MOLECULAR AND CELLULAR BIOLOGY, Nov. 2011, p. 4609–4622 Vol. 31, No. 22
0270-7306/11/$12.00 doi:10.1128/MCB.05766-11

Gata-3 Negatively Regulates the Tumor-Initiating Capacity of Mammary Luminal Progenitor Cells and Targets the Putative Tumor Suppressor Caspase-14▽§

Marie-Liesse Asselin-Labat,[1,2] Kate D. Sutherland,[1]† François Vaillant,[1,2] David E. Gyorki,[1,2] Di Wu,[1,2] Sheridan Holroyd,[1] Kelsey Breslin,[1] Teresa Ward,[1] Wei Shi,[1,2] Mary L. Bath,[1] Siddhartha Deb,[3] Stephen B. Fox,[3] Gordon K. Smyth,[1,2] Geoffrey J. Lindeman,[1,2,4] and Jane E. Visvader[1,2]*

The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria, Australia[1]; The University of Melbourne, Parkville, Victoria, Australia[2]; Peter MacCallum Cancer Centre, East Melbourne, Victoria, Australia[3]; and The Royal Melbourne Hospital, Parkville, Victoria, Australia[4]

Received 9 June 2011/Returned for modification 15 July 2011/Accepted 10 September 2011

The transcription factor Gata-3 is a definitive marker of luminal breast cancers and a key regulator of mammary morphogenesis. Here we have explored a role for Gata-3 in tumor initiation and the underlying cellular mechanisms using a mouse model of "luminal-like" cancer. Loss of a single *Gata-3* allele markedly accelerated tumor progression in mice carrying the mouse mammary tumor virus promoter-driven polyoma-virus middle T antigen (MMTV-PyMT mice), while overexpression of *Gata-3* curtailed tumorigenesis. Through the identification of two distinct luminal progenitor cells in the mammary gland, we demonstrate that *Gata-3* haplo-insufficiency increases the tumor-initiating capacity of these progenitors but not the stem cell-enriched population. Overexpression of a conditional *Gata-3* transgene in the PyMT model promoted cellular differentiation and led to reduced tumor-initiating capacity as well as diminished angiogenesis. Transcript profiling studies identified *caspase-14* as a novel downstream target of Gata-3, in keeping with its roles in differentiation and tumorigenesis. A strong association was evident between GATA-3 and caspase-14 expression in preinvasive ductal carcinoma *in situ* samples, where GATA-3 also displayed prognostic significance. Overall, these studies identify GATA-3 as an important regulator of tumor initiation through its ability to promote the differentiation of committed luminal progenitor cells.

Breast cancer is a highly heterogeneous disease and can be broadly classified into three major types based on histopathology: estrogen and/or progesterone receptor-positive luminal tumors that account for the majority of breast cancer cases, HER2-amplified/overexpressing tumors, and basal-like tumors that are frequently negative for the estrogen receptor (ER), progesterone receptor (PR), and HER2 while positive for cytokeratin 5/6 and epidermal growth factor receptor (EGFR). Molecular profiling analyses have led to the further stratification of breast tumors into six major subtypes that include the luminal A, luminal B, claudin-low, normal-like, basal-like, and HER2-positive subclasses (12, 30, 36–38). Although the subtypes can be prognostic and predictive for patient outcome, responsiveness to chemotherapy or targeted therapy is highly variable. A more comprehensive understanding of the distinct subtypes and intrinsic differences between individual breast tumors of the same type is required to enable the design of more effective therapies.

The zinc finger transcription factor GATA-3 is a defining marker of luminal breast cancers. The highest levels of GATA-3 occur in the luminal subtypes (4, 13, 26, 37, 38, 45, 46), which account for over 70% of breast cancers. GATA-3 is often coexpressed with estrogen receptor α (ERα), a direct target of Gata-3, and vice versa, implying that these factors operate in a positive cross-regulatory loop (9). Although GATA-3 is an important luminal cancer marker, its role as an independent prognostic factor in breast cancer is unclear. Interestingly, frequent somatic mutations in the *GATA-3* gene occur in *BRCAX* (non-BRCA1/BRCA2) familial breast tumors but not in *BRCA1* or *BRCA2* tumors (2). It has been speculated that *GATA-3* mutations may occur earlier in the evolution of *BRCAX* tumors relative to *BRCA1*, *BRCA2*, or sporadic tumors (2). It is noteworthy that mutations in *GATA-3* are found in at least 4 to 5% of sporadic breast tumors (41).

Gata-3 is differentially expressed within the epithelial hierarchy of the normal mouse mammary gland. There are two primary mammary epithelial lineages comprising luminal (both ductal and alveolar) and myoepithelial cells. The lowest levels of Gata-3 occur in the mammary stem cell (MaSC)-enriched population and the highest levels in mature luminal cells, while mature myoepithelial cells are devoid of Gata-3 (3). During normal ontogeny, Gata-3 acts as a critical regulator of mammary gland morphogenesis and luminal differentiation (3, 20). In young developing mice, *Gata-3* loss results in the formation of stunted ductal trees and an augmented luminal progenitor subset, even in heterozygous mice. During pregnancy, severely impaired alveolar development was evident in *Gata-3*-deficient

* Corresponding author. Mailing address: The Walter and Eliza Hall Institute of Medical Research, 1G Royal Parade, Parkville, VIC 3052, Australia. Phone: 61-3-9345-2494. Fax: 61-3-9347-0852. E-mail: visvader@wehi.edu.au.

† Present address: The Netherlands Cancer Institute, Amsterdam, The Netherlands.

§ Supplemental material for this article may be found at http://mcb.asm.org/.

▽ Published ahead of print on 19 September 2011.

TABLE 1. Primer sequences used in this study

Primer use	Primer sequences[a]
Genotyping	
MTB	Fwd, 5′-TGCCGCCATTATTACGACAAGC-3′; Rev, 5′-ACCGTACTCGTCAATTCCAAGGG-3′
Gata-3$^{+/Tg}$	Fwd, 5′-ACGTCTCACTCTCGAGGCAGCATG-3′; Rev, 5′-GAAGTCCTCCAGCGCGTCATGCAC-3′
Gata-3$^{+/nlslacZ}$	Fwd, 5′-CAGGAGTCCGCGGACCTCC-3′; Rev, 5′-CGTTGAGGACCGCGGGGTG-3′; Rev, 5′-CCA AGCTTGGACTCAAAAAAC-3′
MMTV-PyMT	Fwd, 5′-CTCTAGAGGATCTTTGTGAAGG-3′; Rev, 5′-GGACAAACCACAACTAGAATGC-3′
Gene expression analysis	
β-Casein	Fwd, 5′-AAAGGACTTGACAGCCATGAA-3′; Rev, 5′-TAGCCTGGAGCACATCCTCT-3′
Gata-3	Fwd, 5′-AGCCACATCTCTCCCTTCAG-3′; Rev, 5′-AGGGCTCTGCCTCTCTAACC-3′
Caspase-14	Fwd, 5′-GATGAGGTTGCTGTGCTCAA-3′; Rev, 5′-GAACACATCCGTCAGGGTCT-3′
GAPDH	Fwd, 5′-GTATTGGGCGCCTGGTCACC-3′; Rev, 5′-CGCTCCTGGAAGATGGTGATGG-3′
18S rRNA	Fwd, 5′-GTAACCCGTTGAACCCCATT-3′; Rev, 5′-CCATCCAATCGGTAGTAGCG-3′
Chromatin immunoprecipitation	
Control	Fwd, 5′-AGGCATCTCTTTGAGATGGCT-3′; Rev, 5′-TGGGTCATTCACCAGCTTCT-3′
−484	Fwd, 5′-CTGTCTGTGCAGCCAGAGAA-3′; Rev, 5′-AGCAGGGCAAGTGACCCATGTG-3′
−3060	Fwd, 5′-CCTTATAGGGTTGAATGCCAA-3′; Rev, 5′-TGGCCTAGGCAGAGAGTCT-3′
−3772	Fwd, 5′-GACTGGCAGGTAGGGAAGGTGAG-3′; Rev, 5′-ATAGTGAAACCAAAGGCGAAC-3′
−4392	Fwd, 5′-ATCTCACACTTTGGGAGGCCGAGTG-3′; Rev, 5′-ACTCGGTTCGATAGCAACAGCT-3′
Plasmid construction	
Gata-3 Tg (pTet-On-Gata-3)	Fwd, 5′-GGAAGATCTCACCATGGACTACAAAGACGATGACGATAAAGAGGTGACTGCGGAC CAGCCGCGC-3′; Rev, 5′-GGAAGATCTCTAACCCATGGCGGTGACCAT-3′
pFU-TA-GFP Gata-3	Fwd, 5′-TGCGAATTCATGGAGGTGACGGCGGACCAGCCGCGC-3′; Rev, 5′-CCGAGATCTACCC ATGGCGGTGACCATGCTGGA-3′
pFU-TA-GFP caspase-14	Fwd, 5′-CGCGAATTCTCACCATGAGCAATCCGCGGTCTT-3′; Rev, 5′-CCGGGATCCCTGCAGA TACAGCCGTTTCCGGA-3

[a] Fwd, forward; Rev, reverse.

mammary glands, accompanied by expansion of the committed luminal progenitor pool, thus indicating a block in differentiation. Overexpression of *Gata-3* alone in the MaSC-enriched subset induced the expression of milk protein genes in the absence of a lactogenic stimulus, suggesting that Gata-3 is a master regulator of luminal differentiation (3).

In the context of breast cancer, GATA-3 has been implicated in tumor differentiation and metastasis. Significantly, overexpression of GATA-3 suppressed tumor dissemination of mouse mammary tumor cells (19) or xenografted human breast cancer cells to the lung (8). The role of GATA-3 in tumor initiation, however, remains unclear. We therefore examined its function in the early stages of mammary tumorigenesis using mice either lacking a single *Gata-3* allele or those conditionally overexpressing this transcription factor. Through the analysis of distinct mammary epithelial subsets, including a novel alveolar progenitor cell, we show that Gata-3 specifically affects the tumor-initiating ability of the luminal progenitor but not the MaSC-enriched cellular subset. Expression profiling studies also revealed a novel target gene of Gata-3, *caspase-14*, which was found to be frequently overexpressed with GATA-3 in preinvasive ductal carcinoma *in situ* (DCIS) and may represent a new breast tumor suppressor.

MATERIALS AND METHODS

Plasmids. The sequences of all primers used in this study are listed in Table 1. For the pTet-On-Gata-3 construct, mGata-3 was amplified by PCR. Flag-mGata-3 cDNA was cloned into the BglII site of the pTet-On-IRES-GFP vector backbone (a gift from C. Ormandy, Garvan Institute, Australia) and verified by sequencing. The transgene was excised using PvuI and purified for oocyte injection. For the lentiviral constructs, full-length *Gata-3* was amplified by PCR using FastStart high-fidelity polymerase (Roche), digested with EcoRI and BglII, and then cloned into the pFU-TA-GFP vector. Full-length caspase-14 was amplified by PCR (using FastStart high-fidelity polymerase [Roche] from pCMV-XL4-Caspase-14 (a kind gift from H. Hsu [14]), digested with EcoRI and BamHI, and cloned into pFU-TA-GFP. The pBABE-Gata-3 plasmid has been described previously (3).

Mice and tumor monitoring. Transgenic mice carrying the Tet-On-Gata-3 IRES-GFP cassette were generated by oocyte injection on an FVB/NJ background. MTB mice, mice with the mouse mammary tumor virus promoter-driven polyomavirus middle T antigen (MMTV-PyMT [FVB/NJ] mice), and floxed Gata-3 and Gata-3$^{+/nlsLacZ}$ mice have been described previously (3, 11, 22, 43). The MTB; Gata-3$^{+/Tg}$ mice were generated by cross-breeding and given 1 mg/ml doxycycline (Sigma) in their drinking water (with 2% sucrose) at 4 weeks of age. Mouse tail DNA was genotyped by PCR using primers listed in Table 1.

Mice were palpated three times a week for tumor growth. The date that tumors were first detected was recorded for the generation of tumor-free survival curves. Tumors were collected at the ethical endpoint size of 1 cm in diameter. All animal experiments were conducted with approval of the WEHI Animal Ethics Committee and in accordance with its guidelines.

Whole-mount analysis, histology, and immunostaining of mammary tissue. For histological examination, mammary glands were fixed in 4% paraformaldehyde in phosphate-buffered saline (PBS), embedded in paraffin, sectioned (1.5 μm), and stained with hematoxylin and eosin. For whole-mount analysis, mammary glands were harvested from pregnant or lactating mice, fixed in Carnoy's solution, and then stained with hematoxylin. For ERα and milk protein immunohistochemistry, sections were blocked in 10% normal goat serum prior to incubation with specific antibodies followed by a biotin-conjugated anti-rabbit IgG. A streptavidin-biotin peroxidase detection system was used with 3,3′-diaminobenzidine as a substrate (Dako). Cytokeratin 18 and SMA immunostaining was performed using the Mouse-on-Mouse kit (Vector) according to the manufacturer's protocol. For NPt2b immunofluorescence staining, sections were blocked in 10% normal goat serum and then incubated with anti-NPt2b antibody followed by Alexa 594-conjugated anti-rabbit IgG.

Mammary cell preparation, cell sorting, and *in vivo* transplantation. The thoracic and inguinal mammary glands were dissected from 8-week-old virgin female mice, and cell suspensions were prepared as previously described (32). Fluorescence-activated cell sorter (FACS) analysis and cell sorting were performed using a FACS Diva or FACS Aria (Becton Dickinson). Antibodies

against mouse antigens were purchased from BD Pharmingen unless otherwise specified and included CD24-phycoerythrin (PE), biotinylated CD31, TER119 and CD45, CD29-fluorescein isothiocyanate (FITC) (clone HMβ1.1), CD61-Alexa 647 (Biolegend), CD14-biotin (eBiosciences), c-kit-PE (clone ACK4) (gift from A. Strasser), streptavidin-APC-Cy7, and streptavidin-Alexa 594 (Molecular Probes).

For normal and tumor cell transplantation assays, freshly sorted cellular subsets (purity, >90%) were injected into the cleared inguinal glands of 3-week-old syngeneic recipient mice (32). Limiting dilutions were calculated using the ELDA software (15). ELDA performs maximum likelihood analysis of limiting dilution data sets. It is able to return statistically reliable results even for small numbers of assays. ELDA uses likelihood ratio statistics to test for enrichment of active cells in one cell population versus another.

Colony-forming and differentiation assays. For colony formation assays, freshly sorted mammary epithelial cells were plated on irradiated fibroblast feeders for 5 days or embedded in 20 μl of Matrigel (BD Pharmingen) (1,000 cells) as described previously (32). Colonies in 2D assays were scored following Giemsa staining (Merck).

For the HC11 differentiation assay, cells (10^5) were grown for 2 days in RPMI supplemented with 10% fetal calf serum (FCS), 5 μg/ml insulin, and 5 ng/ml EGF. The medium was then replaced with EGF-deficient medium. Differentiation medium (RPMI supplemented with 10% FCS, 5 μg/ml insulin, 5 μg/ml prolactin, and 500 μg/ml hydrocortisone) was added to the cells 24 h later. Cells were harvested 2 and 3 days after differentiation, and RNA was purified.

Retroviral infections. Retroviral infection of Gata-$3^{f/f}$ cells has been described previously (3). For lentiviral transductions, 293T cells (5×10^5) plated in 6-cm dishes were transfected with 5 μg pMDlg.RRE, 2.5 μg pRSV-Rev, 3 μg pMD-G, and 10 μg of the gene of interest using Fugene (Roche). Viral supernatant was collected 2 days after transfection and filtered, and Polybrene was added to a final concentration of 1 μg/ml prior to infection. MDA-MB-231^{Luci} cells (grown in RPMI supplemented with 10% FCS and 5 μg/ml insulin) or HC11 cells (grown in RPMI supplemented with 10% FCS and 5 μg/ml insulin and 5 ng/ml EGF) were spin infected at 2,500 rpm for 1 h. Expression was analyzed by Western blotting 3 days after infection and reassessed on the day of transplantation.

***In vivo* imaging.** MDA-MB-231^{Luci} cells (Caliper Life Sciences) infected with the relevant lentiviruses were injected into the cleared inguinal glands of 3-week-old NOD-SCID-IL-2Rγ$^{-/-}$ recipients. For live imaging, mice were injected intraperitoneally with Luciferin (Caliper Life Sciences, 15 mg/ml) 10 min prior to imaging. Mice were imaged on the day following mammary fat pad transplantation and once a week up to 8 weeks postsurgery using the Xenogen IVIS live imager (Caliper Life Sciences).

Antibodies for Western blot analysis and immunostaining. Anti-Gata-3 antibodies used for Western blotting and immunostaining were from Santa Cruz (sc-268) and Abcam (ab32858), respectively. The following antibodies were used for Western blot analysis: antitubulin (Sigma), anti-ERα (Santa Cruz sc-542), anti-cytokeratin 18 (Progen Biotechnik), anti-SMA (Sigma), anti-caspase-14 (Abcam, ab45415), anti-GAPDH (Sigma), and anti-β-actin (Sigma). The anti-NPt2b antibody was kindly provided by J. Biber (University of Zurich, Switzerland). Alexa 594-conjugated anti-rabbit (Biolegend) and biotin-conjugated anti-rabbit (Vector) secondary antibodies were used for immunofluorescence and immunohistochemistry, respectively.

ChIP analysis. MCF-7 cells (human breast cancer cell line derived from a metastatic breast pleural effusion) were cross-linked with 2% formaldehyde for 10 min at room temperature, and the chromatin immunoprecipitation (ChIP) assay was performed according to the manufacturer's protocol (Upstate 17-371) using anti-Gata-3 antibody (sc-268; Santa Cruz) and normal mouse IgG (Upstate 17-371) as a control. Immunoprecipitated DNA was analyzed by real-time PCR using primers corresponding to the specific region (spanning the four putative GATA-3 binding sites and flanking region lacking GATA-3 binding sites). PCR primers are listed in Table 1.

RNA preparation and quantitative RT-PCR analysis. Reverse transcription-PCR (RT-PCR) for Gata-3 has been described previously (3). Oligonucleotides used for quantitative RT-PCR are listed in Table 1.

Genomic and transcriptome profiling analyses. The DNA sequence of the promoter region of caspase-14 was retrieved from the GABOS database (http://bioinf.wehi.edu.au/gabos). Binding sites for Gata-3 were found using the MATCH program and position weight matrices from the TRANSFAC database.

RNA isolated from GFP$^+$ and GFP$^-$ cells from two PyMT MTB Gata-$3^{+/Tg}$ tumors was profiled on Illumina MouseWG-6 Version 2 BeadChips at the Australian Genome Research Facility (Melbourne), using standard Illumina protocols. Unnormalized data were exported from BeadStudio and analyzed using the limma package of the Bioconductor software project (10). Raw intensities were normexp background corrected with offset 16, quantile normalized, and then $\log_2$ transformed (33). Probes were filtered if not detected in any sample (detection *P* value, 0.01, based on negative-control probes). The correlation between paired samples from the same tumor was estimated as 0.19 (35). An empirical quality weight was estimated for each array (31), and the weights and the within-pool correlation were incorporated into a linear model analysis. Differential expression between GFP$^+$ and GFP$^-$ cells was assessed using empirical Bayes moderated t-statistics (34). The false discovery rate (FDR) was controlled using the Benjamini and Hochberg algorithm. Probes with FDR of <0.2 and fold change of >1.5 were judged to be differentially expressed. Manufacturer probe annotation was used.

Patient samples and scoring. Patient details, including age, tumor size, grade, histology, nodal status, estrogen receptor (ER) status, and HER-2 status, were collected from clinical and pathological records. The median age of patients with DCIS was 56 years (range, 32 to 75 years). The clinicopathological characteristics of the DCIS have been described previously (40). Tissue microarrays (2-mm cores) were created from 64 cases of pure DCIS collected from patients who underwent surgery at the John Radcliffe Hospital (Oxford, United Kingdom). This study has Ethical Committee approval (number C02.216). Using conventional pathological criteria, the arrays were scored by two independent pathologists for intensity: 0, no staining; 1, weak staining; 2, moderate staining; and 3, strong staining. The scoring system for percentages was as follows: 0, no cells staining positive for caspase-14 or Gata-3; 1, ≤10% cells staining positive; 2, 11% to 50% positive cells; 3, 51% to 80% positive cells; and 4, >80% positive cells. A final score (range from 0 to 12) was obtained from the product of the staining intensity and the percentage of cells in both nuclear and cytoplasmic compartments. To stratify patients for analysis, a score of 3 or more was considered positive for cytoplasmic or nuclear Gata-3 or caspase-14 expression in the statistical analysis.

RESULTS

Gata-3 is highly expressed in two distinct luminal progenitor subtypes and appears to be required for stem cell function. As a prerequisite to understanding the cellular mechanisms by which Gata-3 affects breast tumorigenesis, we further dissected epithelial subtypes along the luminal lineage. CD24 and CD29 have previously been used to resolve luminal and MaSC-enriched/basal cell populations (Fig. 1A) (32). CD61 was subsequently used to identify committed luminal progenitor cells in the mouse mammary gland (3), while c-kit has recently emerged as a defining marker of the gene expression signatures of mouse and human luminal progenitor cells (23, 24). CD61 and c-kit were found to be coexpressed by the majority of, but not all, committed luminal progenitor cells (Fig. 1A).

We next used a combination of antibodies against c-kit and CD14 to identify a novel luminal progenitor subpopulation. CD14 was investigated, as it appeared to be a potential marker of specific luminal cells (17, 24, 39). This marker was found to subdivide c-kit-negative/low cells into two distinct subpopulations (Fig. 1A): CD29^{lo} CD24$^+$ CD14$^+$ c-kit$^{-/lo}$ cells (termed CD14$^+$ c-kit$^{-/lo}$) exhibited colony-forming capacity when plated on fibroblast feeder layers or in Matrigel (Fig. 1B), whereas the double-negative CD29^{lo} CD24$^+$ CD14$^-$ c-kit$^-$ subset (designated CD14$^-$ c-kit$^-$) lacked clonogenic activity. The first subset accounts for the progenitor activity previously detected in the CD61$^-$ population (3, 5). The CD29^{lo} CD24$^+$ c-kit$^+$ subset (termed c-kit$^+$) exhibited substantial progenitor activity *in vitro* (Fig. 1B) and is likely to share overlap with the previously described CD29^{lo} CD24$^+$ CD61$^+$ subset. Interestingly, 45% of the colonies arising from CD14$^+$ c-kit$^{-/lo}$ cells produced milk protein in the absence of a lactogenic stimulus, compared to 23% of c-kit$^+$ cells, suggesting that CD14$^+$ c-kit$^{-/lo}$ cells are precommitted to the alveolar lineage (Fig. 1C). Furthermore, the size of the CD14$^+$ c-kit$^{-/lo}$ subset reached a maximum in late pregnancy (18.5 days of pregnancy [dP])

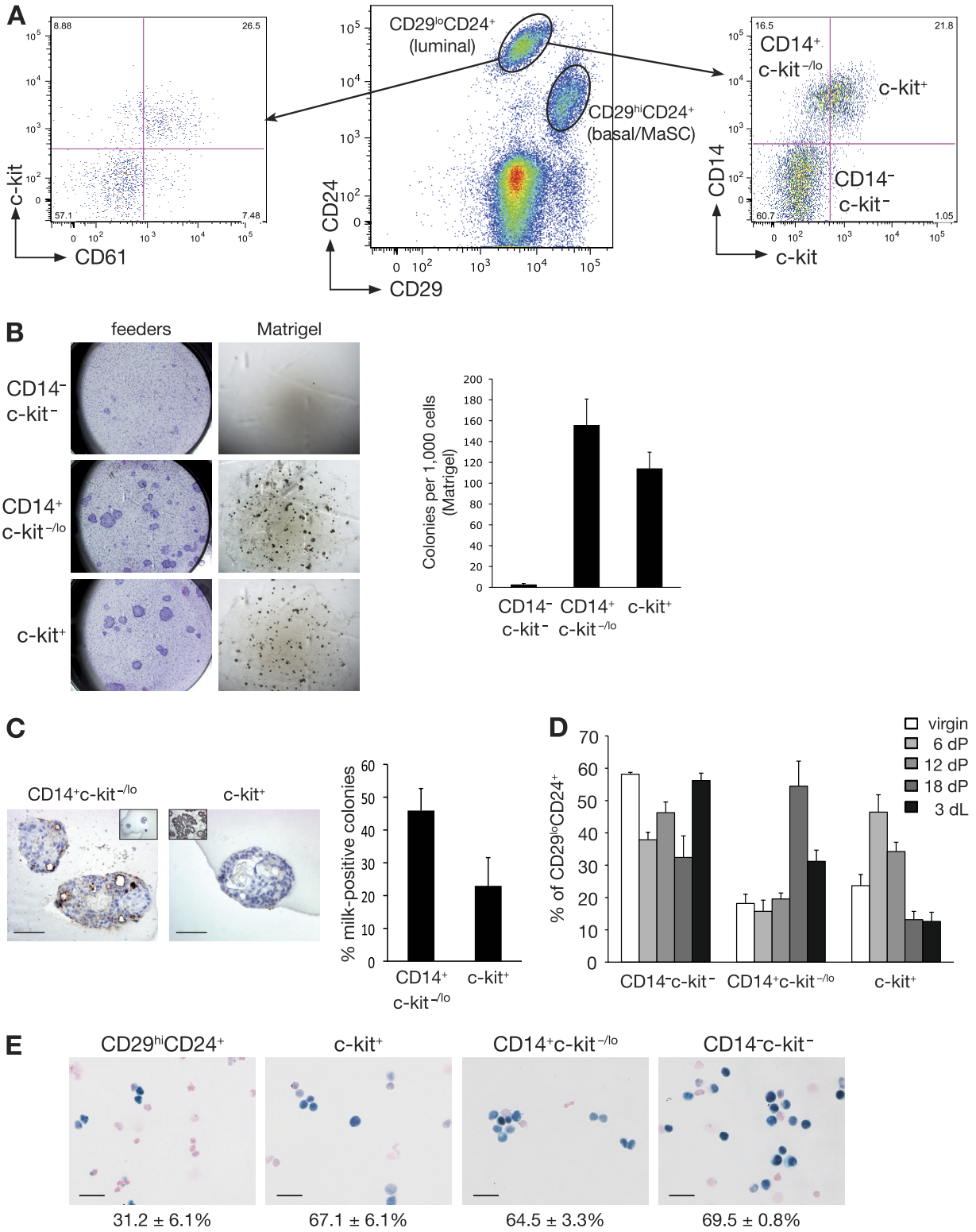

FIG. 1. Isolation of a distinct alveolar progenitor cell using the CD14 and c-kit markers. (A) Representative FACS dot plot showing the expression of CD29 and CD24 in Lin$^-$ (CD31$^-$ CD45$^-$ TER119$^-$) cells isolated from 8-week-old virgin mouse mammary glands (middle panel). Representative FACS dot plots showing CD61 and c-kit expression (left panel) or CD14 and c-kit expression (right panel) in Lin$^-$ CD29^{lo} CD24$^+$ cells in mammary glands taken from an 8-week-old virgin mouse. (B) Colony-forming capacity of c-kit/CD14 subpopulations plated on fibroblast feeder layers or embedded in Matrigel, as indicated. Right panel, histogram showing the colony formation capacity of c-kit/CD14 cells plated in Matrigel. Data represent the means ± standard error of the means (SEM) of results of three independent experiments. (C) Representative anti-milk protein immunostaining of CD14$^+$ c-kit$^{-/lo}$ and c-kit$^+$ colonies grown in Matrigel for 14 days. Left inset shows staining for the isotype control antibody; right inset shows milk staining of a section from an 18.5-day-pregnant-mouse mammary gland. Scale bars, 60 μm. Right panel,

TABLE 2. Limiting dilution analysis of the repopulating frequency of CD29^{hi} CD24^{+} MaSC-enriched cells from control or MMTV-cre; Gata-3$^{f/f}$ mammary glands[a]

Recipient	No. of CD29^{hi} CD24^{+} cells injected per mammary fat pad	No. of positive outgrowths[b]	
		Gata-3$^{+/f}$	MMTV-cre; Gata-3$^{f/f}$
Rag1$^{-/-}$	75	4/9	0/9
	150	1/9	1/9
	250	8/11	2/10
	500	7/8	3/8
Repopulating frequency (95% confidence interval)		1/254 (1/406–1/159)	1/1,232 (1/2,737–1/554)
P value			0.0003
NOD-SCID	250	2/2	1/2
	500	3/4	0/4
Repopulating frequency (95% confidence interval)		1/244 (1/670–1/87)	1/2,373 (1/17,594–1/320)
P value			0.016

[a] Limiting dilution analysis of the repopulating frequency of CD29^{hi} CD24^{+} cells (double sorted) from the mammary glands of Gata-3$^{+/f}$ or MMTV-cre;Gata-3$^{f/f}$ mice. Cells were injected into the cleared mammary fat pads of 3-week-old Rag1$^{-/-}$ or NOD-SCID immunocompromised recipients and fat pads collected 8 weeks posttransplantation. Data are pooled from three (Rag1$^{-/-}$) or two (NOD-SCID) independent experiments for each cell type.

[b] Shown as number of outgrowths per number of injected fat pads.

before decreasing. Conversely, the c-kit^{+} subpopulation expanded in early pregnancy (6 dP; Fig. 1D) but then declined by midpregnancy. This progenitor subset appears to lie upstream of the alveolar progenitor cell and may have the capacity to differentiate into either ductal or alveolar progenitor cells, dependent on the hormonal stimulus.

Gata-3 levels were found to be comparable between the two luminal progenitor subtypes (64.5% ± 3.3% for c-kit^{+} cells, 67.1% ± 6.15% for CD14^{+} c-kit$^{-/lo}$ cells) and CD14^{-} c-kit^{-} mature luminal cells (69.5% ± 0.8%). In contrast, only 31.2% (±6.1%) of the cells within the MaSC-enriched subset expressed Gata-3 as demonstrated by β-galactosidase staining of freshly sorted and cytospun mammary epithelial subsets from Gata-3$^{+/nlslacZ}$ glands, as previously noted (3) (Fig. 1E).

Given that Gata-3 expression was evident in the MaSC-enriched (CD29^{hi} CD24^{+}) population but not in mature myoepithelial cells (3) that also reside in this subset, we hypothesized that this transcription factor might be specifically expressed in MaSCs and regulate their function. Transplantation assays revealed a striking 10-fold decrease in the number or activity of mammary repopulating cells in *Gata-3*-deficient mammary glands (MMTV-cre; Gata-3$^{f/f}$ mice) compared to control Gata-3$^{+/f}$ glands (Table 2 and Fig. 2). In addition, a marked decrease in the capacity of these cells to yield full outgrowths that permeated the fat pad was noted (Fig. 2B). *Ex vivo* deletion of *Gata-3* in CD29^{hi} CD24^{+} cells from *Gata-3$^{f/f}$* mice by pMIG-cre-mediated transduction also led to reduced mammary repopulating capacity (data not shown). The decreased repopulating ability of the *Gata-3*-deficient MaSC-enriched subset may reflect a change in the absolute number of stem or bipotent progenitor cells due to diminished renewal or, alternatively, impaired lineage restriction potential of these cells or their immediate descendants.

Gata-3 haplo-insufficiency accelerates mammary tumor onset. The MMTV-PyMT transgenic model was selected to study the influence of Gata-3 on mammary tumor initiation, since analysis of microarray data sets of different models of tumorigenesis revealed high *Gata-3* expression levels in "luminal-like" MMTV-PyMT tumors (12, 19). In this model, mammary hyperplasias progress to adenocarcinomas, with a short preneoplastic period preceding the development of multiple distinct foci throughout the gland (11, 25).

Loss of a single *Gata-3* allele from MMTV-PyMT mice (MMTV-PyMT; Gata-3$^{+/nlslacZ}$) reduced the latency of tumor onset compared to that of mice that retained both alleles (72 versus 48 days) (Fig. 3A). The lower levels of Gata-3 in MMTV-PyMT; Gata-3$^{+/nlslacZ}$ mammary tissue were verified by Western blot analysis (Fig. 3B). Interestingly, the tumors arising in MMTV-PyMT; Gata-3$^{+/nlslacZ}$ mice were phenotypically different from MMTV-PyMT-derived tumors, with cystic dilatation evident in almost every MMTV-PyMT; Gata-3$^{+/nlslacZ}$ tumor (Fig. 3C). NPt2b, an early marker of alveolar differentiation, was readily detected in MMTV-PyMT tumors, whereas very low levels were apparent in MMTV-PyMT; Gata-3$^{+/nlslacZ}$ tumors, suggesting that they may be more ductal-like lesions (Fig. 3C). The change in histopathology associated with *Gata-3* haplo-insufficiency could reflect a change in the "cell of origin" of transformation or, alternatively, the extent of differentiation. To examine these possibilities further, we assessed the tumor-initiating capacity of the different epithelial subsets by *in vivo* transplantation (see below).

bar chart showing the percentage of milk protein-positive colonies observed 14 days after plating of CD14^{+} c-kit$^{-/lo}$ and c-kit^{+} cells in Matrigel (means ± SEM of results of three independent experiments). (D) Histogram showing the percentage of c-kit/CD14 cells in the Lin^{-} CD29^{lo} CD24^{+} population at different stages of development: virgin (V), pregnancy (dP), and lactation (dL). Results represent the means ± SEM of results for three to five animals per group. (E) β-Galactosidase activity in freshly cytospun CD29^{hi} CD24^{+} (MaSC-enriched), c-kit^{+}, CD14^{+} c-kit$^{-/lo}$, and CD14^{-} c-kit^{-} cells from Gata-3$^{+/nlslacZ}$ mice. Scale bars, 25 μm.

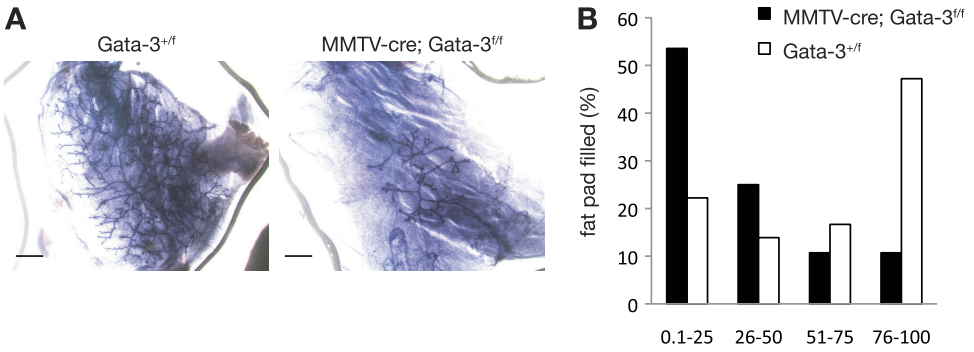

FIG. 2. Gata-3 is important for mammary stem cell function. (A) Representative whole-mounted outgrowths obtained 8 weeks posttransplantation of CD29^{hi} CD24^{+} cells from Gata-3$^{+/f}$ or MMTV-cre; Gata-3$^{f/f}$ glands. Scale bars, 1 mm. (B) Bar chart showing percentage of fat pad filling by outgrowths 8 weeks posttransplantation of CD29^{hi} CD24^{+} cells from Gata-3$^{+/f}$ or MMTV-cre; Gata-3$^{f/f}$ mammary glands.

Augmented tumor-initiating activity of luminal progenitor subsets in Gata-3-deficient MMTV-PyMT mice. Transplantation of total mammary cells from adenomas arising in MMTV-PyMT; Gata-3$^{+/nlslacZ}$ glands revealed a 3.6-fold increase in the frequency of tumor-initiating cells relative to those from MMTV-PyMT mice (Table 3). Cystic dilatation was also observed after transplantation of PyMT; Gata-3$^{+/nlslacZ}$ cells (Fig. 3D), indicating that the phenotypic change is intrinsic to the tumor cells. The early lesions or adenomas occurring in 35- to 42-day-old mice were next analyzed by flow cytometry using antibodies against CD14, c-kit, and CD24. We did not use CD29 for fractionation, since the expression of this marker is downregulated during tumor progression (data not shown). Downregulation of CD29 expression has also been observed for other mouse models of tumorigenesis (42). An expansion of luminal progenitor cells was observed in Gata-3 haplo-in-

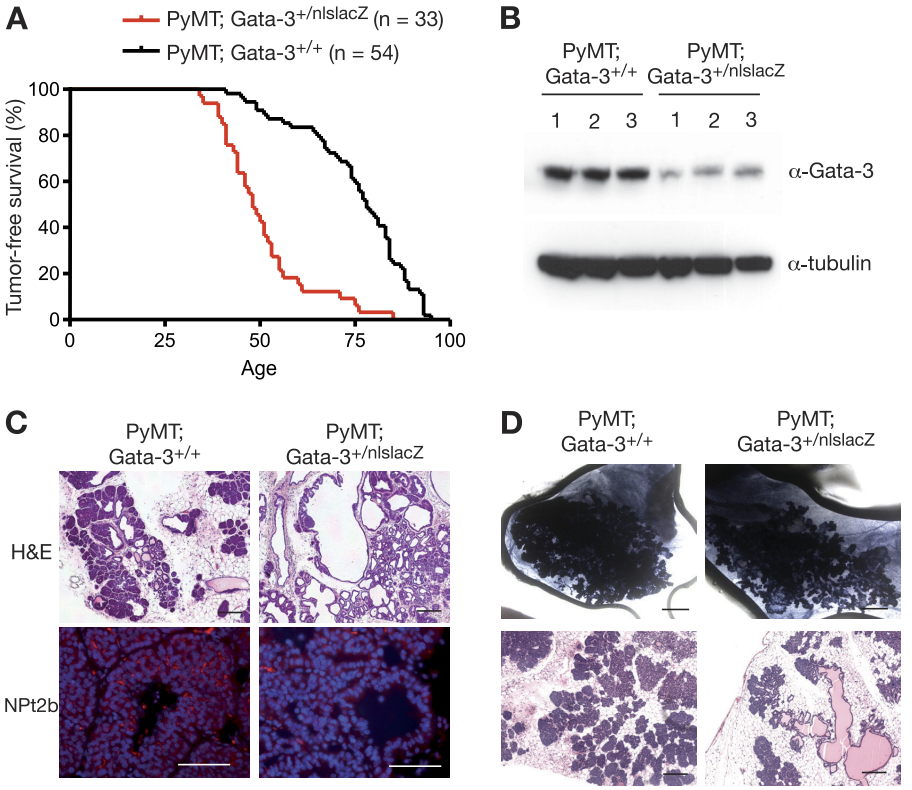

FIG. 3. Loss of one *Gata-3* allele reduces tumor latency and affects tumor morphology. (A) Kaplan-Meier tumor-free survival curves of PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice. The median tumor-free survival times were 78 days versus 48 days, respectively ($P < 0.0001$). (B) Western blot analysis showing expression level of Gata-3 in three representative PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlsLacZ}$ mammary tumors. (C) Hematoxylin and eosin (scale bars, 200 μm) and immunofluorescence staining for NPt2b in mammary lesions arising in 7-week-old PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice (scale bars, 100 μm). (D) Whole-mount (scale bars, 2 mm) and hematoxylin and eosin (scale bars, 200 μm) sections of tumors arising from transplanted unsorted cells (160 cells) from 5- to 6-week-old PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice.

TABLE 3. Frequency of tumor-initiating cells in mammary gland lesions from PyMT; Gata-3$^{+/+}$ or PyMT; Gata-3$^{+/nlslacZ}$ mice[a]

No. of cells injected per mammary fat pad	No. of tumors[b]	
	PyMT; Gata-3$^{+/+}$	PyMT; Gata-3$^{+/nlslacZ}$
10	1/12	0/12
40	0/6	1/6
160	6/12	8/12
640	7/10	10/10
2,560	11/12	12/12
10,240	5/5	6/6
Tumor-initiating frequency (95% confidence interval)	1/556 (1/972–1/318)	1/156 (1/275–1/89)
P value	0.001	

[a] Limiting dilution analysis of the tumor-initiating frequency of unsorted mammary gland cells isolated from PyMT; Gata-3$^{+/+}$ or PyMT; Gata-3$^{+/nlslacZ}$ mice at 5 to 6 weeks of age. Cells were injected into the cleared mammary fat pads of 3-week-old syngeneic FVB/N recipients and fat pads collected 10 weeks post-transplantation. Data are shown for two independent experiments.

[b] Shown as number of outgrowths per number of injected fat pads.

sufficient MMTV-PyMT; Gata-3$^{+/nlslacZ}$ compared to that in MMTV-PyMT glands (*P* value = 0.0077) and largely encompassed the alveolar progenitor population (CD14$^+$ c-kit$^{-/lo}$) (*P* value = 0.025). Furthermore, this was accompanied by a decrease in the proportion of cells sharing the mature luminal cell phenotype (CD14$^-$ c-kit$^-$) (*P* = 0.008) (Fig. 4A and B). Transplantation studies demonstrated that both luminal progenitor subpopulations (CD14$^+$ c-kit$^{-/lo}$ and c-kit$^+$) displayed a 2-fold increase in tumor-initiating capacity in MMTV-PyMT; Gata-3$^{+/nlslacZ}$ mice compared to that of the same population from MMTV-PyMT glands (Fig. 4C). Only c-kit$^+$ cells showed a significant fold difference in tumor-initiating capacity compared to the c-kit$^-$ CD14$^-$ subset. No change in the tumor-initiating frequency of the CD14$^-$ c-kit$^-$ and MaSC-enriched (CD29^{hi} CD24$^+$) subsets was evident for MMTV-PyMT versus MMTV-PyMT; Gata-3$^{+/nlslacZ}$ mice. Collectively, these data indicate that the change in phenotype that accompanies loss of a single *Gata-3* allele reflects the less differentiated nature of these tumors.

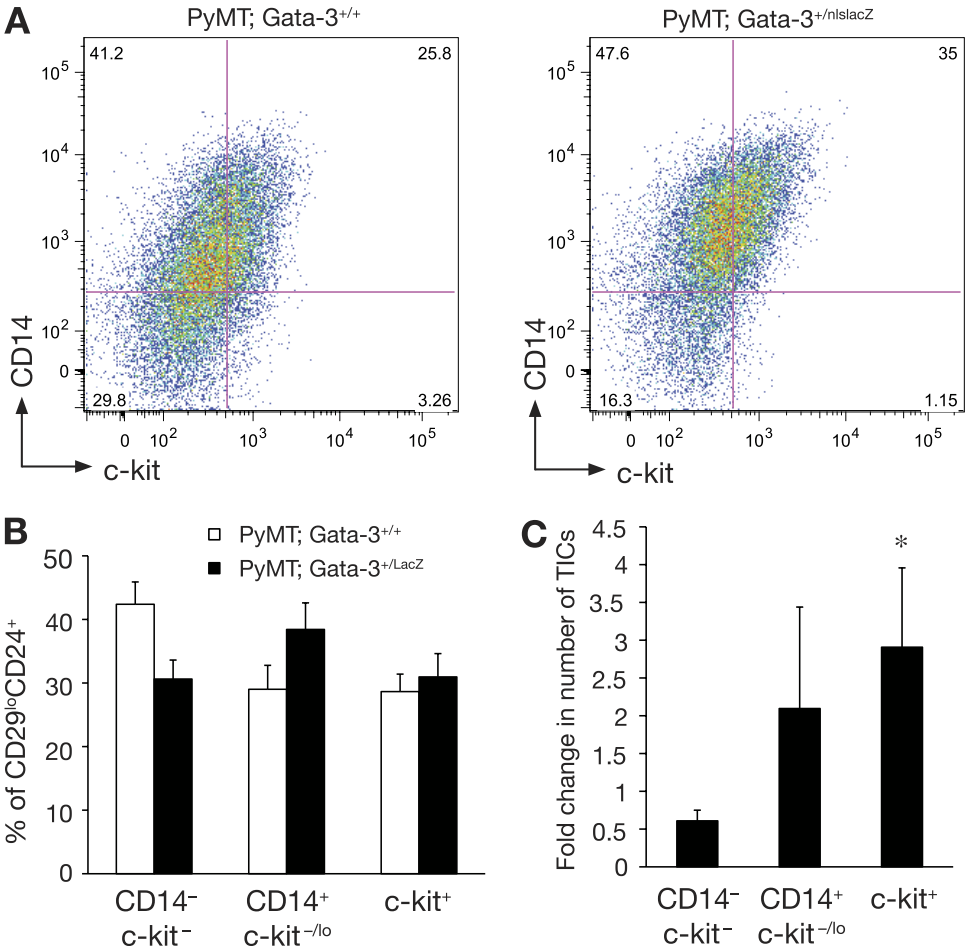

FIG. 4. Expansion of luminal progenitor subsets in PyMT; Gata-3$^{+/nlslacZ}$ mammary lesions. (A) Representative FACS dot plot showing the expression of c-kit and CD14 in Lin$^-$ CD24$^+$ cells of mammary adenomas (5 to 6 weeks) from PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice. (B) Histogram showing percentage of c-kit/CD14 cells in the CD24$^+$ populations isolated from 5- to 6-week-old PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice (means ± standard errors of the mean [SEM] of results for five animals per group). (C) Histogram showing fold change (PyMT; Gata-3$^{+/+}$ over PyMT; Gata-3$^{+/nlslacZ}$) in tumor-initiating cell (TIC) capacity after transplantation of c-kit$^+$, CD14$^+$ c-kit$^{-/lo}$, and CD14$^-$ c-kit$^-$ cells from PyMT; Gata-3$^{+/+}$ and PyMT; Gata-3$^{+/nlslacZ}$ mice (means ± SEM of results from three independent experiments). *, *P* = 0.04. Student *t* test of fold change in TIC capacity for c-kit$^+$ and CD14$^-$ c-kit$^-$ cells.

Delayed tumor onset and decreased angiogenesis in transgenic mice overexpressing Gata-3. To further investigate the role of Gata-3 in mammary tumor initiation, we generated tetracycline-inducible Gata-3-internal ribosome entry site (IRES)-green fluorescent protein (GFP) transgenic mice and crossed these with MMTV-rtTA mice (MTB) to generate double-transgenic mice. Two out of five strains (strains 16 and 28) were selected for further analysis; induction of *Gata-3* expression was observed in sorted $CD24^+$ epithelial cells from these strains following treatment with doxycycline for 4 weeks (Fig. 5A and data not shown). Mammary glands from virgin MTB; Gata-$3^{+/Tg}$ mice of either strain showed precocious alveolar formation reminiscent of early pregnancy and high levels of the alveolar marker NPt2b (Fig. 5B and data not shown). Concomitant with the diminished luminal progenitor populations evident in MTB; Gata-$3^{+/Tg}$ relative to MTB glands, the mature luminal subset ($CD14^-$ c-kit^-) was expanded (data not shown). These data support previous findings that Gata-3 induces luminal cell differentiation and is critical for alveolar differentiation (3, 20).

MTB; Gata-$3^{+/Tg}$ mice were next crossed with MMTV-PyMT mice to generate triple-transgenic mice (PyMT; MTB; Gata-$3^{+/Tg}$). Tumors in these mice developed more slowly, and tumor-free survival was significantly prolonged, with median tumor-free survival times of 86 versus 66 days for Gata-3-overexpressing mice and MMTV-PyMT mice, respectively ($P = 0.003$) (Fig. 5C). Heterogeneous expression of GFP and Gata-3 was noted within all PyMT; MTB; Gata-$3^{+/Tg}$ tumors (Fig. 5C and D) and was particularly evident as tumor progression occurred. During tumorigenesis, selection against tumor cells expressing Gata-3 may occur; this could explain why all mice eventually succumb to tumors. Tumors arising in PyMT; MTB; Gata-$3^{+/Tg}$ mice maintained a ductal architecture and more differentiated phenotype, based on the coexpression of ERα, cytokeratin 18, and smooth muscle actin in regions of the tumor expressing Gata-3 (Fig. 5E). These data are consistent with earlier observations that high levels of Gata-3 drive the differentiation of PyMT tumor cells (19). Moreover, we observed that GFP^+ cells in these triple-transgenic tumors were enriched for differentiated $CD14^-$ c-kit^- cells whereas GFP^- cells contained a higher proportion of progenitor cells ($CD14^+$ c-$kit^{-/lo}$ and c-kit^+) (Fig. 5F and G).

In vivo transplantation of limiting numbers of tumor cells (Table 4) showed that $CD24^+$ GFP^+ cells contained 5-fold fewer tumor-initiating cells than $CD24^+$ GFP^- cells. Notably, tumors arising in PyMT; MTB; Gata-$3^{+/Tg}$ mice were less angiogenic than those in PyMT mice, and an inverse correlation was observed between the percentages of $CD31^+$ and GFP^+ cells (Fig. 6A and B). Together these data indicate that tumor cells expressing high levels of Gata-3 are more differentiated and less angiogenic and have a lower propensity to initiate tumorigenesis.

Identification of a novel target of Gata-3, *caspase-14*, that delays tumor growth. To identify potential Gata-3 target genes that could play a role in abrogating tumor initiation, we compared the gene expression profiles of freshly sorted GFP^+ or GFP^- cells derived from tumors arising in PyMT; MTB; Gata-$3^{+/Tg}$ mice. In this triple-transgenic mouse model, the level of Gata-3 expression achieved via the transgene was estimated to be 3-fold higher than that in MMTV-PyMT glands. Although ERα was not detected as a differentially expressed gene on the Illumina platform, RT-PCR analysis showed a 2.3 (±0.6)-fold increase in the level of *ERα* expression in GFP^+ cells compared to that in GFP^- cells, and higher levels of ERα protein were evident in the more differentiated Gata-3-expressing regions of tumors arising in triple-transgenic mice (Fig. 5E).

Of the 129 probes (representing 121 distinct genes or ESTs) found to be differentially expressed between GFP^+ and GFP^- cells from PyMT; MTB; Gata-$3^{+/Tg}$ mice (see Table S1 in the supplemental material), we focused on the role of caspase-14, since this cysteinyl aspartate-specific protease has previously been reported to induce the differentiation of keratinocytes (7). Quantitative RT-PCR confirmed the microarray expression data, showing a 3- to 4-fold increase in *caspase-14* levels in Gata-3-overexpressing tumor cells (Fig. 7A). *caspase-14* expression was increased in GFP^+ versus GFP^- cell populations for every tumor. Furthermore, the relative increase in *caspase-14* between the cell populations correlated with that of *Gata-3* (correlation coefficient = 0.64). Immortalized mammary epithelial cells (MECs) generated from targeted mice deficient in *Gata-3* also showed increased *caspase-14* expression following retrovirus-mediated reinstallation of Gata-3 expression, suggesting that *caspase-14* may be a direct target of Gata-3 (Fig. 7B). Notably, Western blot analysis of human breast cancer cell lines showed a correlation between GATA-3 and caspase-14 expression, although this was not evident in the case of T47D cells (Fig. 7C).

Analysis of sequences 10 kb upstream and 500 bp downstream of the caspase-14 transcription start site revealed four putative Gata-3 binding sites. Chromatin immunoprecipitation studies with a GATA-3 antibody were next performed in MCF-7 human breast cancer cells since they express both GATA-3 and *caspase-14*. Significant binding to the most proximal site (−484, $P = 0.013$) as well as to one distal site (−3772, $P = 0.028$) was observed. In contrast, specific binding was not evident at an adjacent region lacking GATA-3 binding sites (−422 to −286, $P = 0.195$) or to the other distal sites (−3060, $P = 0.058$; −4392, $P = 0.067$). It is noteworthy that only the −484 and −3772 sites are conserved between human and mouse. These data indicate that *caspase-14* is a likely target gene of GATA-3 (Fig. 7D).

We next evaluated the effect of caspase-14 overexpression on mammary epithelial cell differentiation. Caspase-14 proved to be a potent inducer of differentiation in HC11 mouse mammary epithelial cells with higher levels of β-*casein* expression observed in response to a lactogenic stimulus relative to that seen for control cells or those overexpressing Gata-3 (Fig. 7E). To assess the role of caspase-14 in mammary tumorigenesis, we overexpressed these genes in a human breast cancer cell line, MDA-MB-231^{Luci}, carrying a stably integrated luciferase gene. Caspase-14 or GATA-3 lentivirus-transduced MDA-MB-231^{Luci} cells or control cells (transduced with empty lentivirus) were transplanted into the cleared mammary fat pads of immunocompromised animals and imaged for bioluminescence on the Xenogen live imager over a period of 8 weeks. Overexpression of either caspase-14 or GATA-3 significantly reduced mammary tumor growth (Fig. 7F). Thus, caspase-14 appears to function in a similar manner to its regulator, GATA-3, at least in the MDA-MB-231 xenograft model.

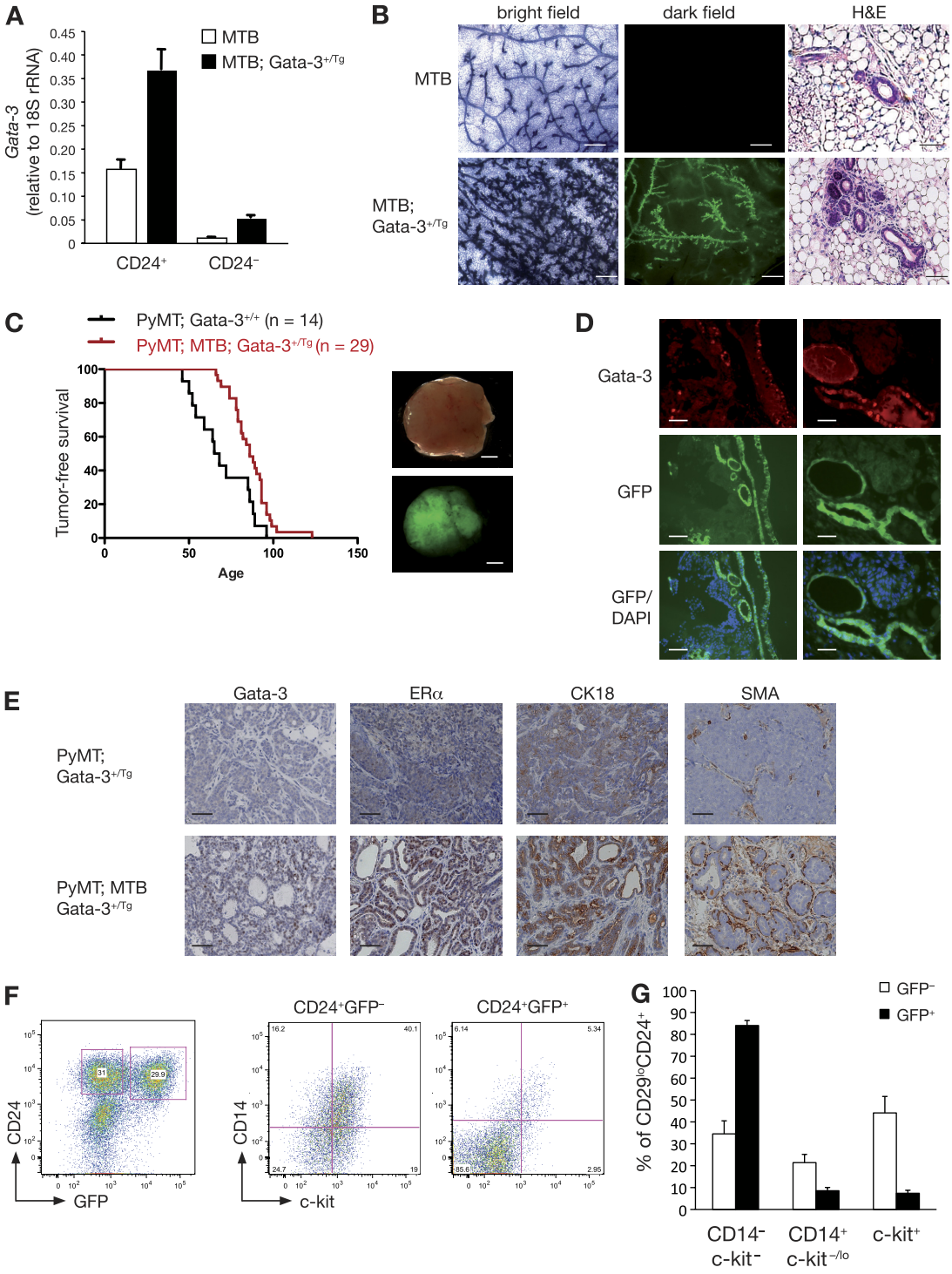

FIG. 5. Forced expression of Gata-3 delays mammary tumor formation in MMTV-PyMT mice and induces a differentiated phenotype. (A) Quantitative RT-PCR analysis showing the expression of *Gata-3* in freshly sorted $CD24^+$ and $CD24^-$ cells from 8-week-old control mice (MTB) and inducible Gata-3 transgenic mice (MTB; Gata-$3^{+/Tg}$), 4 weeks posttreatment with doxycycline. (B) Whole-mount and mammary gland sections from 7-week-old MTB and MTB; Gata-$3^{+/Tg}$ mice 4 weeks after treatment with doxycycline. Bright-field images show hematoxylin staining of mammary whole mounts. Dark-field images show GFP expression in mammary gland whole mounts. Scale bars, 250 μm. Hematoxylin and eosin staining section: scale bars, 50 μm. (C) Kaplan-Meier plots showing tumor-free survival of PyMT; Gata-$3^{+/+}$ and PyMT; Gata-$3^{+/Tg}$ mice treated with doxycycline (from 4 weeks of age). Median tumor-free survival times were 66 versus 86 days, respectively ($P < 0.05$). Right panel: bright-field and dark-field images showing GFP expression in PyMT; MTB; Gata-$3^{+/Tg}$ tumor. Scale bars, 2 mm. (D) GFP expression, DAPI staining, and Gata-3 immunofluorescence staining on consecutive sections of PyMT; MTB; Gata-$3^{+/Tg}$ tumors. Scale bars, 50 μm. (E) Immunohistochemical staining for Gata-3, estrogen receptor α (ERα), cytokeratin 18 (CK18), and smooth muscle actin (SMA) in sections from PyMT; Gata-$3^{+/Tg}$ and PyMT; MTB; Gata-$3^{+/Tg}$ tumors. Scale bars, 50 μm. (F) FACS analysis of PyMT; MTB; Gata-$3^{+/Tg}$ tumors showing the expression of GFP and CD24 in Lin^- cells (left panel). Right panel shows the expression of CD14 and c-kit in GFP^+ and GFP^- cells. (G) Histogram showing percentages of c-kit/CD14 cells in $CD24^+$ GFP^+ or $CD24^+$ GFP^- subpopulations from PyMT; MTB; Gata-$3^{+/Tg}$ mice (means ± SEM of results for three animals per group).

TABLE 4. Frequency of tumor-initiating cells in GFP subpopulations isolated from PyMT; MTB; Gata-3$^{+/Tg}$ mammary glands[a]

No. of cells injected per mammary fat pad	No. of tumors[b]	
	Lin$^-$ CD24$^+$ GFP$^-$	Lin$^-$ CD24$^+$ GFP$^+$
200	4/14	0/14
400	8/20	3/20
800	5/6	1/6
Repopulating frequency (95% confidence interval)	1/648 (1/1,051–1/399)	1/3,643 (1/9,704–1/1,368)
P value	0.0004	

[a] Limiting dilution analysis of the tumor-initiating frequency of Lin$^-$ CD24$^+$ GFP$^-$ and Lin$^-$ CD24$^+$ GFP$^+$ cells isolated from PyMT; MTB; Gata-3$^{+/Tg}$ mice. Cells were injected into the cleared mammary fat pads of 3-week-old syngeneic recipient mice and fat pads collected 10 weeks posttransplantation. Data are shown for three independent experiments.

[b] Shown as the number of tumors per number of injected fat pads. For transplanted GFP$^+$ cells, only GFP-positive tumors were counted.

Correlation between caspase-14 and GATA-3 overexpression in human DCIS samples. Since DCIS samples can express high levels of GATA-3 (28), we evaluated the relationship between caspase-14 and GATA-3 expression in a cohort of 64 human DCIS specimens with extensive follow-up data. Intense GATA-3 expression was associated with more favorable relapse-free survival compared to patients with GATA-3-negative breast tumors (Fig. 8A, $P = 0.038$), highlighting a role for GATA-3 as a key biomarker in the preinvasive stages of breast carcinogenesis. In 22 samples (37.9%), nuclear caspase-14 expression was detected, with a score of >3 in 17 samples (29.3%). Weak granular cytoplasmic caspase-14 was observed in 35 samples (60%) (score, 1 to 2), and high cytoplasmic staining (score, >3) was detected in 10 samples (17%). Consistent with results of previous studies (18), intense staining appeared more frequently in areas of necrosis. Notably, a strong association between nuclear caspase-14 and nuclear GATA-3 expression was observed, in which intense caspase-14 staining (score, >3) was positively associated with strong GATA-3 staining (score, >4) (Fig. 8B, upper panel) and low caspase-14 nuclear staining (score, 1 to 2) with lower GATA-3 levels (score, <2) (Fig. 8B, lower panel) ($P = 6.6 \times 10^{-11}$). Caspase-14-positive tumors frequently expressed GATA-3, but GATA-3-positive tumors did not always express high levels of caspase-14. No association was evident between caspase-14 expression and relapse-free survival, possibly due to the small number of specimens (22) expressing caspase-14 in this patient cohort.

DISCUSSION

GATA-3 is a definitive marker of ER-positive breast cancers in which high levels of GATA-3 correlate with high ER expression, while low levels are strongly associated with ER negativity, higher histological grade, lymph node positivity, and a higher propensity for metastasis (16, 26). Meta-analysis of breast cancer microarray data sets initially indicated that GATA-3 may be an independent predictor of clinical outcome (26), but other studies have suggested that GATA-3 lacks independent prognostic value (1, 45). To further understand the role of Gata-3 in breast tumorigenesis, we have explored its involvement in tumor initiation. Our data reveal that GATA-3 levels affect the tumor-initiating capacity of luminal but not basal stem/progenitor cells, thus indicating a critical role for Gata-3 in regulating the differentiation of luminal progenitor cells in either normal or malignant mammary tissue.

The mammary epithelial hierarchy can serve as a useful tool to understand cell types predisposed to neoplasia (44). In the normal mammary gland, Gata-3 is expressed in multiple epithelial cell types, including MaSCs and those restricted to the luminal lineage. Notably, *Gata-3* expression was found to be essential for the activity of MaSCs or their immediate descendants, as demonstrated through *in vivo* transplantation of *Gata-3*-depleted MaSCs. In the luminal lineage, two luminal progenitor subtypes were identified and found to express comparable levels of Gata-3; the CD14$^+$ c-kit$^{-/lo}$ subset likely corresponds to a committed alveolar progenitor cell, while the c-kit$^+$ subpopulation appears to lie further upstream. Mature luminal cells appear to have the highest levels of Gata-3 protein, commensurate with its role in differentiation. Interestingly, we found that all four of the epithelial cellular subsets exhibited tumor-initiating capacity in the early stages of tumorigenesis in the MMTV-PyMT model, although a considerably lower frequency was observed for the mature luminal subset. Presumably, the middle T antigen elicited dedifferentiation in a proportion of mature cells.

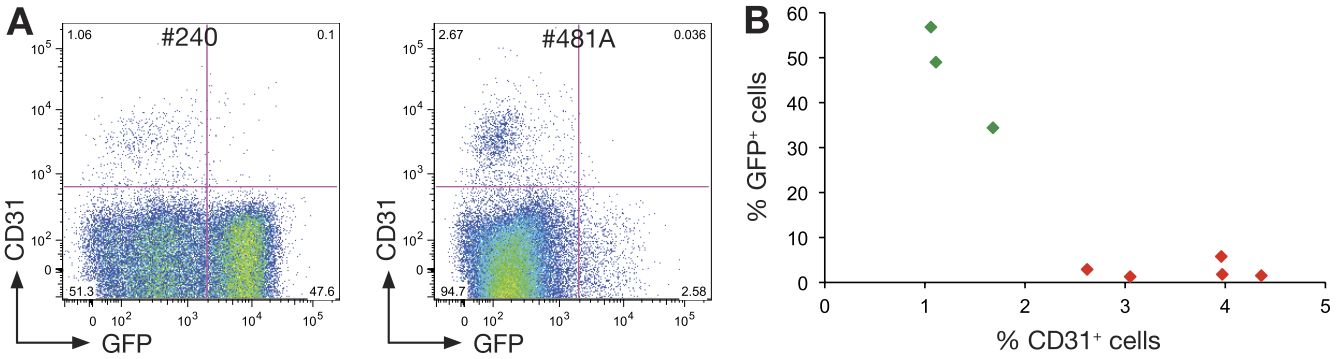

FIG. 6. Overexpression of Gata-3 reduces tumor angiogenesis in MMTV-PyMT mice. (A) Representative dot plot showing the expression of GFP and CD31 in PyMT; MTB; Gata-3$^{+/Tg}$ mammary tumors. (B) Scattered plot showing the inverse correlation between CD31 expression and GFP expression in eight individual PyMT; MTB; Gata-3$^{+/Tg}$ mammary tumors.

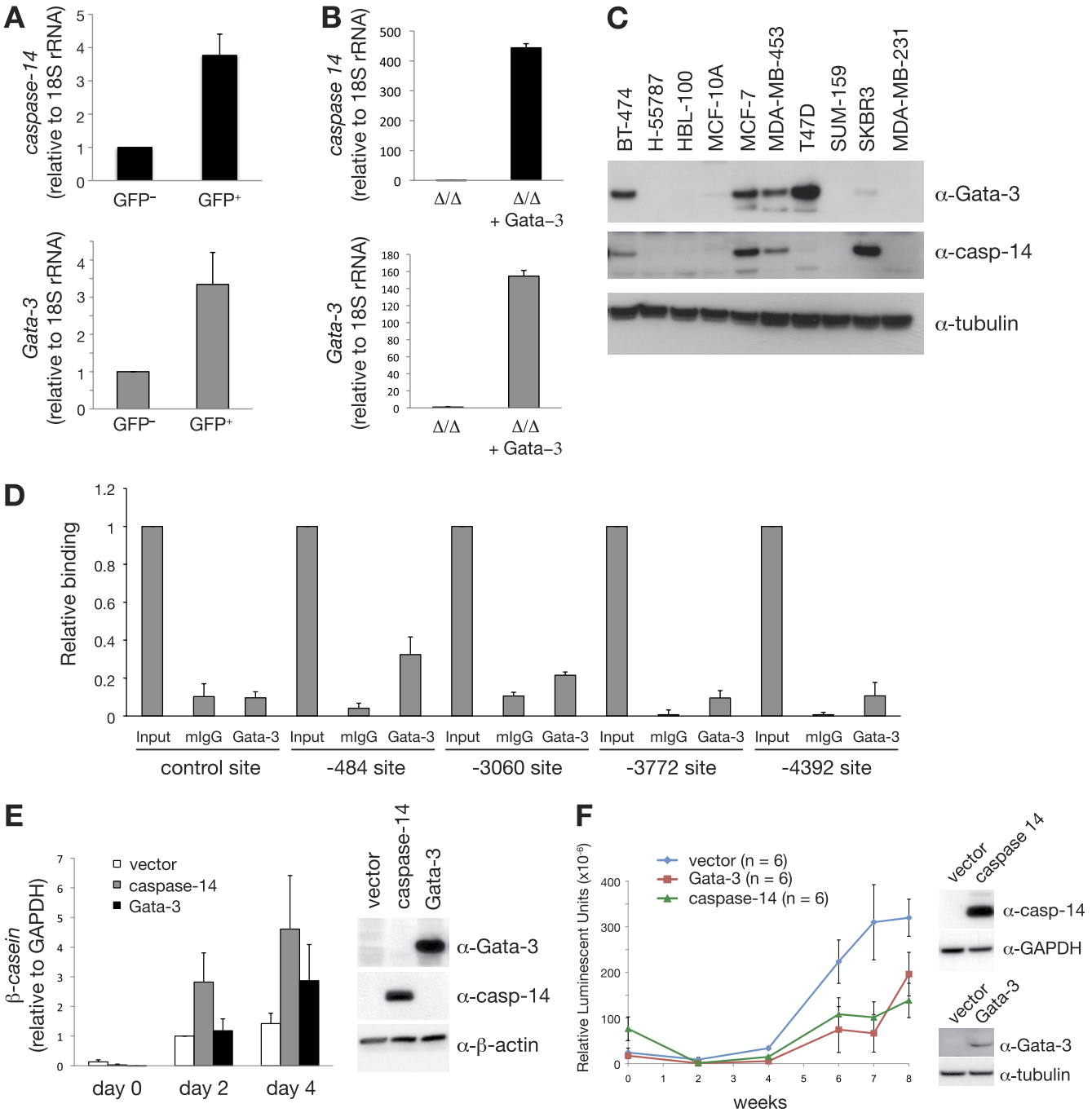

FIG. 7. The Gata-3 target gene, *caspase-14*, promotes differentiation and delays breast tumor formation. (A) Expression analysis of *caspase-14* and *Gata-3* in GFP$^+$ and GFP$^-$ cells from PyMT; MTB; Gata-3$^{+/Tg}$ mammary tumors (means ± standard errors of the mean [SEM] of results for seven tumors per group). (B) Expression analysis of *caspase-14* and *Gata-3* in immortalized mammary epithelial cells isolated from Gata-3$^{f/f}$ mice, transduced sequentially with pMSCV-cre retrovirus and then Gata-3-expressing (or empty) pBABE retrovirus. (C) Western blot showing the expression of GATA-3, caspase-14, and tubulin in human breast cancer cell lines. (D) ChIP analysis of endogenous GATA-3 binding to four putative GATA-3 binding sites identified within a 10-kb upstream regulatory region of caspase-14 (−484, −3,060, −3,772, and −4,392 from the transcription start site) and a flanking region of caspase-14 promoter with no GATA-3 binding site (control) in MCF-7 cells. Unprecipitated chromatin provided the input control (mean ± SEM of results of three experiments). Mouse IgG and Gata-3 ChIP values for each region were compared using the Student *t* test: −484 (P = 0.013), −3,060 (P = 0.058), −3,772 (P = 0.028), −4,392 (P = 0.067), control (P = 0.195). (E) β-*casein* mRNA expression in HC11 cells transduced with control (empty pFU-TA-GFP), pFU-TA-GFP-Gata-3-, or caspase-14-expressing lentiviruses (means ± SEM of results of four independent experiments). Right panel shows Western blot analysis of Gata-3, caspase-14, and β-actin in transduced HC11 cells. (F) Kinetics of tumor formation after orthotopic transplantation of 500,000 MDA-MB-231^{Luci} cells transduced with either pFU-TA-GFP-Gata-3- or caspase-14-expressing lentiviruses or control (vector) virus (means ± SEM of results for six animals per group). Right panel shows Western blot analysis of caspase-14, GATA-3, GAPDH, and tubulin expression in transduced MDA-MB-231^{Luci} cells at the time of transplantation.

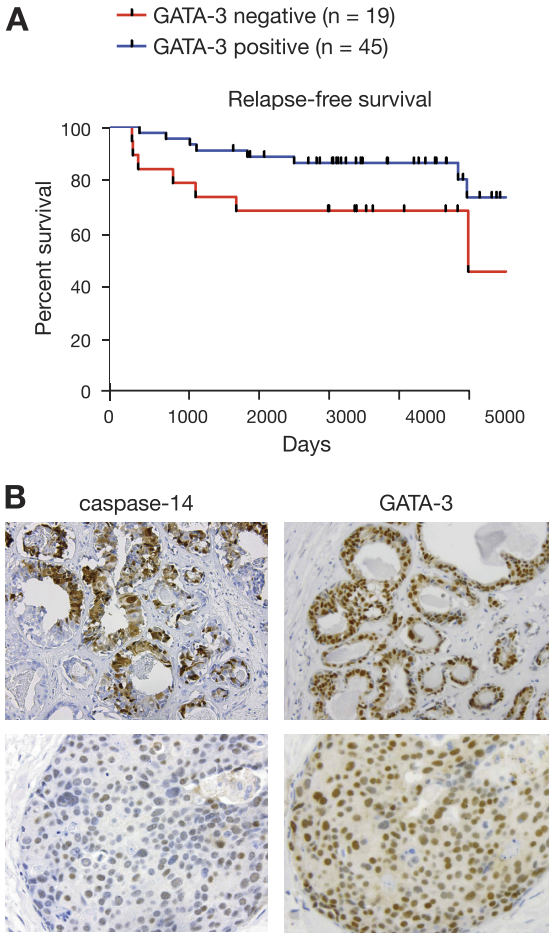

FIG. 8. High GATA-3 levels in human ductal carcinoma *in situ* (DCIS) are associated with better relapse-free survival. (A) Kaplan-Meier curve showing relapse-free survival in GATA-3-positive and -negative DCIS samples. (B) Representative immunohistostaining for GATA-3 and caspase-14 in human DCIS specimens.

Gata-3 expression levels directly influence the tumor-initiating capacity of the luminal progenitor populations. Loss of a single *Gata-3* allele from hyperplastic or early neoplastic lesions arising in MMTV-PyMT mice augmented the tumor-initiating ability of the two luminal progenitor subpopulations but did not alter that of the MaSC-enriched subset. The histopathological change (cystic dilatation) that accompanied Gata-3 haplo-insufficiency in the MMTV-PyMT model does not reflect a change in the tumor-initiating cell type upon loss of *Gata-3*. Rather, it appears to correlate with the proportion of progenitor cells in the adenomas and hence the differentiation status of the lesions. In the converse model, in which Gata-3 was conditionally expressed in triple-transgenic mice, the proportion of progenitor cells declined and tumor-initiating capacity was diminished. These data support previous observations that Gata-3 is a predictor of tumor differentiation in MMTV-PyMT mice (19). However, our findings that Gata-3 plays an important role early in neoplasia differ somewhat from the observation that targeted deletion of Gata-3 in early tumors is not sufficient to induce tumor progression in MMTV-PyMT; WAP-rtTA-cre; Gata-$3^{f/f}$ mice. The apparent discrepancy may result from the rapid and widespread loss of tumor cells following cre-induction in this specific model (19).

Selection against Gata-3-expressing cells occurs during malignant progression. MMTV-PyMT tumors overexpressing Gata-3 were highly differentiated, and they exhibited a ductal phenotype with ERα- and K18-positive luminal cells surrounded by SMA-positive cells visible within Gata-3-positive regions of the tumor. However, loss of Gata-3 expression occurred as tumor progression occurred, reminiscent of late-stage Gata-3-negative carcinomas that have been previously described (19). Interestingly, the expression of ERα, a target gene of Gata-3, also decreased with tumor progression in MMTV-PyMT mice (25). Loss of Gata-3 seems to mark the onset of tumor metastasis, since disseminated tumor cells and lung metastases in MMTV-PyMT tumor-bearing mice were found to be negative for the Gata-3 transcription factor (19). The development of these tumors is unlikely to result from the expansion of a stem-like population in preneoplastic tissue given the implicit expression of Gata-3 in MaSCs (or bipotent progenitor cells) and the striking effect of *Gata-3* loss on mammary repopulating capacity. Indeed, these tumors (19) may develop from luminal cells that have lost Gata-3 expression during progression rather than from more primitive basal precursor cells. Epigenetic silencing of the *Gata-3* gene may represent one mechanism that mediates loss of expression, since a number of methylated cytosine residues have been identified within the *Gata-3* gene in MMTV-PyMT and human breast tumors (19, 47).

Gata-3 expression levels were found to alter the tumor microenvironment in the context of the MMTV-PyMT model. We observed that tumors arising in triple-transgenic mice (PyMT; MTB; Gata-$3^{+/Tg}$) were less angiogenic than those in MMTV-PyMT mice. It is well recognized that the neovasculature plays an essential role in tumor progression and dissemination to other organs (27). However, the mechanism by which Gata-3-overexpressing tumor cells delay the formation of the neovasculature remains to be determined. The reduced neovasculature in these tumors may limit extravasation of tumor cells to the bloodstream and thus reduce their metastatic propensity (8, 19). The tumor-associated endothelial cells may also be indirectly involved in promoting the selection of Gata-3-negative tumor cells.

Several genes have been implicated as targets of the Gata-3 transcription factor, including ERα, FoxA1, and p18 (1, 9, 29). Here we have identified *caspase-14* as a likely transcriptional target of Gata-3 and show that it promotes the differentiation of murine HC11 mammary epithelial cells. Chromatin immunoprecipitation studies in human breast cancer cells showed that GATA-3 binds distal and proximal sites in the promoter of the human *caspase-14* gene, a member of the cysteinyl aspartate-specific protease family. Caspase cleavage at an atypical site has been described only during epidermal differentiation and does not seem to be required for caspase-14 activity (6, 18). In mammary epithelial cells, expression of caspase-14 induced differentiation to milk-producing cells, reminiscent of its role in keratinocyte differentiation and cornification (7). Cleavage of caspase-14 was not detected in our samples, consistent with results of previous studies (6, 18).

Forced expression of caspase-14 in human breast cancer cells significantly delayed tumor growth in an orthotopic breast

tumor xenograft model. Likewise, caspase-14 overexpression in a xenograft model of skin cancer was shown to reduce tumorigenicity (14). Moreover, caspase-14 expression is deregulated in epithelial malignancies and is associated with the differentiation status of lung squamous cell and cervical carcinomas (18, 21). In gastric and ovarian cancers, higher caspase-14 expression correlated with improved patient survival (21), while in breast cancer, high caspase-14 levels were reported in both preinvasive lesions and invasive carcinomas (21). In preinvasive DCIS specimens, we observed a significant correlation between GATA-3 and caspase-14 expression. Although caspase-14 expression did not exhibit prognostic value in this cohort, further studies on a larger cohort of patients may shed light on its role as a predictive marker.

GATA-3 appears to be an important predictive biomarker in DCIS, based on the strong association between GATA-3 expression and relapse-free survival in these patients. Despite treatment of DCIS by surgical lumpectomy together with radiotherapy or hormonal therapy, approximately 5 to 10% of women develop invasive carcinomas within 5 years. The identification of robust biomarkers to predict clinical outcome in patients with DCIS would be a significant advance. Given the prognostic value of GATA-3 in DCIS, it may be possible to use compounds that activate Gata-3 expression as therapeutic agents for the prevention of tumors in these patients.

ACKNOWLEDGMENTS

We are grateful to L. Chodosh for the MTB strain, C. Ormandy and B. Pal for plasmids, K. Stoev and A. Morcom for animal care, E. Takano, B. Helbert, and C. Young for genotyping, S. Mihajlovic and E. Tsui for histology support, and J. Biber for antibody. We thank the Australian Genome Research Foundation for RNA bioanalyses.

This work was supported by Cancer Australia, the National Breast Cancer Research Foundation, the Victorian Breast Cancer Research Consortium, the National Health and Medical Research Council Australia, the Victorian State Government OIS (Operational Infrastructure Support), and the Australian Government NHMRC IRIISS. M.L.-A.L. was supported by an Australian Research Council Queen Elizabeth II Fellowship, and J.E.V. and G.J.L. were supported by NHMRC Fellowships.

REFERENCES

1. **Albergaria, A., et al.** 2009. Expression of FOXA1 and GATA-3 in breast cancer: the prognostic significance in hormone receptor-negative tumours. Breast Cancer Res. **11:**R40.
2. **Arnold, J. M., et al.** 2010. Frequent somatic mutations of GATA3 in non-BRCA1/BRCA2 familial breast tumors, but not in BRCA1-, BRCA2- or sporadic breast tumors. Breast Cancer Res. Treat. **119:**491–496.
3. **Asselin-Labat, M. L., et al.** 2007. Gata-3 is an essential regulator of mammary-gland morphogenesis and luminal-cell differentiation. Nat. Cell Biol. **9:**201–209.
4. **Bertucci, F., et al.** 2000. Gene expression profiling of primary breast carcinomas using arrays of candidate genes. Hum. Mol. Genet. **9:**2981–2991.
5. **Bouras, T., et al.** 2008. Notch signaling regulates mammary stem cell function and luminal cell-fate commitment. Cell Stem Cell **3:**429–441.
6. **Chien, A. J., R. B. Presland, and M. K. Kuechle.** 2002. Processing of native caspase-14 occurs at an atypical cleavage site in normal epidermal differentiation. Biochem. Biophys. Res. Commun. **296:**911–917.
7. **Denecker, G., et al.** 2007. Caspase-14 protects against epidermal UVB photodamage and water loss. Nat. Cell Biol. **9:**666–674.
8. **Dydensborg, A. B., et al.** 2009. GATA3 inhibits breast cancer growth and pulmonary breast cancer metastasis. Oncogene **28:**2634–2642.
9. **Eeckhoute, J., et al.** 2007. Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer. Cancer Res. **67:**6477–6483.
10. **Gentleman, R. C., et al.** 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. **5:**R80.
11. **Guy, C. T., R. D. Cardiff, and W. J. Muller.** 1992. Induction of mammary tumors by expression of polyomavirus middle T oncogene: a transgenic mouse model for metastatic disease. Mol. Cell. Biol. **12:**954–961.
12. **Herschkowitz, J. I., et al.** 2007. Identification of conserved gene expression features between murine mammary carcinoma models and human breast tumors. Genome Biol. **8:**R76.
13. **Hoch, R. V., D. A. Thompson, R. J. Baker, and R. J. Weigel.** 1999. GATA-3 is expressed in association with estrogen receptor in breast cancer. Int. J. Cancer **84:**122–128.
14. **Hsu, S., et al.** 2007. Expression of caspase-14 reduces tumorigenicity of skin cancer cells. In Vivo **21:**279–283.
15. **Hu, Y., and G. K. Smyth.** 2009. ELDA: extreme limiting dilution analysis for comparing depleted and enriched populations in stem cell and other assays. J. Immunol. Methods **347:**70–78.
16. **Jenssen, T. K., W. P. Kuo, T. Stokke, and E. Hovig.** 2002. Associations between gene expressions in breast cancer and patient survival. Hum. Genet. **111:**411–420.
17. **Kendrick, H., et al.** 2008. Transcriptome analysis of mammary epithelial subpopulations identifies novel determinants of lineage commitment and cell fate. BMC Genomics **9:**591.
18. **Koenig, U., W. Sommergruber, and S. Lippens.** 2005. Aberrant expression of caspase-14 in epithelial tumors. Biochem. Biophys. Res. Commun. **335:**309–313.
19. **Kouros-Mehr, H., et al.** 2008. GATA-3 links tumor differentiation and dissemination in a luminal breast cancer model. Cancer Cell **13:**141–152.
20. **Kouros-Mehr, H., E. M. Slorach, M. D. Sternlicht, and Z. Werb.** 2006. GATA-3 maintains the differentiation of the luminal cell fate in the mammary gland. Cell **127:**1041–1055.
21. **Krajewska, M., et al.** 2005. Tumor-associated alterations in caspase-14 expression in epithelial malignancies. Clin. Cancer Res. **11:**5462–5471.
22. **Kurek, D., G. A. Garinis, J. H. van Doorninck, J. van der Wees, and F. G. Grosveld.** 2007. Transcriptome and phenotypic analysis reveals Gata3-dependent signalling pathways in murine hair follicles. Development **134:**261–272.
23. **Lim, E., et al.** 2009. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. **15:**907–913.
24. **Lim, E., et al.** 2010. Transcriptome analyses of mouse and human mammary cell subpopulations reveal multiple conserved genes and pathways. Breast Cancer Res. **12:**R21.
25. **Lin, E. Y., et al.** 2003. Progression to malignancy in the polyoma middle T oncoprotein mouse breast cancer model provides a reliable model for human diseases. Am. J. Pathol. **163:**2113–2126.
26. **Mehra, R., et al.** 2005. Identification of GATA3 as a breast cancer prognostic marker by global gene expression meta-analysis. Cancer Res. **65:**11259–11264.
27. **Nguyen, D. X., P. D. Bos, and J. Massague.** 2009. Metastasis: from dissemination to organ-specific colonization. Nat. Rev. Cancer **9:**274–284.
28. **Park, S. Y., et al.** 2010. Heterogeneity for stem cell-related markers according to tumor subtype and histologic stage in breast cancer. Clin. Cancer Res. **16:**876–887.
29. **Pei, X. H., et al.** 2009. CDK inhibitor p18(INK4c) is a downstream target of GATA3 and restrains mammary luminal progenitor cell proliferation and tumorigenesis. Cancer Cell **15:**389–401.
30. **Perou, C. M., et al.** 2000. Molecular portraits of human breast tumours. Nature **406:**747–752.
31. **Ritchie, M. E., et al.** 2006. Empirical array quality weights in the analysis of microarray data. BMC Bioinformatics **7:**261.
32. **Shackleton, M., et al.** 2006. Generation of a functional mammary gland from a single stem cell. Nature **439:**84–88.
33. **Shi, W., A. Oshlack, and G. K. Smyth.** 2010. Optimizing the noise versus bias trade-off for Illumina whole genome expression BeadChips. Nucleic Acids Res. **38:**e204.
34. **Smyth, G. K.** 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. Stat. Appl. Genet. Mol. Biol. **3:**Article3.
35. **Smyth, G. K., J. Michaud, and H. S. Scott.** 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics **21:**2067–2075.
36. **Sørlie, T., et al.** 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. U. S. A. **98:**10869–10874.
37. **Sorlie, T., et al.** 2003. Repeated observation of breast tumor subtypes in independent gene expression data sets. Proc. Natl. Acad. Sci. U. S. A. **100:**8418–8423.
38. **Sotiriou, C., et al.** 2003. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. Proc. Natl. Acad. Sci. U. S. A. **100:**10393–10398.
39. **Stingl, J.** 2009. Detection and analysis of mammary gland stem cells. J. Pathol. **217:**229–241.
40. **Tan, E. Y., et al.** 2007. BNIP3 as a progression marker in primary human breast cancer; opposing functions in in situ versus invasive cancer. Clin. Cancer Res. **13:**467–474.
41. **Usary, J., et al.** 2004. Mutation of GATA3 in human breast tumors. Oncogene **23:**7669–7678.
42. **Vaillant, F., et al.** 2008. The mammary progenitor marker CD61/beta3 in-

tegrin identifies cancer stem cells in mouse models of mammary tumorigenesis. Cancer Res. **68:**7711–7717.
43. **van Doorninck, J. H., et al.** 1999. GATA-3 is involved in the development of serotonergic neurons in the caudal raphe nuclei. J. Neurosci. **19:**RC12.
44. **Visvader, J. E.** 2009. Keeping abreast of the mammary epithelial hierarchy and breast tumorigenesis. Genes Dev. **23:**2563–2577.
45. **Voduc, D., M. Cheang, and T. Nielsen.** 2008. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. Cancer Epidemiol. Biomarkers Prev. **17:**365–373.
46. **West, M., et al.** 2001. Predicting the clinical status of human breast cancer by using gene expression profiles. Proc. Natl. Acad. Sci. U. S. A. **98:**11462–11467.
47. **Yan, P. S., et al.** 2000. CpG island arrays: an application toward deciphering epigenetic signatures of breast cancer. Clin. Cancer Res. **6:**1432–1438.