Nonlinear progression across the occult transition establishes cancer lethality

Authors. Joshua D. Ginzel, Henry Chapman, Joelle E. Sills, Edwin J. Allen, Lawrence S. Barak, Robert D. Cardiff, Alexander D. Borowsky, H. Kim Lyerly, Bruce W. Rogers* & Joshua C. Snyder*.

*Co-corresponding authors

Highlights

- Growth during the invisible phase of tumorigenesis is a nonlinear process
- Field size and field growth rate are uncoupled from metastatic potential
- Occult transition rates vary by genotype
- Nascent lethal lesions are currently undetectable

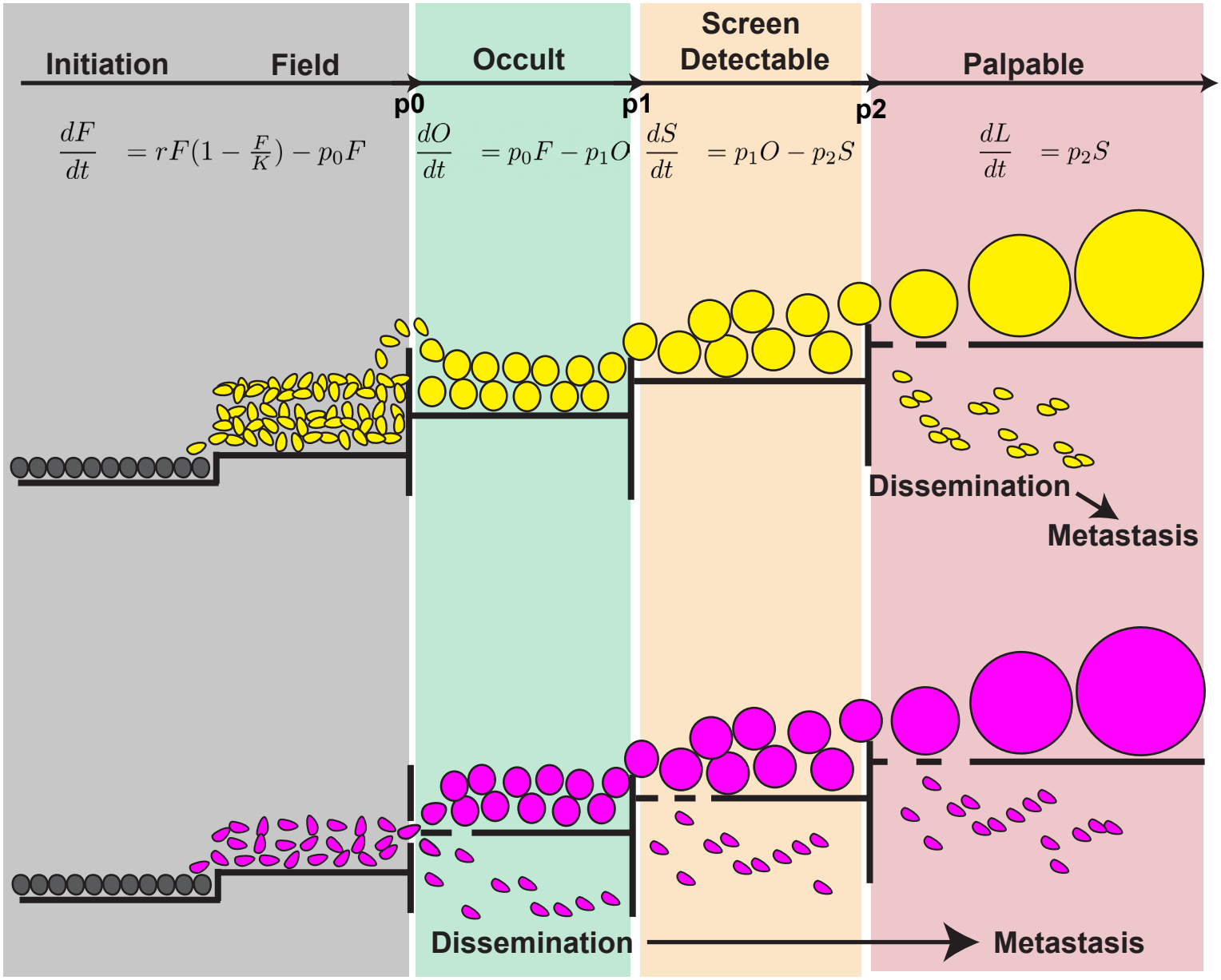
Summary

Cancer screening is based upon a linear model of growth and invasion. Yet, early dissemination during the lengthy pre-diagnostic phase suggests that nonlinearity in growth can also occur. Therefore, we quantitatively traced the invisible and visible phases of tumorigenesis in the mammary gland for more than two-thousand tumors. Dynamic mathematical models of the invisible phase revealed an occult checkpoint resulting in nonlinear progression of transformed field cells. We found that expansile fields have increased dwell time at the occult checkpoint resulting in a large reservoir of image detectable precursors prior to invasion. In contrast, slowly proliferating lesions disseminate early and then transition rapidly through an occult checkpoint in a process we term nascent lethality. Our data illustrate how nonlinear growth across an occult checkpoint can account for a paradoxical increase in early-stage cancer detection without a dramatic reduction in metastatic burden.

Introduction

Although tumor driver gene mutations are among the earliest events in cancer, the mechanisms underlying progression to a clinically detectable and ultimately lethal tumor remains unclear. Cancer initiating mutations alone are not deterministic, as even fully transformed clones can stall or extinguish entirely without ever causing disease (Huels et al. 2018; van Neerven et al. 2021; Brown et al. 2017; Hill et al. 2023). In addition, for any one clone that progresses to cancer, billions of transformed cells exist that will never successfully generate a tumor (Evans and DeGregori 2021). A major conceptual problem is that most of the disease process occurs over the course of several years, if not decades, and often requires cancer cell extrinsic events to promote progression (Gerstung et al. 2020; Swanton et al. 2024). Tumor volume doubling times (TVDT) observed during the visible phase and described by linear models, do not extrapolate to the initiating oncogenic event. Therefore, the formative years of tumor development occur during an “invisible” phase of tumorigenesis that directly impacts the overall effectiveness of screening and intervention (Ian Macdonald 1966; I. Macdonald 1951).

Graphical Abstract



Cancer mitigation strategies seek to reduce metastatic burden by screening and treating at an early-stage (Castle et al. 2024). Early intervention decreases the risk of metastasis in colorectal, stomach, and cervical cancers. However, growing epidemiological evidence suggests the benefits are not shared across all cancers (H. G. Welch, Kramer, and Black 2019). For instance, not all screen-detectable ductal carcinomas (DCIS) progress to invasive ductal carcinoma (IDC). Moreover, despite significant early detection efforts in the breast, metastatic burden remains unchanged (H. G. Welch et al. 2016; H. G. Welch, Kramer, and Black 2019). Unfortunately, the invasive potential of a breast cancer clone is likely established before screen detection (Casasent et al. 2018; Wang et al. 2023; Lips et al. 2022) resulting in early dissemination prior to the detectable or "visible" phase of the primary tumor (Hu et al. 2020; Klein 2020; 2009; Harper et al. 2016). Thus, major inflection points in the disease trajectory are likely established prior to detection, thereby rendering cancer mitigation strategies less effective. Cancer interception strategies seek instead to prevent or intercept a premalignant clone during the invisible phase prior to detection (Castle et al. 2024). However, studying the dynamics of premalignancy in immune intact and autochthonous models has been challenging due to the inability to image and quantify the transition of a premalignant clone to a visible tumor (Domchek and Vonderheide 2024).

We previously developed Cancer rainbow mouse (Crainbow) models for fluorescently visualizing and barcoding premalignant clones (Boone et al. 2019; Ginzel et al. 2021). Our previous work in the mammary gland validated a HER2 Crainbow model that expresses multiple clinically relevant isoforms of the protooncogene HER2 including wild-type (WT), an exon16-splice isoform (d16), and an N-terminally truncated isoform (p95) (Ginzel et al. 2021). We found that the visible phase of tumorigenesis in HER2 Crainbow (HBOW) mice is characterized by isoform dependent differences in tumor phenotypes, suggesting that even in mouse models an early inflection point in the disease process occurs prior to detection. Our goal here is to lineage trace and mathematically model the invisible phase of tumorigenesis. We demonstrate that the transition from premalignancy to malignancy is exceptionally rare, progresses nonlinearly, and is the major inflection point in the overall trajectory of cancer.

Results

The initiation phase of tumorigenesis begins at oncogene expression. In our model, HER2 activation occurs no later than two-weeks of age (Figure 1A). Yet, despite the qualitative increase in transformed WT and d16 cells compared to p95, we previously found that p95 tumors were nearly as abundant and more likely to have invasive properties. Therefore, we first determined if these differences could be attributed to the palpable (i.e. visible) phase of tumorigenesis. HBOW mice were observed for the first sign of a palpable tumor and then measured at least two times weekly until total tumor burden reached 2000mm^3 or a humane endpoint. At endpoint, mice were imaged to retrieve the fluorescent protein (FP) barcode for each palpable tumor (Figure 1B). Tumor free survival and overall survival were both calculated in twenty-eight mice. HBOW mice have a fifty-percent tumor free survival at 113 days and all mice develop tumors by 148 days (Figure 1C). Fifty percent of HBOW mice reach tumor burden by 168 days and all mice reach tumor burden by 198 days (Figure 1D). All mice were imaged using an epifluorescent upright dissecting microscope and palpable tumors were co-registered with tumor genotype using the Crainbow fluorescent barcode (Figure 1E). Of the ninety palpable

Figure 1

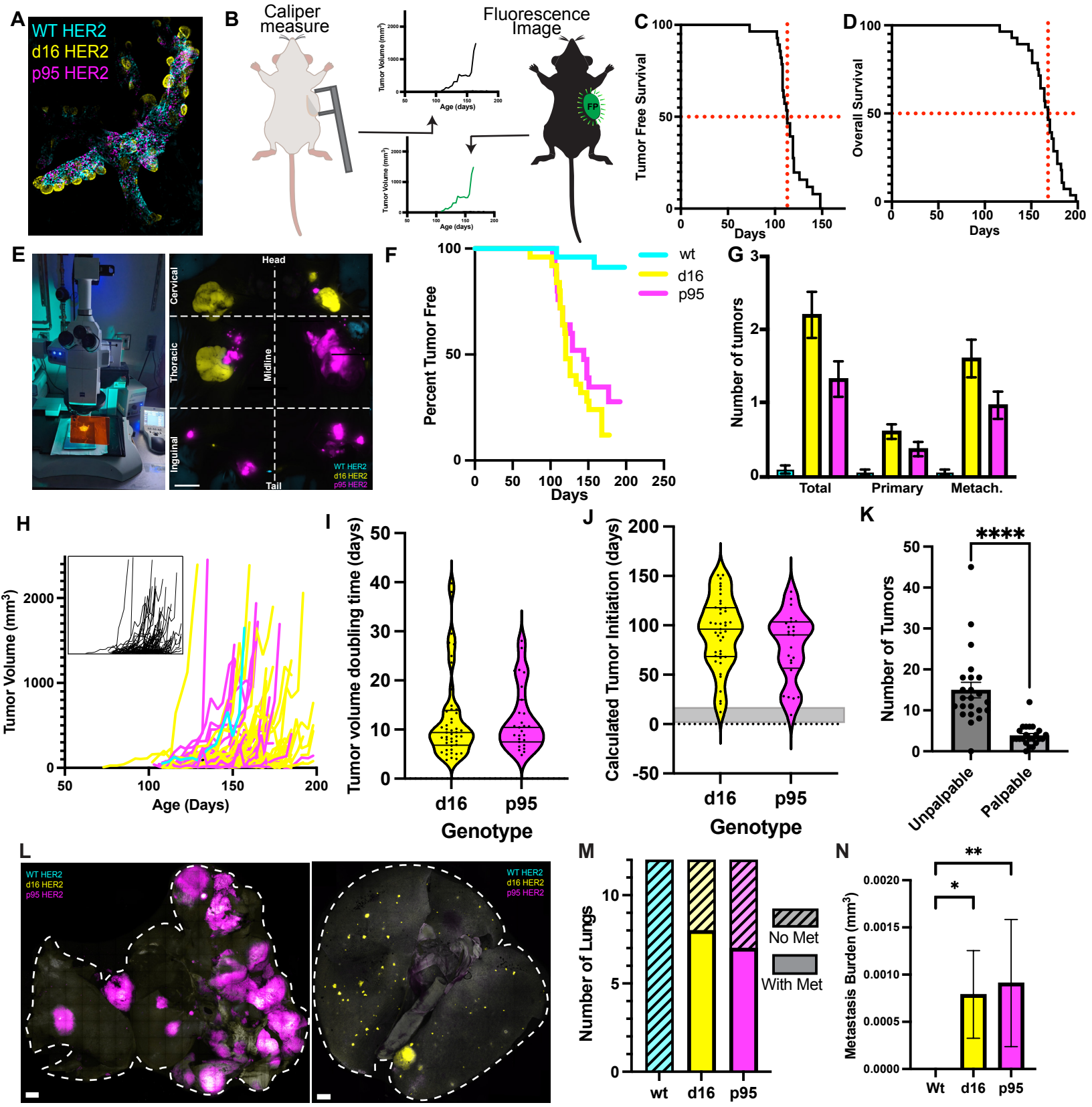

A
WT HER2
d16 HER2
p95 HER2
B
Caliper measure
Fluorescence Image
Tumor Volume (mm³)
Age (days)
FP
C
Tumor Free Survival
Days
D
Overall Survival
Days
E
Head
Cervical
Thoracic
Midline
Inguinal
Tail
WT HER2
d16 HER2
p95 HER2
F
Percent Tumor Free
Days
wt
d16
p95
G
Number of tumors
Total
Primary
Metach.
H
Tumor Volume (mm³)
Age (Days)
I
Tumor volume doubling time (days)
d16
p95
Genotype
J
Calculated Tumor Initiation (days)
d16
p95
Genotype
K
****
Number of Tumors
Unpalpable
Palpable
L
WT HER2
d16 HER2
p95 HER2
M
Number of Lungs
No Met
With Met
wt
d16
p95
N
**
*
Metastasis Burden (mm³)
Wt
d16
p95

Figure 1. Quantifying the visible phase of tumorigenesis. A, Whole mammary gland imaging of fluorescent Crainbow cells expressing mTFP:WT HER2 (cyan), eYFP:d16 HER2 (yellow), mKO:p95 HER2 (magenta) at 2 weeks. **B,** Tumors palpable through the skin were measured twice weekly with calipers until tumor burden endpoint. Mice were imaged at necropsy to retrieve fluorescent protein expression and registered to the palpation data. **C,** Tumor free survival of 28 mice palpated biweekly. Red lines represent 50% on the y-axis and 113 days on the x-axis. **D,** Survival until tumor burden endpoint of 28 mice. Red lines represent 50% on the y-axis and 168 days on the x-axis. **E,** Whole mouse imaging reveals tumor genotype by fluorescence at necropsy. **F,** Tumor free survival for calculated by the time to the first palpable tumor per mouse of each color as determined by whole mouse imaging. **G,** Number of tumors expressing each HER2 isoform per mouse at tumor burden endpoint (N=28 mice). From left to right: Total tumors observed, Primary tumor observed, All metachronous tumors observed. **H,** Observed tumor volume growth over time colored by genotype. Inset shows growth before Crainbow registration. **I,** Tumor volume doubling time of each observed tumor by genotype. Tumors with <3 measurements were not included. **J,** Calculated time of tumor initiation from tumor volume doubling time and first palpation observation. Gray shade indicates actual recombination and oncogene expression between 0 and 14 days old. **K,** Number of tumors per mouse measured by whole mouse imaging that were palpable or too small to palpate (unpalpable). Statistical significance was determined by Mann-Whitney test. P value is represented as follows:
****=p<0.0001 **L,** Representative 3D whole lung imaging for HBOW fluorescence. Lung volume outlined with white dotted line. **M,** Number of lungs with at least one metastatic lesion expressing mTFP:WT HER2 (cyan), eYFP:d16 HER2 (yellow), mKO:p95 HER2 (magenta) from mice at tumor endpoint. **N,** Average area of metastatic burden per lung for each genotype. Statistical significance was determined by Kruskal-Wallis test. P values are represented as follows:
*=p<0.05, **=p<0.01. Data are represented as Mean +/– SEM. Scale bar = 1mm.

tumors, there were two WT expressing tumors (cyan:TFP1), fifty-five d16 expressing tumors (yellow:EYFP), and thirty-three p95 expressing tumors (magenta:mKO). Using this registration, tumor free survival was plotted for each HER2 isoform. Twenty of the surveyed mice had at least one palpable d16 tumor with a median tumor free survival of 120 days compared to the 17 mice with at least one palpable p95 tumor and a median time of 144 days (Figure 1F). On average there were 2.2 d16 tumors, 1.32 p95 tumors, and 0.08 WT tumors per mouse (Figure 1G). HER2 Crainbow mice develop multifocal tumors, so we also reported on the lineage of first palpable tumors and compared this to the lineage of metachronous palpable tumors. First palpable tumors (primary) accounted for twenty-seven percent of total tumor burden, whereas the remaining seventy-three percent were metachronous tumors that developed at another site in the mouse. Similar frequencies for tumor penetrance for each genotype were observed for primary palpated tumors and metachronous tumors (Figure 1G).

Next, growth rates were calculated for d16 and p95 tumors. WT was not included due to the limited number of WT tumors. For each tumor palpated over time, measurements were plotted and registered to the genotype (Figure 1H). Tumor volume doubling time is a commonly used tool to asses tumor growth rate and has been correlated to tumor phenotype (Förnvik et al. 2016; Kay et al. 2019). TVDT was on average 12 days for both p95 HER2 and d16 HER2 tumors (Figure 1I). Assuming that the TVDT is constant, the age of tumor initiation can be calculated by using the first palpable measurement and the calculated TVDT. Inferred tumor initiation is calculated to be 93 days for d16 tumors and 78 days for p95 tumors (Figure 1J). However, in our system, the oncogene initiates no later than 14 days of age and tumors only became palpable between day 113 and 148. This confirms that even in mice most of the tumor growth phase is missed by palpation only and that TVDT is a poor predictor of tumor initiation. Whole mouse imaging also revealed a significant reservoir of lesions, undetectable by palpation, that were approximately 3-fold more abundant than palpable lesions and as small as 0.5mm in diameter (Figure 1K). Lastly, metastatic potential was measured by 3D whole lung imaging at tumor burden endpoint (n=12). Representative images show high p95 and d16 metastatic burden in the lungs of two different mice (Figure 1L). Of the 12 lungs, none had WT metastasis, eight had d16 metastases, and seven had p95 metastases (Figure 1M). Average area of metastatic lesions from whole lung images revealed that the metastatic burden for d16 and p95 were similar (Figure 1N).

Phenotype must therefore be established prior to the palpable phase of tumor growth. Quantitative imaging confirmed previous qualitative observations that each HER2 isoform possesses substantial differences in somatic clone growth and carrying capacity in the early epithelium of HBOW mice (Figure 1A, 2A) (Ginzel et al. 2021). Female HBOW mice were sacrificed at 3, 4, 6, and 10 weeks of age. Single cell suspensions from every gland (10 glands / mouse) were prepared and a representative sample of cells counted by confocal to estimate total cells by HER2 isoform. The total carrying capacity – that is the number of epithelial cells present within the duct – increased by almost three-fold from 3 weeks to adulthood as expected due to the pubertal development of the mammary duct (Figure 2B). Concomitant increases in WT and d16 cell numbers were observed through 6 weeks but p95 cell numbers remained stagnant and only started to variably increase at 10 weeks of age (Figure 2B,C). Throughout gland development, d16 HER2 expressing cells make up roughly half of the recombined mammary epithelium while WT HER2 increases from 30% to 44% during pubertal development

Figure 2

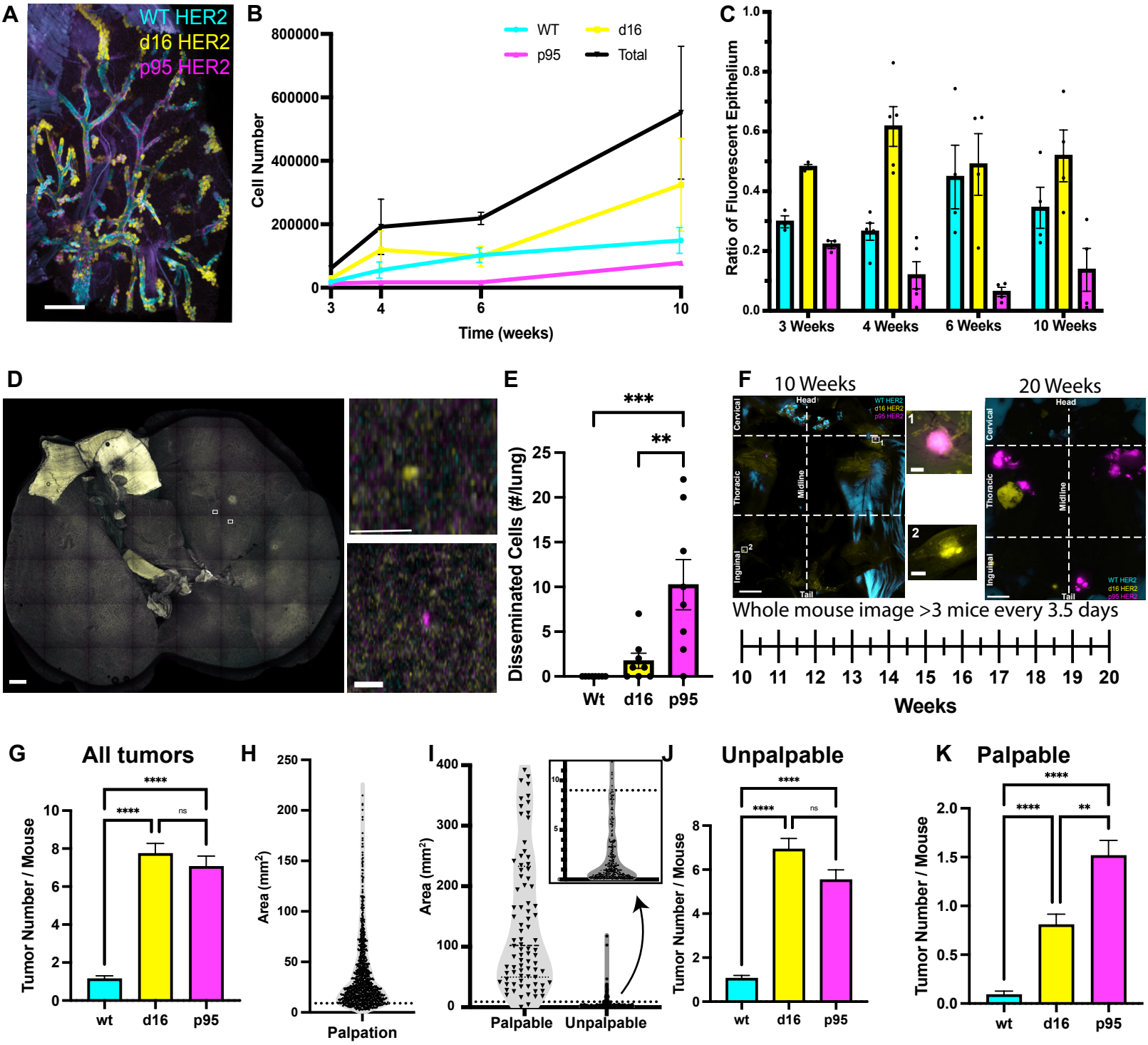

A
WT HER2
d16 HER2
p95 HER2
B
WT
d16
p95
Total
Cell Number
800000
600000
400000
200000
0
3
4
6
10
Time (weeks)
C
Ratio of Fluorescent Epithelium
1.0
0.8
0.6
0.4
0.2
0.0
3 Weeks
4 Weeks
6 Weeks
10 Weeks
D
E
Disseminated Cells (#/lung)
***
**
25
20
15
10
5
0
Wt
d16
p95
F
10 Weeks
20 Weeks
Head
Cervical
Thoracic
Midline
Inguinal
Tail
WT HER2
d16 HER2
p95 HER2
Whole mouse image >3 mice every 3.5 days
10 11 12 13 14 15 16 17 18 19 20
Weeks
G
All tumors
Tumor Number / Mouse
****
****
ns
wt
d16
p95
H
Area (mm²)
Palpation
I
Area (mm²)
Palpable
Unpalpable
J
Unpalpable
Tumor Number / Mouse
****
****
ns
wt
d16
p95
K
Palpable
Tumor Number / Mouse
****
****
**
wt
d16
p95

Figure 2. Quantifying the invisible phase of tumorigenesis. A, Whole mammary gland imaging of fluorescent Crainbow cells expressing WT (cyan), d16 (yellow), p95 (magenta) at 10 weeks of age. **B,** Quantification of the number of fluorescent cells from all ten mammary glands of HBOW mice after dissociation to single cells at 3, 4, 6, and 10 weeks of age (n=2-3 mice). **C,** Proportion of cell number for each genotype from **B.** **D,** Representative image from whole lung imaging of HBOW mice at 10-12 weeks of age. Inset: High magnification imaging shows single disseminated cells in the lung expressing p95 (magenta) or d16 (yellow). **E,** Quantification of the number of disseminated cells from whole lung imaging at 10-12 weeks of age (n=8) Statistical significance was determined by One-way ANOVA with Bonferroni's multiple comparison test. **F,** Whole mouse imaging of HBOW mice at 10 weeks of age and 20 weeks of age shows representative tumor number and size. To survey occult tumor growth, HBOW were sacrificed every 3.5 day from 10 weeks to 20 weeks of age to survey tumor size and genotype. **1&2** are high magnification insets from the 10 week old whole mouse image. **G,** Quantification by genotype of all tumors based on Crainbow fluorescence from whole mouse imaging (n= 120 mice). Statistical significance was determined by Kruskal-Wallis test. **H,** Calculated area of each palpation measurement taken in Figure 1C,D. Dotted line marks the 15th percentile at $9mm^2$. **I,** Tumor area from whole mouse imaging of palpated mice in Figure 1D,E. Dotted line marks the putative cutoff between palpable and unpalpable at $9mm^2$. Inset shows the number of unpalpable lesions that fall below the $9mm^2$ cutoff. **J,** Quantification by genotype of small, unpalpable tumors less than $9mm^2$ (**J)** and large, palpable tumors greater than $9mm^2$ (**K**) based on Crainbow fluorescence from whole mouse imaging (n= 120 mice). Statistical significance was determined by Kruskal-Wallis test. **E,G,J&K:** **=$p<0.01$, *** = $p<0.001$, **** = $p<0.0001$. Data are represented as Mean+/– SEM. Scale bar = 1mm for **A,D,F**; 0.4mm for **F insets**.

(Figure 2C). In contrast, p95 HER2 expressing cells fall from 22% to 7% of the recombined epithelium during puberty and variably increase to 14% on average at 10 weeks of age (Figure 2C).

In order to directly measure the invasive phenotype of early, occult tumorigenesis, whole lung imaging was performed to quantify the number of Crainbow cells that disseminated to the lung. Single disseminated cells could be found before palpable tumors could be observed (between 10 and 12 weeks of age) and p95 HER2 expressing cells were 5-fold more numerous in the lung than d16 HER2 (Figure 2D-E). This data demonstrates how field growth and size do not predict overall tumorigenic or metastatic potential.

Next, we sought to model the screen-detectable phase of tumorigenesis by examining the total number of small, unpalpable lesions over time that were only detectable by whole mouse imaging. HBOW mice were necropsied and imaged at time points every 3-4 days from 10 to 20 weeks of age to quantify unpalpable and image detectable lesions (Figure 2F). Representative images at 10 weeks show a few small lesions within the Crainbow labeled mammary ducts at a time when dissemination had already occurred(Figure 2F insets 1-2). These lesions continued to increase in number and size through 20 weeks (Figure 2F). Tumor burden for d16 HER2 and p95 HER2 was nearly identical, and each were seven-fold more likely than WT HER2 (Figure 2G). In order to differentiate between a palpable lesion and an unpalpable one from whole mouse imaging alone, volume measurements from the previous palpation experiments were converted to area and used to find the lower 15th percentile of palpable tumors ($9mm^2$)(Figure 2H). Applying this cutoff to whole mouse imaging from Figure 1K revealed that 98% of palpable tumors and 88% of unpalpable tumors were correctly classified (Figure 2I). After stratification of whole mouse imaging data using the $9mm^2$ cutoff, mice had on average 1 WT tumor, 7 d16 tumors and 5.5 p95 HER2 tumors that were smaller than $9mm^2$ (Small, Unpalpable:S) and 0.09 WT tumor, 0.8 d16 HER2 and 1.5 p95 HER2 tumors that were larger than $9mm^2$ (Large, Palpable:L) (Figure 2J,K). Despite having a significantly lower carrying capacity in the field, p95 exhibits an equivalent number of screen-detectable tumors. These data suggest that the critical turning point in the tumorigenic process occurs even before screen detection.

Tumors grow along a continuum from the earliest initiation event, to small, image detectable lesions, and finally to large, palpable tumors (Figure 3A). Based on the assumptions that all small tumors arise from a transformed cell in the field and that all large tumors were once small tumors, a series of ordinary differential equations (ODEs) are defined to capture the surveyed data from Figure 2 and estimate progression rates from one state to the next (Figure 3A,B). Due to the limited number of WT HER2 tumors we once again removed them from this analysis. Based on measurements from Figure 2B, a logistic growth model of the fields of HER2 expressing cells (*F*) was used to describe the exponential growth rate of HER2+ cells (*r)* that occurs during puberty from 4 to 6 weeks of age and then approaches a carrying capacity within the mammary duct (*K*) (Figure 3B,C). d16 HER2 transformed field cells were approximately 7-fold greater than p95 HER2 and their growth was closely associated with pubertal development of the gland (Figure 3C). The number of small and large tumors over time are plotted by fitting the ODEs to the whole mouse tumor measurements with 90% confidence intervals. (Figure 3D,E). d16 small tumors are more prevalent over time compared to p95 (Figure 3D). In contrast,

Figure 3

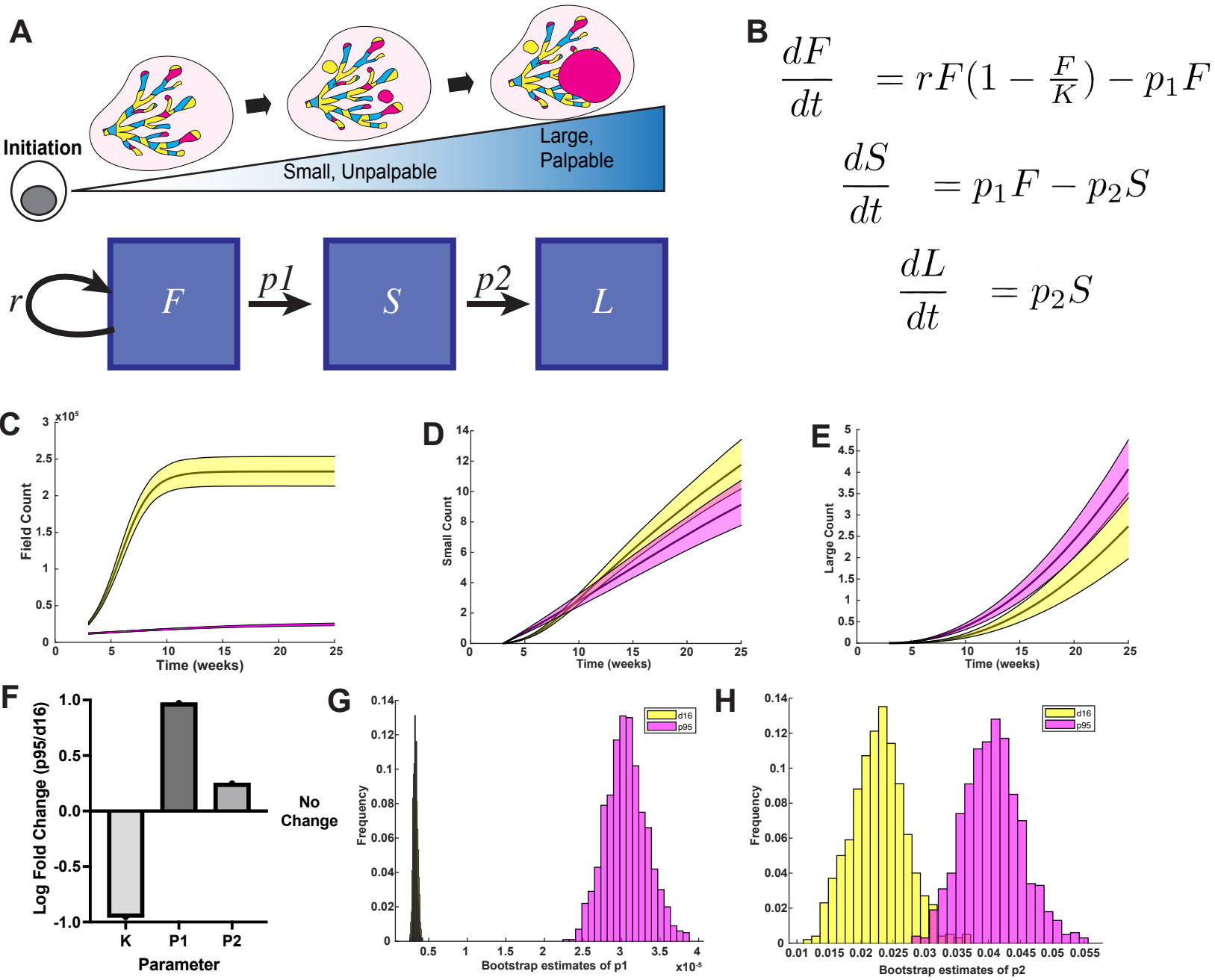


Figure 3. Three compartment model of tumorigenesis

A, Schematic of discrete states of tumorigenesis in HBOW mice. Fields of recombined cells (*F*) grow at a rate (*r*) before progressing to small, unpalpable lesions (*S*) at rate *p1*. Small lesions then progress to large, palpable tumors (*L*) at rate, *p2*. **B,** Ordinary differential equations defining progression rates from one compartment to the next. **C,** Estimated logistic growth of fields of cells expressing d16 (yellow) or p95 (magenta) in HBOW mice over time. **D,E** Logarithmic fit of the number of small (**D**) and large lesions (**E**) per mouse across time. **F,** Fold difference between p95 HER2 and d16 HER2 in field size, p1 rate, and p2 rate. **G-H,** Bootstrap estimates of p1 (**G**) and p2 (**H**). **C-E:** Error bars represent 90% confidence intervals.

p95 HER2 large tumor number remains higher than d16 HER2 throughout the observed time frame (Figure 3E).

Our model was then used to determine which state transition best explained the high p95 tumor burden despite the lack of a large field. Bootstrapping is a powerful statistical tool we used to make more robust statistical inferences from a relatively small biological sample. Bootstrap estimates of progression rate (*p1*) from transformed cell (*T*) to small tumor (*S*) is a log order higher for p95 compared to d16 HER2 (Figure 3F,G; Table1). Progression rate (*p2*) from small (*S*) to large tumors (*L*) is also two-fold higher than for p95 HER2 (Figure 3F,H; Table 1). Comparison of the progression rates between compartments demonstrates how each successive state transition becomes more likely. The model demonstrates that the transition from a transformed cell to a small tumor appears to be the earliest inflection point that begins to differentiate the phenotypic trajectory of p95 HER2 from d16 HER2.

Next, we reasoned that an occult state may also exist prior to image detection but after field growth. Two putative modalities of occult precancers in the mouse mammary gland have been established: hyperplastic alveolar nodules(HAN) and intraductal lesions (Medina 2000; Robert D. Cardiff and Wellings 1999; R. D. Cardiff, Wellings, and Faulkin 1977; Medina 1979). Confocal imaging of Crainbow fluorescence in FFPE sections from HBOW identifies extensive side-budding reminiscent of HAN that predominantly expresses d16 HER2 as well as lesions within the duct expressing both d16 and p95 HER2. (Figure 4A). Due to the rarity and labile nature of these lesions it is difficult to quantify and as such we define this state as the “Occult”. Using our mathematical model, a new compartment for occult dysplastic lesions (O) was incorporated to develop a four-state model of tumorigenesis (Figure 4B,C).

The four-state model revealed that transition out of the field compartment, *p0,* is roughly an order of magnitude higher for p95 than d16, and the transition into the palpable compartment, *p2,* is roughly twice as high for p95 compared to d16 (Figure 4D,E,G; Table 2). Not present in the FSL model, the transition rate from an occult dysplasia to a small neoplasia, *p1,* is about two-fold greater for d16 than p95, though this is not significantly different (Figure 4D,F; Table 2). We used this model to estimate the number of occult (“O”) lesions over time. Estimates of the number of occult lesions are not predictive but suggest that p95 more frequently progresses from the field (F) and generates more occult lesions that readily advance to screen-detectable lesions. This occurs despite a relatively small reservoir of field cells (Figure 4H). Despite the uncertainty in the rate *p1* and the state O, inclusion of an occult compartment highlights that the transition from a field cell to a detectable tumor is not a linear process and each genotype encounters a critical rate limiting step in the tumorigenic process that occurs long before becoming detectable.

The cutoff between the S and L compartment was based on palpation data. In order to verify that this cutoff did not significantly alter the robustness of the model and to correlate this data to our previously published description of MIN, we utilized a 1.5mm^2 cutoff to stratify the S and L compartments (Ginzel et al. 2021). When applied to our four-compartment model, the log fold change in parameter estimates is within 10% of the 9mm^2 cutoff model (Supplementary Figure 1A). The bootstrap estimate for p0 and p1 were very similar between the two cutoffs and while the ratio of p95 to d16 was similar the p2 estimates were ~3 fold larger in the 1.5 cutoff compared to the 9 cutoff (Supp. Figure 1B-D, Table 2). This is reflective of the significantly larger pool of

Figure 4

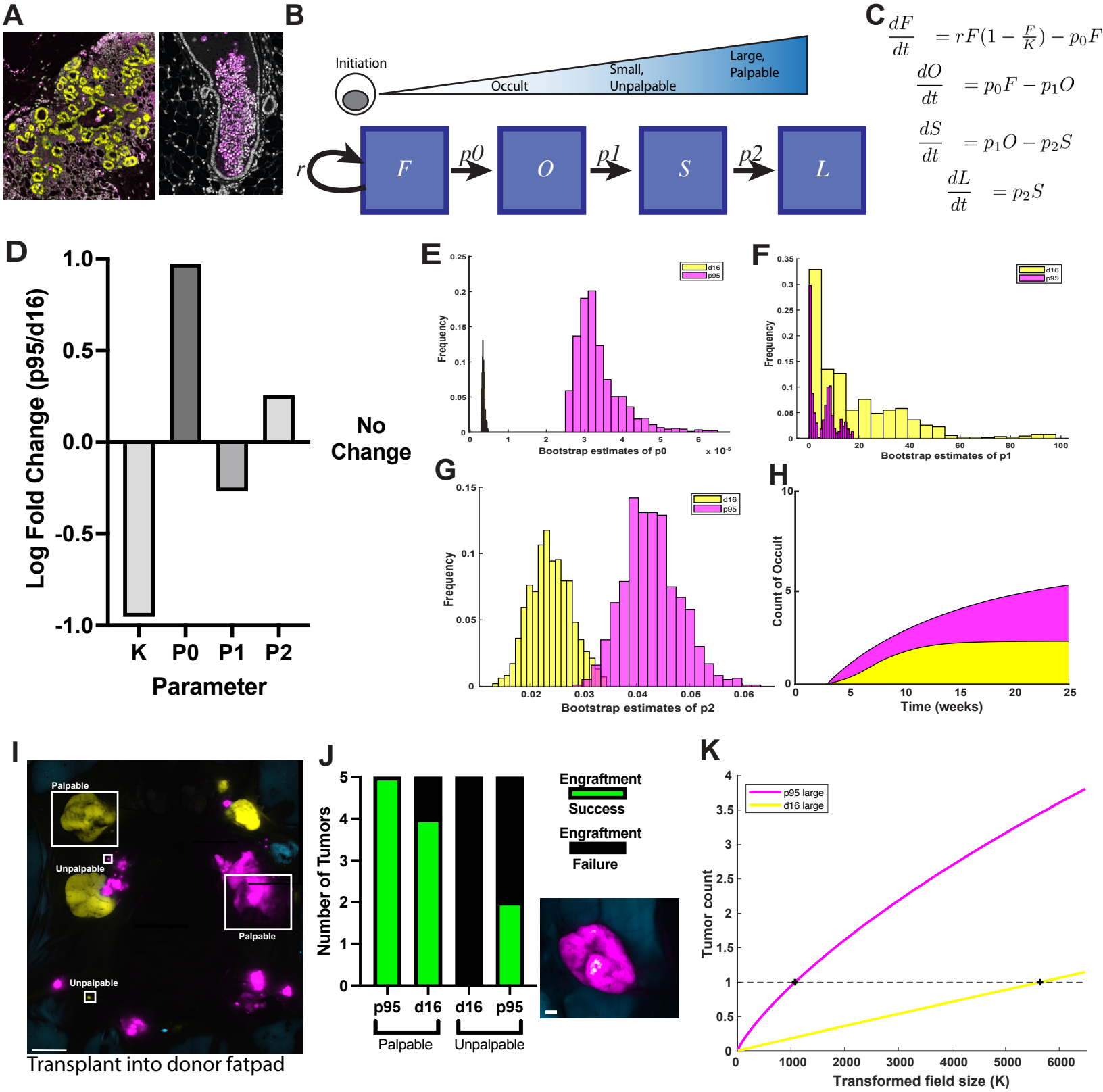

A
B
Initiation
Occult
Small, Unpalpable
Large, Palpable
r
F
p0
O
p1
S
p2
L
C
$\frac{dF}{dt} = rF(1-\frac{F}{K}) - p_0F$
$\frac{dO}{dt} = p_0F - p_1O$
$\frac{dS}{dt} = p_1O - p_2S$
$\frac{dL}{dt} = p_2S$
D
Log Fold Change (p95/d16)
1.0
0.5
0.0
-0.5
-1.0
K
P0
P1
P2
Parameter
No Change
E
Frequency
Bootstrap estimates of p0
x 10^-5
d16
p95
F
Frequency
Bootstrap estimates of p1
d16
p95
G
Frequency
Bootstrap estimates of p2
d16
p95
H
Count of Occult
Time (weeks)
I
Palpable
Unpalpable
Palpable
Unpalpable
Transplant into donor fatpad
J
Number of Tumors
p95
d16
d16
p95
Palpable
Unpalpable
Engraftment Success
Engraftment Failure
K
p95 large
d16 large
Tumor count
Transformed field size (K)

Figure 4. Four compartment model of occult tumorigenesis. A, Confocal imaging of potential occult lesions in HBOW mammary glands expressing d16 (yellow) or p95 (magenta). **B,** Schematic of discrete states of tumorigenesis in HBOW mice. Fields of recombined cells (*F)* grow at a rate (*r)* before progressing to occult dysplasia (O) at rate (*p0*). Dysplastic lesions progress to small, unpalpable tumors (*S*) at rate *p1*. Small lesions then progress to large, palpable tumors (*L*) at rate, *p2*. **C,** Ordinary differential equations defining progression rates from one compartment to the next. **D,** Fold difference between p95 HER2 and d16 HER2 in field size, p0 rate, p1 rate, and p2 rate. **E-G,** Bootstrap estimates of p0 (**E**), p1 (**F**) and p2 (**G**). **H,** Estimates of the number of occult dysplastic lesions for d16 HER2 and p95 HER2. Error bars represent 80% confidence intervals. **I,** Whole mouse image with representative palpable and unpalpable tumors in boxes. Tumors were removed for digestion and transplantation into donor fatpads. **J,** Engraftment rate for tumor cells dissociated from unpalpable and palpable tumors and transplanted into donor mammary fat pads. Green denotes a successful engraftment and black a transplant that failed to grow (n=5). Lower right shows a representative outgrowth from a palpable p95 tumor transplant (magenta). **K,** Simulated data using the FOSL model to estimate the number of cells expressing either d16 (yellow) or p95 (magenta) needed to make a tumor within the lifetime of a mouse. **I&J,** Scale bar = 3mm

small lesions in the 1.5mm^2 cutoff. Overall, this affirms that our image size cutoffs are reasonable.

Next, we determined if the differences observed in the occult transition could, in part, account for the differences in overall tumorigenic potential for each genotype. We reasoned that if all p95 tumors are invasive out of the occult state, then the smallest screen detectable lesions should be transplantable. Whole mouse imaging was performed to identify paired palpable and unpalpable tumors of the same genotype from the same mouse (N=5 per genotype). These tumors were then dissected and transplanted into the orthotopic fatpad of immune intact and syngeneic FVB/N female mice between 4 and 8 weeks of age (Figure 4I). Successful engraftment was determined by a tumor growing to humane endpoint and maintaining HER2 expression as confirmed by whole mouse imaging. Evaluation of engraftment rate revealed that all palpable p95 tumors successfully engrafted while 80% of d16 tumors transplanted successfully (Figure 4J). 40% of p95 unpalpable lesions transplanted successfully while none of d16 unpalpable lesions could be transplanted (Figure 4J). This suggests that the lethality of a tumor clone is established prior to detection – what we call here “nascent lethality”.

Lastly, sensitivity analysis was used to understand how variations in parameters (i.e. r, K, p1, p2) affect the uncertainty in the states (F, O, S, and L) (Dickinson and Gelinas 1976; Morris 1991). In this analysis, we see the most sensitive parameter for a large, detectable tumor (*L*) is *p0*, the transition rate from the field to an occult dysplastic lesion, for both the p95 and d16 models (Supplementary Figure 2A,B). An interesting difference between the genotypes is that for d16, the field carrying capacity *K* has effects of the same magnitude of *p0* (Supplementary Figure 2A,B) whereas for p95, K is one of the least important parameters driving progression to large tumors. We can visualize the importance of the field size, or carrying capacity (*K*), by simulating the number of large tumors that would form from varied field sizes within an average lifetime of a captive mouse (2 years). A field of p95 HER2 cells will make a detectable tumor from a relatively small field of cells (~1,000) compared to d16 HER2 which requires nearly 6 times as large of a field to generate one tumor (Figure 4K). Thus, the progression of invasive large tumors appears to be uncoupled from field growth.

Discussion

Three main phases of cancer intervention have recently been summarized and include: prevention, interception, and mitigation (Castle et al. 2024; Domchek and Vonderheide 2024). The conventional paradigm of lethal cancer progression is based upon the sequential accumulation of genetic changes, sufficient to alter normal cellular behavior to a malignant phenotype of invasion, dissemination, and metastasis (Fearon and Vogelstein 1990; Armitage and Doll 1954; Nordling 1953). This paradigm supports the current cancer control strategy of mitigation using surveillance to detect anatomically small tumors with the expectation that clinically detecting and treating these smaller tumors results in decreased mortality. Unfortunately, even after decades of screening for clinically detectable tumors, this strategy has not resulted in the expected positive benefits of reducing mortality or late-stage disease (H. G. Welch, Kramer, and Black 2019).

The benefits and risks of screening are under increasing scrutiny, and in the breast, there is evidence to suggest net neutrality in survival despite decades of intense screening (H. Gilbert Welch and Dey 2023). Our data demonstrates the paradox of cancer screening using a

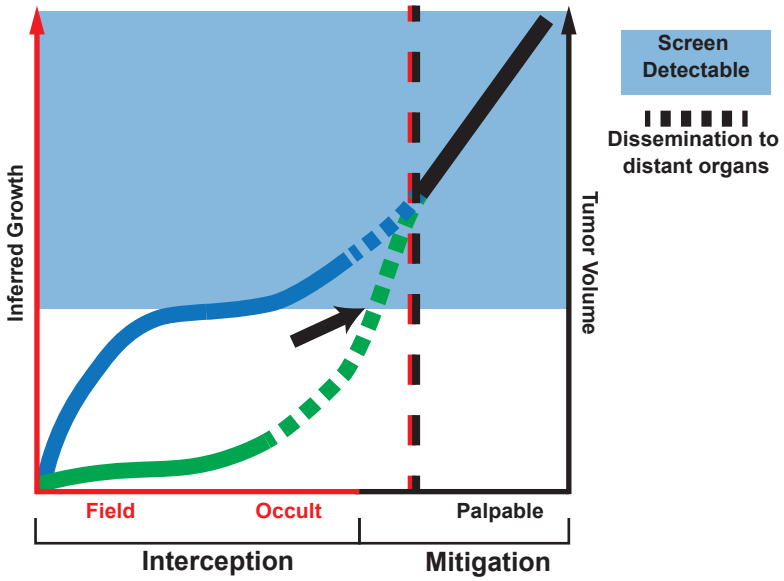


Figure 5. Graphical representation of 2 cases of non-linear tumor progression. Both cases occur prior to the continuous, clinically observable growth of large tumors. The blue line represents a rapidly growing field and its progression to a small screen-detectable tumor. In this case, cancer mitigation by screen detection (blue shade) can occur before dissemination to distant organs (dotted line) and reduce the risk of metastasis. In contrast, the green line represents the slow growing field that rapidly progresses to an invasive lesion that disseminates (dotted line) before becoming screen detectable (Black arrow). This nascent lethal phenotype reduces the efficacy of Mitigation. Thus, cancer interception strategies must be pursued to prevent nascent lethal progression.

genetically tractable mouse model system (Figure 5). We show that tumors can originate from rapidly growing premalignant fields (i.e. d16-HER2). Larger field sizes increase the likelihood that a rare clone advances to a small image detectable event in the mouse that is unlikely to disseminate but if left untreated may eventually metastasize. Therefore, early screening and appropriate interventions could reduce the likelihood of a malignant transition and prevent recurrent, late-stage disease. On the other hand, our modeling also shows how quiescent clones (i.e. p95-HER2) are unable to expand. Yet, their ability to transition from premalignancy to malignancy is high. These nascent lethal clones have already disseminated and are much more likely to escape, prior to anatomic disruption sufficient for clinical detection.

We also provide experimental evidence for a model of “Biological predeterminism” (Ian Macdonald 1966; I. Macdonald 1951), whereby the lengthy preclinical process establishes lethality prior to detection. Curiously, the most detectable lesions in our model are associated with an isoform of HER2 that is also predictive of favorable therapeutic responses (Castagnoli et al. 2014). Thus, d16 driven tumors are likely to have favorable outcomes regardless of early detection. In contrast, a tumor lacking a treatable domain, like p95 HER2 (Abraham et al. 2019; Scaltriti et al. 2007; Pedersen et al. 2009), is also the most unlikely to be detected, and most likely to disseminate. Thus, in our model, p95 represents the worst-case scenario for a patient – a treatment insensitive tumor that is not screen detectable. This continues to support the notion that p95 is associated with worse clinical outcomes (Abraham et al. 2019; Carvajal-Hausdorf et al. 2015; Scaltriti et al. 2007) and further warrants the need to clinically validate this isoform as a predictive biomarker.

Our model points to an important occult transition where a premalignant clone becomes a malignant clone capable of invasion and growth at a distant site. We chose the term occult because of the inability to image the exact moment of this transition. This leads to a challenging question, what is the occult transition? The occult transition could represent a location in the form of a privileged site of growth. We speculate that an occult compartment could reside inside the duct, similar to the ductal hyperplasias described several decades ago (Medina 2000), in which growth inside the duct provides for a unique microenvironment compared to growth into the surrounding fat. For instance, mammary ducts are surrounded by a layer of ductal macrophages (Dawson et al. 2020). This could provide a mechanism whereby disseminating cells could be in close proximity to macrophage populations necessary for early invasion (Linde et al. 2018; Sidani et al. 2006). The occult transition could also represent a moment in time where malignancy is established by changes in the host that impact fitness of tumor cell phenotypes. These could be systemic in nature and may range from immune ageing, microbiome changes, obesity, injury and many others (Swanton et al. 2024). The occult transition could also occur through oncogene driven differences in cell state plasticity. We also have evidence for early state changes in the p95 expressing mammary epithelium that results in EMT-like states and an early invasive histological phenotype (Ginzel et al. 2021) This suggests that functionally selective signaling across HER2 isoforms may also be important (Dankort et al. 2001).

There are limitations and assumptions to our models that must also be considered. First, our system is an abstraction of a much more complicated tumor ecosystem. Our mouse model overexpresses human HER2 isoforms. Although this is a useful model system of HER2 amplification, it is still not naturalistic. Similarly, the system of ODEs is a mathematical

abstraction of the complex tumor ecosystem and has several key limitations, some of which would be common to any similar mathematical model. Primarily, spatial relationships cannot be described by the current ODEs. Second, the parameters of the dynamical system are estimated using survey data (mouse necropsy) instead of longitudinal data. Thus, the use of certain statistical frameworks, e.g. nonlinear mixed-effects modeling, that describe the variation between individuals are not applicable (Pinheiro and Bates 1995; Comets, Lavenu, and Lavielle 2017). Lastly, the models describe observations of 2-dimensional tumor size (mm^2), not the biological state (e.g. metastatic or benign) or interactions between tumor genotypes.

Overall, we demonstrate the power of genetically tractable model systems for investigating the premalignant phase of tumorigenesis in immune intact animals. Our data provide evidence for an occult transition as a major inflection point in the natural history of cancer that shapes the overall trajectory of the disease. Intercepting the occult transition may provide a new strategy for transforming cancer intervention.

Materials and Methods

Animal Studies

All animal studies were performed according to an approved protocol from the Duke Institutional Animal Care and Use Committee. HER2 Crainbow mice were crossed to MMTV Cre (MMTV-cre/Line D (MGI, catalog no. 3581599, RRID:MGI:3581599) to generate HBOW (Ginzel et al. 2021). Both lines were fully backcrossed to the FVB/N background. All experimental animals were pre-emptively enucleated at 3 weeks of age due to off target harderian gland tumors (Garner et al. 2022). All experimental animals used were female.

Mammary Gland Epithelial Cell Analysis

Digestion of mammary glands for analysis of epithelial cell number was performed following the protocol from Ludwik et al, 2021(Ludwik et al. 2021). Briefly, all 10 mammary fat pads were removed from M-HER2 mice, finely minced, and incubated in Collagenase A(2mg/mL) in RPMI media for 3 hours at 37C with 5% CO2. Next, digested tissue was incubated in DNase I followed by RBC lysis. Then digested tissue was spun down and resuspended in Accumax for 10 minutes on a rotating shaker at 37C. Next, cells were pelleted and resuspended in 5X Trypsin for 5 minutes at 37C before being filtered through a 40um filter. Samples of single cell suspensions were imaged on a Zeiss 880 confocal and counted using Imaris(v10.1).

Whole Mouse Imaging

Mice were euthanized with CO2 and pinned on vinyl dissecting pads to expose all 10 mammary glands. Each mouse was then imaged on a Zeiss Axiozoom. Images were imported into either QuPath(Bankhead et al. 2017) or ImageJ2 (Rueden et al. 2017) for manual measurement and classification of all tumors.

Whole Organ Imaging

Mammary glands and lungs were dissected and placed in or inflated with 10% neutral buffered formalin (NBF) for fixation. After 2 hours of fixation, tissues were briefly washed with 0.5% Tween with PBS for 30 minutes before being submerged in FUNGI clearing solution for at least 4 hours (Rios et al. 2019). After clearing, tissues were sandwiched between a glass slide and coverslip for imaging on a Zeiss 880 confocal. For whole lung imaging, spectra were obtained from lambda mode imaging for each Crainbow fluorescent protein and stored as “online fingerprints” for imaging. This process ensures only Crainbow fluorescence was captured and reduces background fluorescence. Imaging of occult lesions was done on FFPE sections of mammary glands fixed in NBF. FFPE sections were deparaffinized and mounted with DAPI nuclear stain before imaging on a Zeiss 880 confocal microscope. Crainbow fluorescence is imageable following FFPE processing with no further amplification or staining (Ginzel et al. 2021). Analysis was performed using ImageJ2 (Rueden et al. 2017) and Imaris (v10.1).

Mathematical Modeling

<u>Sensitivity Analysis</u> Sensitivity analysis is an analytical technique that explains the uncertainty in the model output as a function of the uncertainty in the model parameters. Sensitivity analysis provides insight into the relationship between the parameters (*r, K, p1, p2*) and the values of the

states (F, S, and L). The more sensitive a parameter is, the more effect it has on the state value. If a parameter is *insensitive,* large changes in its value result in small changes to the state values. Here the Morris method is employed to screen the parameters of the system, using the SAFE software package in MATLAB (v9.14; 2023a) (Pianosi, Sarrazin, and Wagener 2015). Morris screening varies one parameter at a time while the others remain constant, and an *elementary effect* is calculated. The elementary effect is the absolute difference in the change in state, divided by the change in the parameter value. The elementary effect is computed at each point in time, and the result is plotted (Supplementary Figure 2).

Parameter Estimation Parameters for the logistic equation describing the field (r, K, and the initial conditions F(0)) are estimated first using mammary gland digestion data (Figure 2B,3C). Ordinary least squares (OLS) regression is used to fit the solution of the logistic equation to the week 3, 4 and 6 data. The MATLAB function lsqcurvefit is used to find the values *r, K,* and *F*(0) that minimize the least-squared error between the data points and the solution of the logistic equation. For each of the estimated parameters, we must set a range of allowable values, and the ranges are chosen to be roughly the same as the observed data. The time coordinates are shifted so that the initial condition *F*(0) corresponds to an age of 3 weeks, the time of the earliest digestion studies. We assume that growth of the mammary duct is essentially complete by 6 weeks of age, so we choose the upper and lower bounds of *K* according to the range of data at 6 weeks. After deriving estimates of the logistic growth parameters of the field, these values are set and parameter estimates of the transition rates in the S and L compartments, p1 and p2, are calculated (and p0 in the FOSL model). The state transition rates are also estimated by OLS using the MATLAB function lsqcurvefit. Here, we solve the ODEs (Fig 3B,4C) using MATLAB's ode15s function and then minimize the sum-of-squared differences with the mouse survey data. Every mouse produces 4 data points: (S,L) for both p95 and d16, and the parameters for the p95 and d16 models are fit independently. To produce confidence intervals for the estimations, we employ a hierarchical bootstrap method. There are 120 mice in the survey data set, and we sample N = 120 mice with replacement. Then from each sampled mouse, we re-sample the tumors in that mouse, recording the area of each resampled tumor. The re-sampled tumors of this synthetic mouse are then stratified by genotype and state (S/L) so that each sampled mouse produces 4 data points. The hierarchical bootstrap is repeated 1000 times. Parameters are estimated using median values, and confidence intervals are given by percentiles of the bootstrapped distribution.

To account for additional uncertainty, the field parameters *r,* K, and F(0) are resampled at each bootstrap iteration, using a normal distribution with mean given by the preliminary OLS estimate and standard deviation equal to 5% of the mean. Uncertainty in the mouse age is modeled by adding a normally distributed random number with mean 0 and std. 0.05 (weeks) to each synthetic mouse's age at every bootstrap iteration.

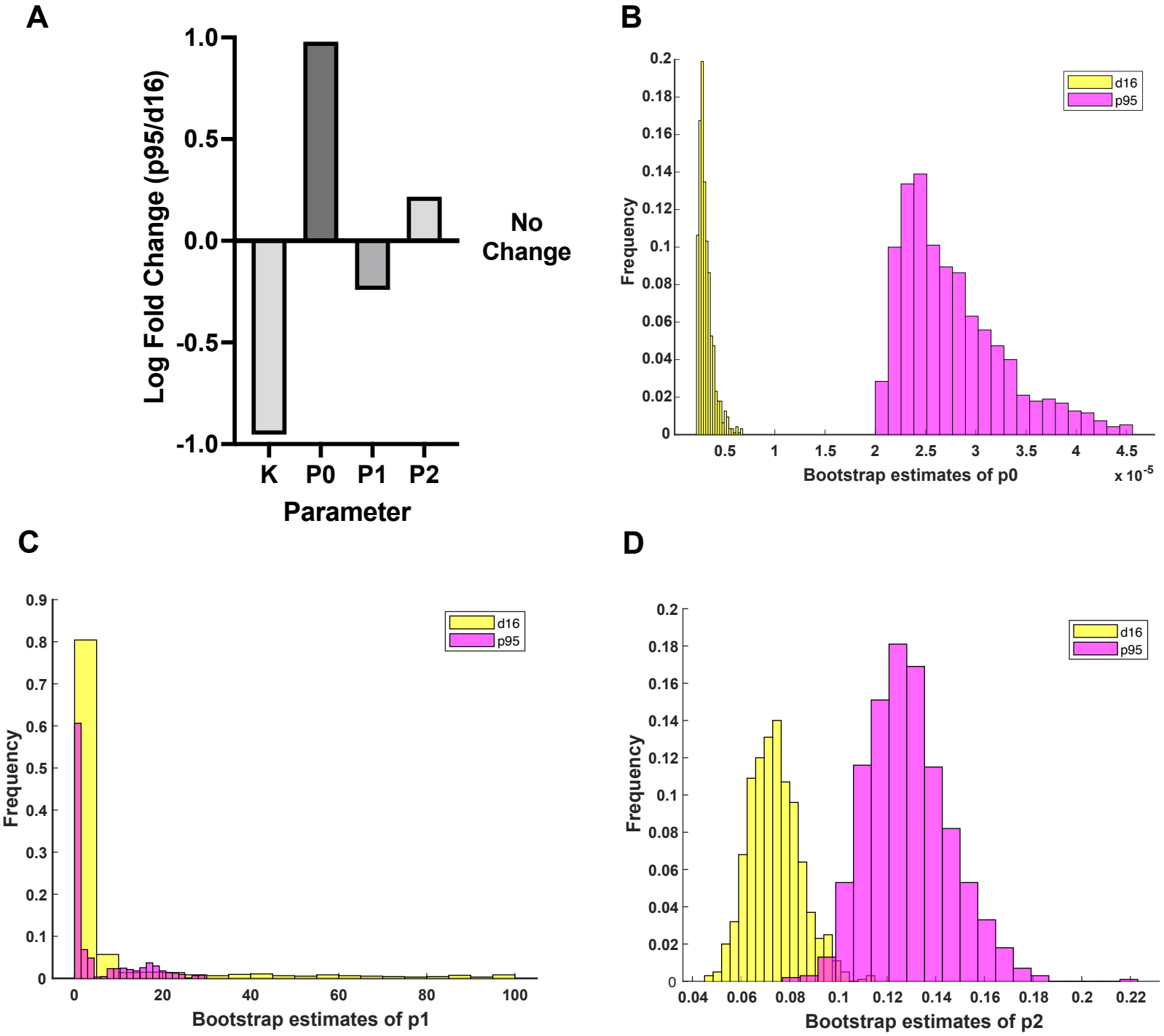
Transplant Studies

Tumors to be used for transplant were selected from Axiozoom imaging and digested following the same digestion protocol as described in **Mammary Gland Epithelial Cell Analysis**. Single cells were resuspended in USP saline. Female FVBN/J mice (Jackson Labs) between 4 and 8 weeks of age were anesthetized using Isoflurane and 500,000 isolated tumor cells were injected

into the right, fourth(inguinal) mammary gland. Tumors were measured once weekly. At tumor burden endpoint, **Whole Mouse Imaging** was performed to assess genotype.

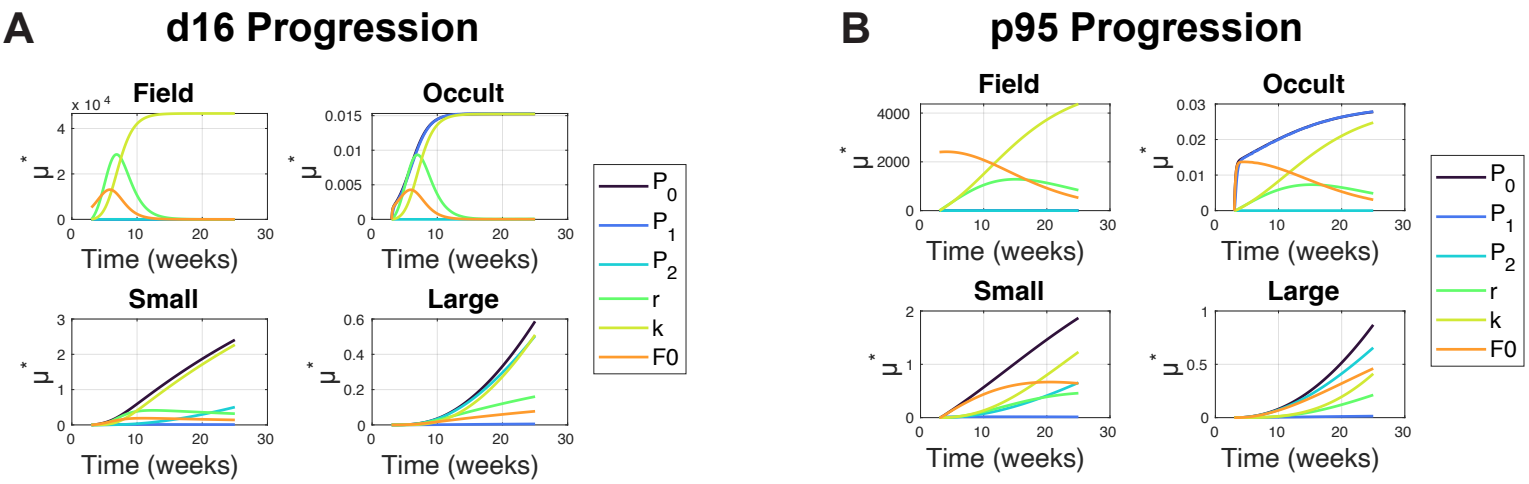
Acknowledgements
Funding for this work was provided by NCI 1R01CA255372-01 to JCS and DoD W81XWH-21-2-0031to HKL.

Supplementary Figure 1



Supplementary Figure 1. Effect of S:L cutoff in four compartment model of tumorigenesis. A, Fold change difference in parameter estimates between p95 and d16 when the cutoff from S:L is 1.5mm^2. **B-D,** Bootstrap estimates of *p0*(**B**), *p1*(**C**), and *p2*(**D**).

Supplementary Figure 2



Supplementary Figure 2. Sensitivity analysis of the four compartment model. A-B, Sensitivity plots for each parameter of the four compartment model for d16(**A**) and p95(**B**).

Table 1. Three compartment model parameters with 95% confidence interval

	d16	CI 95%	p95	CI 95%
F0	26000	N/A	12000	N/A
K	233000	N/A	26000	N/A
r	0.7313	N/A	0.1346	N/A
p1*	0.000003259	2.786e-6, 3.846e-6	0.0003048	2.544e-5, 3.599e-5
p2*	0.0229	0.0148, 0.0315	0.0405	0.0321, 0.0502

* denotes significance

Table 2. Four compartment model parameters with 95% confidence intervals

		d16	95% CI	p95	95% CI
	F0	26000	N/A	12000	N/A
	K	233000	N/A	26000	N/A
	r	0.7313	N/A	0.1346	N/A
C=9mm	p0*	3.40E-06	2.870e-6, 4.339e-6	3.20E-05	2.652e-5, 4.952-5
	p1	10.4332	0.4220, 65.5811	5.6362	0.2117, 15.5382
	p2*	0.0237	0.0178, 0.0310	0.0426	0.0348, 0.0522
C=1.5mm	p0*	3.38E-06	2.861e-6, 6.940e-6	3.21E-05	2.632e-5, 5.079e-5
	p1	8.2566	0.1101, 99.8215	4.7475	0.2064, 17.1339
	p2*	0.0764	0.0614, 0.0981	0.1258	0.1051, 0.1586

* denotes significance

References

Abraham, J., A. J. Montero, R. C. Jankowitz, M. A. Salkeni, J. H. Beumer, B. F. Kiesel, F. Piette, et al. 2019. "Safety and Efficacy of T-DM1 Plus Neratinib in Patients With Metastatic HER2-Positive Breast Cancer: NSABP Foundation Trial FB-10." *J Clin Oncol* 37 (29): 2601–9. https://doi.org/10.1200/JCO.19.00858.

Armitage, P., and R. Doll. 1954. "The Age Distribution of Cancer and a Multi-Stage Theory of Carcinogenesis." *Br J Cancer* 8 (1): 1–12. https://doi.org/10.1038/bjc.1954.1.

Bankhead, Peter, Maurice B. Loughrey, José A. Fernández, Yvonne Dombrowski, Darragh G. McArt, Philip D. Dunne, Stephen McQuaid, et al. 2017. "QuPath: Open Source Software for Digital Pathology Image Analysis." *Scientific Reports* 7 (1): 16878. https://doi.org/10.1038/s41598-017-17204-5.

Boone, P. G., L. K. Rochelle, J. D. Ginzel, V. Lubkov, W. L. Roberts, P. J. Nicholls, C. Bock, et al. 2019. "A Cancer Rainbow Mouse for Visualizing the Functional Genomics of Oncogenic Clonal Expansion." *Nat Commun* 10 (1): 5490. https://doi.org/10.1038/s41467-019-13330-y.

Brown, S., C. M. Pineda, T. Xin, J. Boucher, K. C. Suozzi, S. Park, C. Matte-Martone, et al. 2017. "Correction of Aberrant Growth Preserves Tissue Homeostasis." *Nature* 548 (7667): 334–37. https://doi.org/10.1038/nature23304.

Cardiff, R. D., S. R. Wellings, and L. J. Faulkin. 1977. "Biology of Breast Preneoplasia." *Cancer* 39 (6): 2734–46. https://doi.org/10.1002/1097-0142(197706)39:6<2734::AID-CNCR2820390661>3.0.CO;2-U.

Cardiff, Robert D., and Sefton R. Wellings. 1999. "The Comparative Pathology of Human and Mouse Mammary Glands." *Journal of Mammary Gland Biology and Neoplasia* 4 (1): 105–22. https://doi.org/10.1023/a:1018712905244.

Carvajal-Hausdorf, D. E., K. A. Schalper, L. Pusztai, A. Psyrri, K. T. Kalogeras, V. Kotoula, G. Fountzilas, and D. L. Rimm. 2015. "Measurement of Domain-Specific HER2 (ERBB2) Expression May Classify Benefit From Trastuzumab in Breast Cancer." *J Natl Cancer Inst* 107 (8). https://doi.org/10.1093/jnci/djv136.

Casasent, A. K., A. Schalck, R. Gao, E. Sei, A. Long, W. Pangburn, T. Casasent, F. Meric-Bernstam, M. E. Edgerton, and N. E. Navin. 2018. "Multiclonal Invasion in Breast Tumors Identified by Topographic Single Cell Sequencing." *Cell* 172 (1–2): 205-217 e12. https://doi.org/10.1016/j.cell.2017.12.007.

Castagnoli, L., M. Iezzi, G. C. Ghedini, V. Ciravolo, G. Marzano, A. Lamolinara, R. Zappasodi, et al. 2014. "Activated d16HER2 Homodimers and SRC Kinase Mediate Optimal Efficacy for Trastuzumab." *Cancer Res* 74 (21): 6248–59. https://doi.org/10.1158/0008-5472.CAN-14-0983.

Castle, Philip E., Jessica M. Faupel-Badger, Asad Umar, and Timothy R. Rebbeck. 2024. "A Proposed Framework and Lexicon for Cancer Prevention." *Cancer Discovery* 14 (4): 594–99. https://doi.org/10.1158/2159-8290.CD-23-1492.

Comets, Emmanuelle, Audrey Lavenu, and Marc Lavielle. 2017. "Parameter Estimation in Nonlinear Mixed Effect Models Using Saemix, an R Implementation of the SAEM Algorithm." *Journal of Statistical Software* 80 (August): 1–41. https://doi.org/10.18637/jss.v080.i03.

Dankort, D., B. Maslikowski, N. Warner, N. Kanno, H. Kim, Z. Wang, M. F. Moran, R. G. Oshima, R. D. Cardiff, and W. J. Muller. 2001. "Grb2 and Shc Adapter Proteins Play Distinct Roles in Neu (ErbB-2)-Induced Mammary Tumorigenesis: Implications for Human Breast Cancer." *Mol Cell Biol* 21 (5): 1540–51. https://doi.org/10.1128/MCB.21.5.1540-1551.2001.

Dawson, C. A., B. Pal, F. Vaillant, L. C. Gandolfo, Z. Liu, C. Bleriot, F. Ginhoux, et al. 2020. "Tissue-Resident Ductal Macrophages Survey the Mammary Epithelium and Facilitate Tissue Remodelling." *Nat Cell Biol* 22 (5): 546–58. https://doi.org/10.1038/s41556-020-0505-0.

Dickinson, Robert P., and Robert J. Gelinas. 1976. "Sensitivity Analysis of Ordinary Differential Equation Systems—A Direct Method." *Journal of Computational Physics* 21 (2): 123–43. https://doi.org/10.1016/0021-9991(76)90007-3.

Domchek, Susan M., and Robert H. Vonderheide. 2024. "Advancing Cancer Interception." *Cancer Discovery* 14 (4): 600–604. https://doi.org/10.1158/2159-8290.CD-24-0015.

Evans, E. J., and J. DeGregori. 2021. "Cells with Cancer-Associated Mutations Overtake Our Tissues as We Age." *Aging Cancer* 2 (3): 82–97. https://doi.org/10.1002/aac2.12037.

Fearon, E. R., and B. Vogelstein. 1990. "A Genetic Model for Colorectal Tumorigenesis." *Cell* 61 (5): 759–67. https://doi.org/10.1016/0092-8674(90)90186-i.

Förnvik, D., K. Lång, I. Andersson, M. Dustler, S. Borgquist, and P. Timberg. 2016. "ESTIMATES OF BREAST CANCER GROWTH RATE FROM MAMMOGRAMS AND ITS RELATION TO TUMOUR CHARACTERISTICS." *Radiation Protection Dosimetry* 169 (1–4): 151–57. https://doi.org/10.1093/rpd/ncv417.

Garner, Angela, Joshua D Ginzel, Joshua C Snyder, Jeffrey I Everitt, and Chelsea D Landon. 2022. "Veterinary Management of Harderian Gland Tumors in Cancer Rainbow (Crainbow) HER2-Positive Mice." *Comparative Medicine* 72 (6): 403–9. https://doi.org/10.30802/AALAS-CM-22-000061.

Gerstung, M., C. Jolly, I. Leshchiner, S. C. Dentro, S. Gonzalez, D. Rosebrock, T. J. Mitchell, et al. 2020. "The Evolutionary History of 2,658 Cancers." *Nature* 578 (7793): 122–28. https://doi.org/10.1038/s41586-019-1907-7.

Ginzel, J. D., C. R. Acharya, V. Lubkov, H. Mori, P. G. Boone, L. K. Rochelle, W. L. Roberts, et al. 2021. "HER2 Isoforms Uniquely Program Intratumor Heterogeneity and Predetermine Breast Cancer Trajectories During the Occult Tumorigenic Phase." *Mol Cancer Res*, June. https://doi.org/10.1158/1541-7786.MCR-21-0215.

Harper, K. L., M. S. Sosa, D. Entenberg, H. Hosseini, J. F. Cheung, R. Nobre, A. Avivar-Valderas, et al. 2016. "Mechanism of Early Dissemination and Metastasis in Her2(+) Mammary Cancer." *Nature* 540 (7634): 588–92. https://doi.org/10.1038/nature20609.

Hill, William, Emilia L. Lim, Clare E. Weeden, Claudia Lee, Marcellus Augustine, Kezhong Chen, Feng-Che Kuan, et al. 2023. "Lung Adenocarcinoma Promotion by Air Pollutants." *Nature* 616 (7955): 159–67. https://doi.org/10.1038/s41586-023-05874-3.

Hu, Z., Z. Li, Z. Ma, and C. Curtis. 2020. "Multi-Cancer Analysis of Clonality and the Timing of Systemic Spread in Paired Primary Tumors and Metastases." *Nat Genet* 52 (7): 701–8. https://doi.org/10.1038/s41588-020-0628-z.

Huels, D. J., L. Bruens, M. C. Hodder, P. Cammareri, A. D. Campbell, R. A. Ridgway, D. M. Gay, et al. 2018. "Wnt Ligands Influence Tumour Initiation by Controlling the Number of Intestinal Stem Cells." *Nat Commun* 9 (1): 1132. https://doi.org/10.1038/s41467-018-03426-2.

Kay, Katherine, Keith Dolcy, Robert Bies, and Dhaval K. Shah. 2019. “Estimation of Solid Tumor Doubling Times from Progression-Free Survival Plots Using a Novel Statistical Approach.” *The AAPS Journal* 21 (2): 27. https://doi.org/10.1208/s12248-019-0302-5.

Klein, Christoph A. 2009. “Parallel Progression of Primary Tumours and Metastases.” *Nature Reviews Cancer* 9 (4): 302–12. https://doi.org/10.1038/nrc2627.

———. 2020. “Cancer Progression and the Invisible Phase of Metastatic Colonization.” *Nature Reviews. Cancer* 20 (11): 681–94. https://doi.org/10.1038/s41568-020-00300-6.

Linde, Nina, Maria Casanova-Acebes, Maria Soledad Sosa, Arthur Mortha, Adeeb Rahman, Eduardo Farias, Kathryn Harper, et al. 2018. “Macrophages Orchestrate Breast Cancer Early Dissemination and Metastasis.” *Nature Communications* 9 (1): 21. https://doi.org/10.1038/s41467-017-02481-5.

Lips, Esther H., Tapsi Kumar, Anargyros Megalios, Lindy L. Visser, Michael Sheinman, Angelo Fortunato, Vandna Shah, et al. 2022. “Genomic Analysis Defines Clonal Relationships of Ductal Carcinoma in Situ and Recurrent Invasive Breast Cancer.” *Nature Genetics* 54 (6): 850–60. https://doi.org/10.1038/s41588-022-01082-3.

Ludwik, Katarzyna A., Zachary M. Sandusky, Eric B. Wright, and Deborah A. Lannigan. 2021. “FACS Protocol for Direct Comparison of Cell Populations Isolated from Mice.” *STAR Protocols* 2 (1): 100270. https://doi.org/10.1016/j.xpro.2020.100270.

Macdonald, I. 1951. “Biological Predeterminism in Human Cancer.” *Surg Gynecol Obstet* 92 (4): 443–52.

Macdonald, Ian. 1966. “The Natural History of Mammary Carcinoma.” *The American Journal of Surgery* 111 (3): 435–42. https://doi.org/10.1016/S0002-9610(66)80023-5.

Medina, Daniel. 1979. “Serial Transplantation of Chemical Carcinogen-Induced Mouse Mammary Ductal Dysplasias2.” *JNCI: Journal of the National Cancer Institute* 62 (2): 397–405. https://doi.org/10.1093/jnci/62.2.397.

———. 2000. “The Preneoplastic Phenotype in Murine Mammary Tumorigenesis.” *Journal of Mammary Gland Biology and Neoplasia* 5 (4): 393–407. https://doi.org/10.1023/A:1009529928422.

Morris, Max D. 1991. “Factorial Sampling Plans for Preliminary Computational Experiments.” *Technometrics* 33 (2): 161–74. https://doi.org/10.1080/00401706.1991.10484804.

Neerven, S. M. van, N. E. de Groot, L. E. Nijman, B. P. Scicluna, M. S. van Driel, M. C. Lecca, D. O. Warmerdam, et al. 2021. “Apc-Mutant Cells Act as Supercompetitors in Intestinal Tumour Initiation.” *Nature*, June. https://doi.org/10.1038/s41586-021-03558-4.

Nordling, C. O. 1953. “A New Theory on Cancer-Inducing Mechanism.” *Br J Cancer* 7 (1): 68–72. https://doi.org/10.1038/bjc.1953.8.

Pedersen, K., P. D. Angelini, S. Laos, A. Bach-Faig, M. P. Cunningham, C. Ferrer-Ramon, A. Luque-Garcia, et al. 2009. “A Naturally Occurring HER2 Carboxy-Terminal Fragment Promotes Mammary Tumor Growth and Metastasis.” *Mol Cell Biol* 29 (12): 3319–31. https://doi.org/10.1128/MCB.01803-08.

Pianosi, Francesca, Fanny Sarrazin, and Thorsten Wagener. 2015. “A Matlab Toolbox for Global Sensitivity Analysis.” *Environmental Modelling & Software* 70 (August): 80–85. https://doi.org/10.1016/j.envsoft.2015.04.009.

Pinheiro, José C., and Douglas M. Bates. 1995. “Approximations to the Log-Likelihood Function in the Nonlinear Mixed-Effects Model.” *Journal of Computational and Graphical Statistics* 4 (1): 12–35. https://doi.org/10.2307/1390625.

Rios, Anne C., Bianca D. Capaldo, François Vaillant, Bhupinder Pal, Ravian van Ineveld, Caleb A. Dawson, Yunshun Chen, et al. 2019. “Intraclonal Plasticity in Mammary Tumors Revealed through Large-Scale Single-Cell Resolution 3D Imaging.” *Cancer Cell* 35 (6): 953. https://doi.org/10.1016/j.ccell.2019.05.011.

Rueden, Curtis T., Johannes Schindelin, Mark C. Hiner, Barry E. DeZonia, Alison E. Walter, Ellen T. Arena, and Kevin W. Eliceiri. 2017. “ImageJ2: ImageJ for the next Generation of Scientific Image Data.” *BMC Bioinformatics* 18 (1): 529. https://doi.org/10.1186/s12859-017-1934-z.

Scaltriti, Maurizio, Federico Rojo, Alberto Ocaña, Judit Anido, Marta Guzman, Javier Cortes, Serena Di Cosimo, et al. 2007. “Expression of p95HER2, a Truncated Form of the HER2 Receptor, and Response to Anti-HER2 Therapies in Breast Cancer.” *JNCI: Journal of the National Cancer Institute* 99 (8): 628–38. https://doi.org/10.1093/jnci/djk134.

Sidani, Mazen, Jeffrey Wyckoff, Chengsen Xue, Jeffrey E. Segall, and John Condeelis. 2006. “Probing the Microenvironment of Mammary Tumors Using Multiphoton Microscopy.” *Journal of Mammary Gland Biology and Neoplasia* 11 (2): 151–63. https://doi.org/10.1007/s10911-006-9021-5.

Swanton, Charles, Elsa Bernard, Chris Abbosh, Fabrice André, Johan Auwerx, Allan Balmain, Dafna Bar-Sagi, et al. 2024. “Embracing Cancer Complexity: Hallmarks of Systemic Disease.” *Cell* 187 (7): 1589–1616. https://doi.org/10.1016/j.cell.2024.02.009.

Wang, Kaile, Tapsi Kumar, Junke Wang, Darlan Conterno Minussi, Emi Sei, Jianzhuo Li, Tuan M. Tran, et al. 2023. “Archival Single-Cell Genomics Reveals Persistent Subclones during DCIS Progression.” *Cell* 186 (18): 3968-3982.e15. https://doi.org/10.1016/j.cell.2023.07.024.

Welch, H. G., B. S. Kramer, and W. C. Black. 2019. “Epidemiologic Signatures in Cancer.” *N Engl J Med* 381 (14): 1378–86. https://doi.org/10.1056/NEJMsr1905447.

Welch, H. G., P. C. Prorok, A. J. O’Malley, and B. S. Kramer. 2016. “Breast-Cancer Tumor Size, Overdiagnosis, and Mammography Screening Effectiveness.” *N Engl J Med* 375 (15): 1438–47. https://doi.org/10.1056/NEJMoa1600249.

Welch, H. Gilbert, and Tanujit Dey. 2023. “Testing Whether Cancer Screening Saves Lives: Implications for Randomized Clinical Trials of Multicancer Screening.” *JAMA Internal Medicine* 183 (11): 1255–58. https://doi.org/10.1001/jamainternmed.2023.3781.