, D. M. & Campbell, I. G. Transforming growth factor beta receptor 1 polyalanine polymorphism and exon 5 mutation analysis in breast and ovarian cancer. *Cancer Epidemiol. Biomark. Prev.* **11**, 211–214 (2002).
37. Song, B. et al. TGFBR1(\*)6A and Int7G24A variants of transforming growth factor-beta receptor 1 in Swedish familial and sporadic breast cancer. *Br. J. Cancer* **97**, 1175–1179 (2007).
38. Jin, Q. et al. Polymorphisms and haplotype structures in genes for transforming growth factor beta1 and its receptors in familial and unselected breast cancers. *Int. J. Cancer* **112**, 94–99 (2004).
39. Jakubowska, A. et al. BRCA1-associated breast and ovarian cancer risks in Poland: no association with commonly studied polymorphisms. *Breast Cancer Res. Treat.* **119**, 201–211 (2010).
40. Colleran, G. et al. The TGFBR1\*6A/9A polymorphism is not associated with differential risk of breast cancer. *Breast Cancer Res. Treat.* **119**, 437–442 (2010).
41. Joshi, N. N., Kale, M. D., Hake, S. S. & Kannan, S. Transforming growth factor  $\beta$  signaling pathway associated gene polymorphisms may explain lower breast cancer risk in western Indian women. *PLoS ONE* **6**, e21866 (2011).
42. Kamali, E., Hemmati, S., Safari, F. & Tavassoli, M. TGFBR1 polymorphism and risk of breast cancer in Iranian women. *Int. J. Biol. Markers* **30**, e414–e417 (2015).
43. Durbin, R. M. et al. A map of human genome variation from population-scale sequencing. *Nature* **467**, 1061–1073 (2010).
44. Kaklamani, V. G. et al. TGFBR1\*6A and cancer risk: a meta-analysis of seven case-control studies. *J. Clin. Oncol.* **21**, 3236–3243 (2003).
45. Liao, R. Y. et al. TGFBR1\*6A/9A polymorphism and cancer risk: a meta-analysis of 13,662 cases and 14,147 controls. *Mol. Biol. Rep.* **37**, 3227–3232 (2010).
46. Ou, C. et al. Meta-analysis of transforming growth factor  $\beta$  receptor 1 6A/9A gene polymorphism and breast cancer risk: the picture remains murky. *Biomarkers* **20**, 487–494 (2015).
47. Krishna, B.M., Jana, S. & Panda, A.K. Association of TGF- $\beta$ 1 polymorphisms with breast cancer risk: a meta-analysis of case-control studies †. *Cancers* **12**, 471 (2020).
48. Zhou, R., Huang, Y., Cheng, B., Wang, Y. & Xiong, B. TGFBR1\*6A is a potential modifier of migration and invasion in colorectal cancer cells. *Oncol. Lett.* **15**, 3971–3976 (2018).
49. Zhou, S., Zawal, L., Lengauer, C., Kinzler, K. W. & Vogelstein, B. Characterization of human FAST-1, a TGF beta and activin signal transducer. *Mol. Cell* **2**, 121–127 (1998).
50. Wrana, J. L. et al. TGF beta signals through a heteromeric protein kinase receptor complex. *Cell* **71**, 1003–1014 (1992).
51. Ponik, S. M., Trier, S. M., Wozniak, M. A., Eliceiri, K. W. & Keely, P. J. RhoA is down-regulated at cell-cell contacts via p190RhoGAP-B in response to tensional homeostasis. *Mol. Biol. Cell* **24**, 1688–1699 (2013). s1–3.
52. Stiegler, A. L. & Boggon, T. J. The N-terminal GTPase domain of p190RhoGAP proteins is a pseudoGTPase. *Structure* **26**, 1451–1461.e4 (2018).
53. Dong, G. et al. SIRT1 suppresses the migration and invasion of gastric cancer by regulating ARHGAP5 expression. *Cell Death Dis.* **9**, 977 (2018).
54. Tian, T. et al. Investigation of the role and mechanism of ARHGAP5-mediated colorectal cancer metastasis. *Theranostics* **10**, 5998–6010 (2020).
55. Gen, Y. et al. A novel amplification target, ARHGAP5, promotes cell spreading and migration by negatively regulating RhoA in Huh-7 hepatocellular carcinoma cells. *Cancer Lett.* **275**, 27–34 (2009).
56. Christensen, L., Nielsen, M., Andersen, J. & Clemmensen, I. Stromal fibronectin staining pattern and metastasizing ability of human breast carcinoma. *Cancer Res.* **48**, 6227–6233 (1988).
57. Takei, H. et al. Angiogenesis and stromal fibronectin expression in invasive breast carcinoma. *Int. J. Oncol.* **12**, 517–523 (1998).
58. Horii, Y. et al. The regulatory effect of tamoxifen on fibronectin expression in estrogen-dependent MCF-7 breast carcinoma cells. *Oncol. Rep.* **15**, 1191–1195 (2006).
59. Bhandari, V. Genetic Influences in Lung Development and Injury. In *The Newborn Lung: Neonatology Questions and Controversies* (ed. Bancalari, E.), Chapter 2, 29–55 (Saunders, Philadelphia, PA, 2012).
60. Tasha E., F. & David A., S. Genetics of Lung Disease. In *Murray and Nadel's Textbook of Respiratory Medicine, Vol. 1, part 1, Section A* (ed. Courtney, B. V.), Chapter 3, 35–37 (Elsevier, Philadelphia, PA, 2016).
61. Sirugo, G., Williams, S. M. & Tishkoff, S. A. The missing diversity in human genetic studies. *Cell* **177**, 1080 (2019).
62. Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).
63. Coignard, J., Lush, M., Beesley, J. & O'Mara, T.A. A case-only study to identify genetic modifiers of breast cancer risk for BRCA1/BRCA2 mutation carriers. *Nat. Commun.* **12**, 1078 (2021).
64. Sekhar, D., Pooja, S., Kumar, S. & Rajender, S. RAD51 135G>C substitution increases breast cancer risk in an ethnic-specific manner: a meta-analysis on 21,236 cases and 19,407 controls. *Sci. Rep.* **5**, 11588 (2015).
65. Chen, X. et al. Associations between RAD51D germline mutations and breast cancer risk and survival in BRCA1/2-negative breast cancers. *Ann. Oncol.* **29**, 2046–2051 (2018).
66. Kadouri, L. et al. A single-nucleotide polymorphism in the RAD51 gene modifies breast cancer risk in BRCA2 carriers, but not in BRCA1 carriers or noncarriers. *Br. J. Cancer* **90**, 2002–2005 (2004).
67. Antoniou, A. C. et al. RAD51 135G->C modifies breast cancer risk among BRCA2 mutation carriers: results from a combined analysis of 19 studies. *Am. J. Hum. Genet.* **81**, 1186–1200 (2007).
68. Wang, W. W. et al. A single nucleotide polymorphism in the 5' untranslated region of RAD51 and risk of cancer among BRCA1/2 mutation carriers. *Cancer Epidemiol. Biomark. Prev.* **10**, 955–960 (2001).
69. Levy-Lahad, E. et al. A single nucleotide polymorphism in the RAD51 gene modifies cancer risk in BRCA2 but not BRCA1 carriers. *Proc. Natl Acad. Sci. USA* **98**, 3232–3236 (2001).
70. Zondervan, K. T. & Cardon, L. R. Designing candidate gene and genome-wide case-control association studies. *Nat. Protoc.* **2**, 2492–2501 (2007).
71. Tian, C., Gregersen, P. K. & Seldin, M. F. Accounting for ancestry: population substructure and genome-wide association studies. *Hum. Mol. Genet.* **17**, R143–R150 (2008).
72. Price, A. L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat. Genet.* **38**, 904–909 (2006).
73. Helgason, A., Yngvadottir, B., Hrafnkelsson, B., Gulcher, J. & Stefansson, K. An Icelandic example of the impact of population structure on association studies. *Nat. Genet.* **37**, 90–95 (2005).
74. Narod, S. A. Which genes for hereditary breast cancer? *N. Engl. J. Med.* **384**, 471–473 (2021).
75. Dorling, L. et al. Breast cancer risk genes - association analysis in more than 113,000 women. *N. Engl. J. Med.* **384**, 428–439 (2021).
76. Mavaddat, N. et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am. J. Hum. Genet.* **104**, 21–34 (2019).
77. Phillips, P. C. Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nat. Rev. Genet.* **9**, 855–867 (2008).
78. Tolcher, A. W. et al. A phase 1 study of anti-TGF $\beta$  receptor type-II monoclonal antibody LY3022859 in patients with advanced solid tumors. *Cancer Chemother. Pharm.* **79**, 673–680 (2017).



79. de Gramont, A., Faivre, S. & Raymond, E. Novel TGF- $\beta$  inhibitors ready for prime time in onco-immunology. *Oncoimmunology* **6**, e1257453 (2017).
80. Morris, J. C. et al. Phase I study of GC1008 (fresolimumab): a human anti-transforming growth factor-beta (TGF $\beta$ ) monoclonal antibody in patients with advanced malignant melanoma or renal cell carcinoma. *PLoS ONE* **9**, e90353 (2014).
81. Zhu, H. et al. A novel TGF $\beta$  trap blocks chemotherapeutics-induced TGF $\beta$ 1 signaling and enhances their anticancer activity in gynecologic cancers. *Clin. Cancer Res.* **24**, 2780–2793 (2018).
82. Muraoka, R. S. et al. Blockade of TGF-beta inhibits mammary tumor cell viability, migration, and metastases. *J. Clin. Invest.* **109**, 1551–1559 (2002).
83. Padua, D. et al. TGFbeta primes breast tumors for lung metastasis seeding through angiopoietin-like 4. *Cell* **133**, 66–77 (2008).
84. Lehmann, B. D. et al. Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J. Clin. Invest.* **121**, 2750–2767 (2011).
85. Wahdan-Alaswad, R. et al. Metformin attenuates transforming growth factor beta (TGF- $\beta$ ) mediated oncogenesis in mesenchymal stem-like/claudin-low triple negative breast cancer. *Cell Cycle* **15**, 1046–1059 (2016).
86. Vanpouille-Box, C. et al. TGF $\beta$  is a master regulator of radiation therapy-induced antitumor immunity. *Cancer Res.* **75**, 2232–2242 (2015).
87. Zhang, D. et al. Proteomic characterization of differentially expressed proteins in breast cancer: expression of hnRNP H1, RKIP and GRP78 is strongly associated with HER-2/neu status. *Proteom. Clin. Appl.* **2**, 99–107 (2008).
88. Scriven, P. et al. Activation and clinical significance of the unfolded protein response in breast cancer. *Br. J. Cancer* **101**, 1692–1698 (2009).
89. Bai, W. D. et al. MiR-200c suppresses TGF- $\beta$  signaling and counteracts trastuzumab resistance and metastasis by targeting ZNF217 and ZEB1 in breast cancer. *Int. J. Cancer* **135**, 1356–1368 (2014).
90. Wang, S. et al. MicroRNA-133b targets TGF $\beta$  receptor I to inhibit TGF- $\beta$ -induced epithelial-to-mesenchymal transition and metastasis by suppressing the TGF- $\beta$ /SMAD pathway in breast cancer. *Int. J. Oncol.* **55**, 1097–1109 (2019).
91. Albain, K. S. et al. Prognostic and predictive value of the 21-gene recurrence score assay in postmenopausal women with node-positive, oestrogen-receptor-positive breast cancer on chemotherapy: a retrospective analysis of a randomised trial. *Lancet Oncol.* **11**, 55–65 (2010).
92. Dowsett, M. et al. Prediction of risk of distant recurrence using the 21-gene recurrence score in node-negative and node-positive postmenopausal patients with breast cancer treated with anastrozole or tamoxifen: a TransATAC study. *J. Clin. Oncol.* **28**, 1829–1834 (2010).
93. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
94. Milanese, J. S. et al. Germline variants associated with leukocyte genes predict tumor recurrence in breast cancer patients. *NPJ Precis. Oncol.* **3**, 28 (2019).
95. Xu, X. et al. Association of germline variants in natural killer cells with tumor immune microenvironment subtypes, tumor-infiltrating lymphocytes, immunotherapy response, clinical outcomes, and cancer risk. *JAMA Netw. Open* **2**, e199292 (2019).
96. Foroutan, M., Cursons, J., Hediyyeh-Zadeh, S., Thompson, E. W. & Davis, M. J. A transcriptional program for detecting TGF $\beta$ -induced EMT in cancer. *Mol. Cancer Res.* **15**, 619–631 (2017).
97. Wang, X., Liu, Q. & Zhang, B. Leveraging the complementary nature of RNA-Seq and shotgun proteomics data. *Proteomics* **14**, 2676–2687 (2014).

98. Li, Y. et al. RNA-Seq and network analysis revealed interacting pathways in TGF- $\beta$ -treated lung cancer cell lines. *Cancer Inf.* **13**, 129–140 (2014).

## ACKNOWLEDGEMENTS

The authors acknowledge Michael Olivier, Ph.D. (Center for Precision Medicine, Department of Internal Medicine, Wake Forest Baptist Health, Winston-Salem, USA) for his contribution to discussing parts of the manuscript. All images in Figs. 1 and 3 were created with Biorender (<https://biorender.com/>) online visual drawing tool.

## AUTHOR CONTRIBUTIONS

All authors contributed in parts to the conception, writing, and revision of the manuscript. The first author K.A. and corresponding author B.P. conceived and wrote the manuscript. The co-authors A.M.J., G.W.B., M.J.P., K.S., H.J., and C.B., contributed to discussing and revising the cancer genetics sections, S.S., P.A.F., R.S., K.L.C., H.-K.L., W.Z., H.-W.L., K.W., and P.S. contributed to discussing and revising the molecular signaling sections, and C.D.L. contributed to discussing and revising the cancer genetics and genome-wide association (GWAS) implications of the article.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Correspondence** and requests for materials should be addressed to Boris Pasche.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022