AWARD NUMBER: W81XWH-17-1-0174

TITLE: Characterization of Tumor Initiation Using a Novel Mouse Model of High-Grade Serous Ovarian Cancer

PRINCIPAL INVESTIGATOR: Dr. Jill Slack-Davis

CONTRACTING ORGANIZATION: Rector and Visitors of the University of Virginia

REPORT DATE: SEPTEMBER 2020

TYPE OF REPORT: Final Report

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
SEPTEMBER 2020	Final	15MAY2017 - 14MAY2020

4. TITLE AND SUBTITLE

Characterization of Tumor Initiation Using a Novel Mouse Model of High-Grade Serous Ovarian Cancer

5a. CONTRACT NUMBER
W81XWH-17-1-0174

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jill K. Slack-Davis; Hui Zong

E-Mail: jks6a@virginia.edu; hz9s@virginia.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Rector & Visitors of the University of Virginia
1001 N Emmett St.
Charlottesville, VA 22903-4833

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Early detection remains the most reliable approach for successful treatment of cancer. Unfortunately, the majority (80%) of ovarian cancer patients are diagnosed with advanced-stage high-grade serous cancer (HGSOC) and have a 5-year survival rate of <30%, due to the limitation of current screening methods. While information gleaned from research with patient samples and existing mouse models has provided invaluable information about malignant ovarian cancer including the identification of precursor lesions within the fallopian tube (i.e., serous tumor in situ carcinoma or STIC), it remains challenging to visualize, let alone analyze, mutant cells prior to the manifestation of pathological features. We developed a novel mouse model to facilitate the investigation of early molecular and cellular events leading to the development of HGSOC. The model immediately and permanently labels mutant cells upon the occurrence of initiating mutations and results in the formation of high-grade serous carcinoma. Importantly, the observed long latency (~40 weeks) of premalignant mutant cell expansion prior to the formation of STICs and serous tumors provides ample opportunity to 1) evaluate molecular and cellular processes involved in the early stages of STIC formation and 2) identify potential biomarkers for early detection.

15. SUBJECT TERMS
High-grade serous ovarian cancer, Serous tubal in situ carcinoma, Early detection, Cancer prevention, Mouse model, Gene expression analysis, RNA-Seq

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	44	19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

Page

1. INTRODUCTION

Early detection remains the most reliable approach for successful treatment of cancer. For ovarian cancer, patients diagnosed with Stage I ovarian cancer experience a 5-year survival of 90%. Unfortunately, the majority (80%) of ovarian cancer patients are diagnosed with advanced-stage high-grade serous cancer (HGSOC) and have a 5-year survival rate of <30%, due to the limitation of current screening methods. Therefore, the identification of biomarkers for early detection and an understanding of the progressive kinetics of pre-malignant mutant cells is vitally important to allow physicians to determine precisely when and how to intervene to achieve a more favorable outcome. While information gleaned from research with patient samples and existing mouse models has provided invaluable information about malignant ovarian cancer including the identification of precursor lesions within the fallopian tube, it remains challenging to visualize, let alone analyze, mutant cells prior to the manifestation of pathological features. Therefore, *essentially nothing is known about the molecular events leading from the occurrence of initiating mutations to the development of precursor lesions, the ideal period for developing sensitive and specific biomarkers for early detection.* This project addresses the gap in knowledge by developing a *novel* mouse model for ovarian cancer that can immediately and permanently label mutant cells upon the occurrence of initiating mutations. Using this system, we will first map the expansion kinetics of tumor-initiating cells in the fallopian tube to gain insights of pre-malignant tumor progression at an unprecedented temporal and spatial resolution. After we pinpoint the critical time points during early tumor progression, we will purify labeled mutant cells at these time points to identify changes in gene expression in comparison to the normal tissue. Finally, we will validate the most promising biomarkers in a cohort of human precursor lesions and early stage HGSOC identified by our group. The successful completion of this project will identify biomarkers prior to the development of pathologically identifiable tumors that can lead to clinical translation for cancer early detection. Ultimately, such information would enable the prevention of ovarian cancer in women, including military personnel and family members.

2. KEYWORDS

High-grade serous ovarian cancer
Serous tubal in situ carcinoma
Early detection
Cancer prevention
Mouse model
Gene expression analysis
RNA-Seq

3. ACCOMPLISHMENTS

A. Major Goals

Statement Of Work
Project Start Date: May 15, 2017

Specific Aim 1: Analyze the expansion kinetics of $Trp53^{-/-}$; $Brca1^{-/-}$; $Nf1^{-/-}$ cells in the fallopian tube.	**Timeline (Months)**	**Percent complete**
Major Task 1: Produce experimental mice		
Subtask 1: USAMRMC Animal Care and Use regulatory review	1-3	100%
Subtask 2: Expand two mouse stock lines for the project and maintain 10-20 breeding pairs to produce MADM-Trp53,Brca1,Nf1 mice for experimentation	4-36	100%
Subtask 3: Among all MADM-Trp53,Brca1,Nf1 mice produced, dissect 90 for experiments in Aim 1.	6-18	100%
Milestone(s) Achieved: Obtain USAMRMC regulatory approval and maintain steady supply of MADM-Trp53,Brca1,Nf1 mice for analysis		100%
Major Task 2: Analyze kinetics of cancer progression		
Subtask 1: Determine G/R ratio over time by counting GFP+ vs RFP+ cells in frozen sections from the fallopian tube	6-18	100%
Subtask 2: Immunohistologically stain fallopian tubes and ovarian cancer nodules for markers of HGSOC (Pax8, Stathmin, etc), apoptosis and proliferation	6-18	100%
Milestone(s) Achieved: Identify the kinetics of mutant cell expansion		100%
Specific Aim 2: Identify changes in gene expression upon loss of heterozygosity in $Trp53^{-/-}$; $Brca1^{-/-}$; $Nf1^{-/-}$ mutant mice.		
Major Task 3: Identify candidate biomarkers using the mouse model		
Subtask 1: Among all MADM-Trp53,Brca1,Nf1 mice produced, dissect 24 for gene expression studies.	9-18	100%
Subtask 2: USAMRMC Animal Care and Use regulatory review to add the ROSA26-LSL-tdTomato line to collect control cells for the LCM experiment	25-27	100%

Subtask 3: Breed and dissect ROSA26-LSL-tdTomato; Pax8-Cre mice for gene expression studies.	28-30	100%
Subtask 4: Laser-capture microdissection of tissue sections, RNA extraction and cDNA synthesis for gene expression analysis	31-32	20%
Subtask 5: RNA-Seq and gene expression analysis	33	Not started
Milestone(s) Achieved: Identify list of candidate genes that may participate in ovarian cancer initiation		
Major Task 4: Validate candidate biomarkers in human patient samples		
Subtask 1: Validate candidate genes in a separate group of MADM-Trp53,Brca1,Nf1 mice	34	Not started
Subtask 2: Evaluate protein expression of candidate genes in 30 de-identified human specimens including 10 Stage IA ovarian cancers, 10 STICs or neoplasms confined to the fallopian tube, and 10 non-transformed fallopian tubes by IHC	35-36	Not started
Milestone(s) Achieved: Identify unique gene expression patterns that contribute to the early events in HGSOC tumorigenesis; Publish 1-2 papers in peer-reviewed journals		

B. Accomplished to date

The goals of this research project were to develop and characterize a novel mouse model to study the initiation phase of high-grade serous ovarian cancer (HGSOC) with the ultimate goal of understanding the earliest events of HGSOC initiation and identifying specific markers that could facilitate early detection research. Early detection remains the most reliable approach for successful treatment of cancer. Indeed, patients diagnosed with Stage I ovarian cancer, which is confined to the ovary, experience a 5-year survival of 90%. Unfortunately, the majority (80%) of women with ovarian cancer are diagnosed at advanced stage, which is characterized by extensive spread within the peritoneal cavity and <30% 5-year survival rate. Therefore, the identification of biomarkers for early detection and an understanding of the progressive kinetics of pre-malignant mutant cells would be vitally important to allow physicians to determine precisely when to intervene to achieve a more favorable outcome.

We employed the genetic system called MADM (mosaic analysis with double markers), which immediately and permanently labels mutant cells upon the loss of heterozygosity (LOH) of tumor suppressor genes (TSGs), the tumor initiating event. From a heterozygous mouse, MADM generates sparse, GFP-labeled TSG-null cells and RFP-labeled wild type sibling cells in somatic tissues. This model is ideal for studying tumor initiation and early progression because 1) the rarity of TSG-null cells mimics sporadic LOH in human patients and the clonal origin of cancer, 2) the labeling of mutant cells with GFP allows us to investigate the entire process of tumorigenesis, and 3) the direct comparison of cellular behaviors between green mutant and red wild type cells allows the detection of even the subtlest anomalies at early stages. We have established a MADM-based model that inactivates three TSGs in Pax8+ fallopian tube cells (origin of HGSOC): p53, BRCA1 and NF1, all of which are frequently

observed in the most common and aggressive form of ovarian cancer (high-grade serous ovarian cancer).

Aim 1. Analyze the expansion kinetics of Trp53$^{-/-}$;Brca1$^{-/-}$;Nf1$^{-/-}$ cells in the fallopian tube. Using MADM, we proposed to investigate the effects of loss of *Trp53, Brca1,* and *Nf1* on tumor initiation and progression in fallopian tube epithelial cells.

a. *Expand and maintain breeding pairs to produce MADM-Trp53,Brca1,Nf1 mice.* Using the Pax8-rtTA, tetO-Cre transgenic mouse, Cre expression was induced in the secretory epithelium during mitotic expansion of the fallopian tube by providing doxycycline in the drinking water of nursing mothers. We have generated all experimental mice aged 8 – 52 weeks for the project.

b. *Determine the G/R ratio over time.* Evaluation of the fallopian tubes in experimental mice at various ages revealed significant mutant cell expansion and Ki67 positivity potentially identifying serous tubal in situ carcinoma (STIC). Understanding the kinetics of mutant cell expansion in this model provided critical information for LCM experiments in Specific Aim 2 (for details, see Year 2 Report, and the attached manuscript).

c. *Immunohistologically stain tissues for markers of HGSOC (Pax8, WT1, etc), apoptosis and proliferation.* One uterine horn, fallopian tube and ovary from each of 83 mutant and 10 wild type mice aged 8 – 52 weeks was fixed, embedded in paraffin, sectioned, stained with H&E or immunohistochemically (IHC) for Ki67 and evaluated by Dr. Atkins (surgical pathologist collaborator) for pathological abnormalities within the reproductive tract (Table 1). STICs and high grade serous (HGS) tumors of the uterus emerged at 32 weeks of age and increased in incidence through 52 weeks (Table 1), demonstrating that this is a valid HGSOC model.

d. *Conclusions.* Taken together, the data validate the MADM model for high-grade serous carcinoma. The long latency (~32 weeks) of premalignant mutant cell expansion prior to the emergence of STIC and serous tumors provides ample opportunity to 1) evaluate molecular and cellular processes involved in the early stages of STIC formation and 2) identify potential biomarkers for early detection; both of which are critical for successful completion of Aim 2.

Age (wks)	Total (n)	Histology – n (%)						
		Fallopian tube		Uterus				
		Normal	***STIC***	Normal	Hyper*	Cancer		
						HGS	Endo.**	Invasive***
8	7	7 (100)	***0***	7 (100)	0	***0***	0	0
12	5	6 (100)	***0***	6 (100)	0	***0***	0	0
16	6	6 (100)	***0***	6 (100)	0	***0***	0	0
20	9	9 (100)	***0***	9 (100)	0	***0***	0	0
24	16	16 (100)	***0***	15 (94)	1 (6)	***0***	0	0
32	7	6 (86)	***1 (14)***	5 (71)	1 (14)	***1 (14)***	0	0
40	9	7 (78)	***2 (22)***	0	2 (22)	***3 (33)***	0	4 (44)
48	5	3 (60)	***2 (40)***	1 (20)	0	***2 (40)***	0	2 (40)
52	19	16**** (84)	***1 (6)***	6 (32)	1 (6)	***8 (42)***	2 (10)	2 (10)

Table 1. Histopathological evaluation of fallopian tubes and uteri from TSG mutant mice. *Hyperplasia +/- atypia; ** Endometrioid; ***Histological subtype not specified; ****2 mice – fallopian tube engulfed by tumor. HGS = High Grade Serous

Aim 2. Identify changes in gene expression upon LOH in Trp53$^{-/-}$;Brca1$^{-/-}$; Nf1$^{-/-}$ mutant mice. The goals of this aim were to capitalize on the unequivocal, differential labeling of mutant (GFP+) and wild type (RFP+) cells and use laser-capture microdissection (LCM) to isolate GFP+ and RFP+ populations of cells to obtain comparative gene expression data using RNASeq and then to evaluate candidate genes in human specimens. To date, we have dissected all 24 MADM mice and stored (-80^{o} C) their fallopian tubes for LCM. Through the course of obtaining G/R ratio data, we realized that it would be difficult to obtain sufficient RFP+ cells from MADM mice because GFP+ cells greatly outnumber RFP+ ones. Therefore, during this period, we took an alternative approach to acquire wild type cells from a different mouse line after being approved by ACURO. Most importantly, our careful analysis revealed that GFP+ mutant clones progressed heterogeneously. While some grew into large clones, others remained small and appeared to be stalled. With this important discovery, we planned our LCM/RNAseq analysis with two critical considerations: 1) we will focus on large clones since they are most likely to progress into malignancy; 2) rather than combining samples, we will perform clone-by-clone analysis to increase our chance to identify unique signatures of early cancer progression. Unfortunately, while we were fully ready to perform the LCM experiment, the university shut down core facilities due to the COVID-19 situation. Therefore, we now have all the samples but will have to wait for a future time to analyze them.

C. <u>Opportunities for training and professional development</u> – Nothing to report

D. <u>Dissemination of results</u> – Nothing to report

E. <u>Future plans</u> – Nothing to report

4. IMPACT

A. Impact on the development of the principal discipline

While we weren't able to finish the last step of the project, we have drafted a manuscript to describe this unique mouse model and to unveil the critical finding of uneven progression of cells carrying genetic mutations. This finding suggests that gene mutations alone are not enough to drive tumor progression, and that most likely only some cells in the fallopian tube (such as stem cells) could progress upon genetic mutations. Identifying these cells and defining their signature genes would be very important to distinguish progressing from non-progressing early lesions, and reduce the risk of over-diagnosis and over-treatment. Furthermore, understanding the unique biology of these cells should shed light on effective cancer prevention strategies. Finally, we expect that our mouse model would be heavily sought after upon the publication of our manuscript to test drug candidates for cancer prevention of HGSOC.

B. **Impact on other disciplines** – Nothing to report

C. **Impact on technology transfer** – Nothing to report

D. **Impact on society** – Nothing to report

5. CHANGES/PROBLEMS

A. Changes in approach and reason for change

While we weren't able to perform the final experiment, we shifted from the original plan to analyze all mutant clones collectively to the new plan to profile each mutant clone individually. This change should help us identify key signatures that underlie the clonal heterogeneity, and pinpoint the biomarkers for progressing clones specifically.

B. Actual or anticipated problems of delays and actions or plans to resolve

COVID-19 stopped everything in its track.

C. Significant impact on expenditures

Unfortunately our fund ran out before we could perform the final experiment.

D. Impact on human subjects, vertebrate animals, biohazards and/or select agents

Nothing to report.

6. PRODUCTS

Nothing to report

7. PARTICIPANTS AND OTHER COLLABORATING ORGANIZATIONS

Name:	Jill K. Slack-Davis
Role on Project:	Principal Investigator
Researcher identification (e.g., ORCID ID):	
Nearest person month worked:	0.6
Contribution to project:	Headed studies characterizing ovarian cancer development in the MADM mouse model.
Funding support:	None

Name:	Hui Zong
Role on Project:	Co-Principal Investigator
Researcher identification (e.g., ORCID ID):	0000-0002-4263-9633
Nearest person month worked:	1.2
Contribution to project:	Obtained IACUC approval, oversaw the development and breeding of mouse models, and directed experiments.
Funding support:	None

Name:	Ying Jiang
Role on Project:	Research Scientist
Researcher identification (e.g., ORCID ID):	
Nearest person month worked:	3
Contribution to project:	Managed the breeding of mouse models including genotyping, weaning and distribution. Collected samples for the LCM experiment.
Funding support:	None

Name:	Eli Casarez
Role on Project:	Laboratory Technician
Researcher identification (e.g., ORCID ID):	
Nearest person month worked:	3

Contribution to project:	Dissected, processed and managed the distribution of all tissues.
Funding support:	None

Has there been a change in the active other support of the PD/PI or senior / key personnel since the last reporting period? No.

8. Special Reporting Requirements: None to Report

9. Appendices: manuscript that describes the findings from this project

A genetic mosaic mouse for modelling ovarian cancer initiation and early progression.

Jianhao R. Zeng [a], Eli Casarez [a], Ying Jiang [a], Ali Rohani [a], Kristen A. Atkins [b, c], Brianna E. Kelly [a], Eugene Ke [a], Shua Jeong [a], David Kashatus [a, c], Jill K. Slack Davis [a, c *], Hui Zong [a, c*].

[a] Department of Microbiology, Immunology, and Cancer Biology, [b] Department of Pathology, and [c] Cancer Center, University of Virginia, Charlottesville, Virginia, United States of America.

* Corresponding author. **Email:** hz9s@virginia.edu (Hui. Zong) jks6a@virginia.edu (Jill K. Slack- Davis)

Classification

Biological Science; Medical Sciences.

Keywords

Ovarian Cancer; Novel Mouse Model; Premalignancy; Clonal Heterogeneity.

Author Contributions

- J.R.Z. performed most of the experiments, analyzed the data, and contributed to team management for this project.
- E.C., B.E.K., and S.J. provided substantial technical assistance to this project and gained significant understanding of the project that merits authorship.
- E.K. performed the mutation spectrum analysis of human HGSOC patients.
- A.R. & D.K. developed the automated quantification algorithm for the quantification of green and red cells from light-sheet microscopy imaging data.
- Y.J. established the mouse model, bred the mice, and fed Dox according to the experimental design, and coordinated experiments within the team.
- K.A. performed the pathological review of STICs in fallopian tube.
- J.S-D. & H.Z. conceived the project idea, guided the experimental design and data interpretation, and wrote the manuscript.

Highlights

- A novel HGSOC mouse model that reveals initiation and pre-malignant progression.
- Streamlined analytical pipeline to visualize mutant cell expansion dynamics and to test efficacy of cancer prevention drugs with intact tissue.
- Clonal-resolution analysis of mutant cells expansion: heterogeneity due to cell intrinsic differences, or spatial/environmental differences.

Abstract

High-grade serous ovarian cancer (HGSOC) is the most common and lethal ovarian cancer subtype, with a 5-year survival rate of only 30% (1). The high mortality is mainly due to the typically late-stage diagnosis. While effective early diagnosis markers and preventive methods are in urgent need, progresses in the past decades are not significant. Mouse models that well recapitulate the early HGSOC development could extend our understanding of mechanism driving disease progression. Here, we utilized a genetic mosaic system to generate a mouse model which not only mimics the clonal origin of human cancer but also enables tracking of cancer cells with unprecedented spatiotemporal resolution. We mapped the expansion kinetics of mutant cells with a self-developed streamlined analytical pipeline, and found profound heterogeneity of clonal expansion caused by non-genetic heterogeneity.

Significance Statement

Animal models are crucial for studying pathogenic processes of cancer development. While most existing animal models focus on malignant stages, we describe a genetic mosaic tumor mouse model for studying the pre-malignant progression of cancerous cells. We also established a streamlined pipeline to analyze the mutant cell expansion dynamics at clonal level and identified profound heterogeneity caused by non-genetic heterogeneity.

Introduction

High-grade serous ovarian cancer (HGSOC) is the most prevailing and lethal ovarian cancer subtype which causes about 14,000 deaths annually in the United States (1, 2). The high mortality is partially attributed to the inability of early diagnosis due to the lack of effective detection methods. Indeed, the five-year survival rate for patients detected in stage I is 90%; and dramatically drops to <30% for those in stage III/IV (3). Early diagnosis markers have long been imperative, and such marker identification would require understanding of the cellular and molecular mechanisms underlying the abnormalities at early progression stages (4, 5). Since the access to human tissue from asymptomatic patients at early stages is limited, a mouse model that accurately recapitulates HGSOC initiation and early progression should significantly facilitate the discovery of early diagnostic markers.

Most current mouse models are powerful for studying tumor stage events, but have limitations in modelling human cancer initiation and early progression due to the inability to recapitulate the de novo clonal evolution of human cancer in which a single tumor-initiating cell gradually gains expansion advantages over the surrounding normal cells and eventually dominates the tissue (4, 6). The ability to manipulate genetic mutations at the single-cell level in mice could in principle overcome this limitation through generating single mutant cells for cancer initiation. Our lab has developed a mouse genetic system called mosaic analysis with double markers (MADM) that could satisfy this need (7). The MADM system utilizes Cre-mediated interchromosomal recombination in cell mitosis to concomitantly generate a single mutant cell and permanently label it with GFP fluorescence (**Fig. 1*A***). The rarity of single mutant cell islets in the tissue achieved from the low frequency of interchromosomal recombination (0.1%–1% or much lower in Cre active cells) closely mimics the clonal origin of human cancer (**Fig. 1*B***). The permanent GFP-labeling and the sparseness of mutant clones enable visualization, isolation and analysis of mutant cell clones, allowing exploration of clonal heterogeneity.

Besides mimicking the clonal nature of human cancer, an ideal genetically engineered mouse (GEM) model for human cancer initiation and early progression should also target the cancer cell of origin with genomic lesions seen in patients, as mutations and cell of origin both prominently affect tumor phenotype (8). The mutation spectrum of human HGSOC is characterized by the inactivation of *TP53* (96% of cases) as well as defects in homologous recombination due to alterations of *BRCA1/2* (50%) (9-12). Additionally, aberrant activation of the *Ras-MAPK* pathway through multiple mechanisms, including the loss of *NF1,* a Ras GAP, is also frequently found in over 60% of HGSOC patients (10, 13).

As for the cancer cell of origin of HGSOC, the fallopian tube (FT) secretory epithelial cell is becoming the stronger candidate over the ovarian surface epithelial cells (14, 15). Evaluation of ovary and FT from prophylactic surgery of women harboring germline *BRCA1* or *BRCA2* mutations revealed serous tubal in situ carcinoma (STIC) in the fimbria and distal FT. The STICs not only show p53 mutation, DNA damage, and increased proliferation, but also share the pathological and immunophenotypic features with HGSOC, thus are thought to be precursor lesions of HGSOC (16-19) . Moreover, transgenic mouse studies showed that deletion of *Trp53*, *Brca1* and *Pten* in the FT secretary epithelium could strongly lead to HGSOC that phenocopy human diseases, which further strengthened the idea that the FT is at least one of, if not the more likely site of origin for HGSOC (14, 20).

In this study, with the objective to develop a mouse for faithful modeling of human HGSOC initiation and early progression, we have established a MADM-based mouse model, in which sparse mutant cells in the cancer cell of origin group initiate cancer. We have also characterized the early progression kinetics and interrogated the clonal heterogeneity during oncogenesis. Our novel GEM model presents a powerful tool for studying the cellular and molecular changes in HGSOC initiation and early progression, which could pave the way for developing more efficient early diagnostic methods.

Results

Establishment of a MADM-based HGSOC mouse model.

To establish a MADM-based mouse model for HGSOC, we took two critical prerequisites into considerations: 1) a Cre transgene that could specifically induce mutations in the cell of origin for HGSOC; 2) genes that are commonly mutated in HGSOC and are located on the telomeric side to the MADM cassettes (**Fig. 2*A***). With such considerations, we chose a Pax8-rtTA; TetO-Cre system that drives Cre expression with temporal control specifically in the FT secretary cells (14, 21) (**Fig. 2*B***). Next, we assessed the mutation spectrum in HGSOC patients with TCGA data-sets, and found that *TP53* (96%), *BRCA1/2* (50%) and genes within the Ras-MAPK pathway (60%) (**Fig. 2*C***) are among the most frequently mutated genes (9, 10, 13). Therefore, we chose *TP53, BRCA1* and *NF1,* a negative regulator of Ras, since they all reside on mouse chromosome 11 between the telomere and MADM cassettes (**Fig. 2*D***) (22). Subsequently, we bred mouse stocks containing the MADM cassettes, the *TP53, BRCA1,* and *NF1* mutant alleles, and Pax8-rtTA; TetO-Cre to established the MADM model for HGSOC (**Fig. *S*1**).

Verification of the sparseness of MADM generated GFP+ cells in the FT.

A unique advantage of the MADM-based cancer model is the sparseness of initial mutant cells, which closely mimics the clonal origin of human cancers. To determine the frequency and specificity of PAX8 promotor driven MADM recombination events, we induced Cre activity in MADM-WT mice (no mutations associated, **Fig. *S*1**) from P1 to P21, and harvested FTs one day after (**Fig. 3*A***). We observed sparse GFP+ cells in the FT and uterus but not in the ovary (**Fig. 3*B***). Among all the FT epithelial cells, less than 1% were GFP+ or tdTomato+ (**Fig. 3*C***), which is the typical sparseness of MADM labeling. We then examined faithfulness of PAX8 promoter driven Cre by staining for PAX8 protein in MADM labelled cells. We found that most MADM colored cells are also PAX8+ (**Fig. 3*D***), suggesting that the right cell type is targeted. It is noteworthy that a considerable portion of MADM-labeled cells were PAX8-, which could due

to the differentiation of PAX8+ cells to PAX8- ciliated cells as reported in literature (23). In summary, we confirmed the low frequency and the specificity of MADM labeling in the FT secretary cells, which is the cancer cell of origin of HGSOC.

Sparse mutant cells gradually progressed into STIC lesions In Vivo

Next, we investigated whether the sparse GFP+ mutant cells could progress to premalignant STIC lesions. Based on histopathological features and Ki67 status, FTs from mice younger than 32 weeks displayed little abnormality; however, after 40 weeks, FT epithelial cells showed increased nuclear atypia and Ki67+ staining, characteristic of STICs, which are recognized as the precursor of HGSOC (**Fig. 4*A***). In 82 mice that are 32 weeks or older, we reviewed one FT slides from each mouse and identified 6 STICs lesion (7%) (**Fig. 4*B***). Consider the low sampling coverage with this method, the true incidence of STICs could be higher. Importantly, these STIC regions were GFP+, Ki67+ and PAX8+ (**Fig. 4*C***), demonstrating that they develop from GFP+ mutant cells generated by MADM. These data suggest that the initial GFP+ mutant cells can develop into STIC lesions in vivo.

MADM-based model allows monitoring of mutant cell expansion kinetics.

The mutant cell progression from onset to STICs is long, we next intend to monitor mutant cell expansion kinetics to get a glance on the cellular activities in this slow progression process. MADM mouse model provides a unique parameter for this goal: the green to red cell number ratio (**G/R** ratio). Because MADM simultaneously generates GFP+ (green) mutant cells and tdTomato+ (red) wildtype cells with equal number initially, the **G/R** ratio could quantitatively evaluate the extent of mutant cell abnormal expansion. The larger the G/R ratio, the more over-expansion the mutant cells have undergone. The G/R ratio change with time could reflect the expansion kinetics of mutant cells. We measured the G/R ratio of FTs from mice at a series of ages between 8-52 weeks old with sectioned slides. We found that even at 16 weeks old, the

G/R ratio was small (< 10), and there was only a slight increase by 24 weeks. In contrast, the G/R ratio rapidly increased after 32 weeks (**Fig. 5*A***). The G/R ratio curve with time showed a punctuated two-phase increase: a relatively stable phase before 32 weeks followed by a steep increasing phase (**Fig. 5*B***). This kinetics could imply a transition of the mutant cells from a quiescent phase to an escalated phase. The escalated phase of disease progression could be the more critical window for identifying early diagnosis markers and testing efficacy of cancer prevention strategies. Overall, with the G/R ratio parameter, our MADM-based tumor mouse model presents a unique tool for studying mutant cell expansion kinetics at an unprecedented early stage with a high sensitivity.

A pipeline combined large-scale 3D imaging and computational image-processing for fast and accurate G/R ratio analysis.

G/R ratio counting with tissue sections is time-consuming, as intensive sampling is required for good coverage, because the FT is a complex organ with four distinct segments and the MADM-colored clones distribute stochastically. Moreover, this method loses the spatial distribution of clones and fails to resolve the heterogeneity among clones. Hence, it is not optimal for a large number of samples and for clonal level analysis. To overcome these hurdles, we developed a pipeline that combines large-scale 3D imaging with computational processing for rapid census of G/R ratio within the entire FT and for characterization of the clonal heterogeneity. In the pipeline, we first adopted the CUBIC tissue-clearing method to remove the opaqueness of FTs (from five weeks old mutant mice) (**Fig. 6*A***). We then used the light-sheet microscope to acquire a full view of the abundance and distribution of clones **(Fig. 6*B*)**. After that, we developed an image processing software package called "**G**/**R**-count" to quantify the G/R ratio of image stacks obtained from light-sheet microscopy (**Fig. *6C,*** method for G/R-count software). The estimated overall G/R ratio, the G/R ratio at each Z-depth, and the green/red area at each Z-depth can be generated within minutes (**Fig. *6D, E, F***). We verified the pipeline's accuracy by

comparing the G/R ratio reads from the pipeline with those counted manually and found no significant differences (**Fig. 6*G***). This pipeline estimates the G/R ratio not only at the population level, but also at the clonal level. Interestingly, the output—green area at each Z-depth curve—showed peaks of distinct height， two of which are prominently higher. Each peak is likely to represent an independent clone because of the sparseness of MADM labeling and minimal lateral migration. The variances in peak height should reflect size differences of the expanded mutant clones (**Fig. *6E***). Indeed, with 3D imaging of FTs, we observed multiple expanded mutant clones of distinct clonal size (**Fig. *6B***) suggesting heterogeneous expansion of mutant clones even though derived from cells with the same mutations. Notably, high peaks representing the wildtype red clones coexisted with green peaks (**Fig. *6E***), suggesting the expansion of wildtype clones. This indicates that genetic mutations are not required for clonal expansion, although mutations do boost the extent of clonal expansion as the green curve tends to be higher than red curves. In summary, this pipeline not only reduced the workload of population level G/R ratio counting from days to hours without compromising on accuracy, but also revealed important insights on clonal level heterogeneity.

MADM-based model allows clonal analysis of cellular heterogeneity.

Intrigued by the heterogeneous size of expanded mutant clones, we then thoroughly assessed the size of all mutant clones in 5, 8 and 20-week-old mice, and surprisingly found that even shortly after birth, the mutant clones showed prominent heterogeneity: the majority of clones remained at small size (<10 cells) and about ten clones had dramatically expanded (**Fig. 7*A1***). This trend of clonal size heterogeneity was consistently observed as mice got older (**Fig. 7*A2*, *A3***). Notably, the majority of clones were consistently small without expanding, even though the mice are getting older (**Fig. 7*B***). These data suggest that only a small portion of mutant cells is progressing, whereas the vast majority of mutant cells were stalled at the beginning. Next, we

asked whether the heterogeneous expansion ability of initiated cells is caused by genetic heterogeneity or non-genetic heterogeneity or both. The genetic heterogeneity due to differentially acquired extra mutations of each mutant clone, and the non-genetic heterogeneity inherited from the distinct cell states of the initial cells (epigenetic modifications or other mechanism), are both possible for inducing the heterogeneous clonal expansion (**Fig. 7C**).

We tested these possibilities by analyzing clonal size distribution in wild type MADM mice, in which all the clones are genetically identical. We performed 3-day DOX treatment (P19-P21) in wild type MADM mice to induce clones with synchronized birth time, and then assessed the clonal size distribution. Interestingly, we still observed prominent size heterogeneity (**Fig. 7*D***), (use standard deviation to characterize the heterogeneity), implying that acquired mutations are unlikely responsible for size heterogeneity. However, the absolute size of expanded clones in these wildtype mice are smaller than those expanded mutant clones (**Fig. 7*E***), suggesting that genetic mutations did augment the extent of clonal expansion. In summary, these data suggest that non-genetic heterogeneity leads to the heterogeneous clonal expansion ability and genetic mutations could exaggerate the extent of variance. Collectively, these experiments demonstrate the strength of our model for analyzing cellular heterogeneities at a clonal level of an unprecedented early stage.

Discussion

Early detection and prevention of HGSOC is invaluable to people with inherited mutations while little progress has been made mainly due to the lack of understanding of the early phase progression of disease. The recent discovery that at least a portion of HGSOCs originate from the distal fallopian tube has shift the attention of preventive research from ovary to fallopian tube. Here, we described a novel mouse model in which carcinogenesis starts from the fallopian tube and gradually progresses to STICs. Our model accurately mimics cancer initiation and faithfully recapitulates the early phase progression of disease, thus present to be superb for studying early prevention of HGSOCs.

As a preliminary proof of principles, we performed the testing of a potential cancer preventive drug PARPi, which could induce synthetic lethality in the BRCA1 mutated cells. With drug treatment between 20-40 weeks, we did not observe significant decline of both the green mutant cell percentage and the G/R ratio. The reasons for ineffectiveness of PARPi for prevention which however has been approved for the treatment of advanced ovarian cancer (24) could be at these the following three possibility 1 PARPi are known to be more effectively killing cells that are in cycle (25) , although mutant cells in the FT at this stage have expanded, the proliferation is slow, hence not sensitive to PARPi; 2) the extra NF1 mutation rescued the synthetic lethality effect of PARPi; 3) the clones evolved and resistant clones occur. Although we did not get positive result from this experiment, we have demonstrated how to use our model for testing cancer preventive drug, and how to use the green cell percentage and G/R ratio as alternative high-resolution and time-saving readouts of the drug effects other than the latency change of tumor formation. Other promising chemo-preventive drugs like fenretinide could be test with our platform in the future.

We also demonstrate how the clonal resolution in our model allows better understanding of the

heterogeneity of early phase disease. The inter-clonal heterogeneity could fail the cancer preventive interventions, as treatments not only deplete sensitive clones, but also select out resistant clones. Mapping out the spectrum of the heterogeneity among pre-cancer clones would facilitate the design of optimized of interventions, such as higher coverage cancer preventive chemo “cocktail”. We utilized our model to assess the heterogeneity of expansion kinetics of cancer initiating cells clonally with temporal resolution. We found that the PAX8+ cells are more complex than our understanding, there are a subset of cells that have higher potential to progress to tumor once initiated with mutation. A better understanding of these cells is important for narrowing down the target for preventive studies. Assess the transcriptional profile and the abnormality of highly expanded mutant clones rather than all mutant cells would be more likely to identify a good early diagnosis marker with lower false positive chance. This is an example of how our model could be used to study other aspect of clonal heterogeneity such as drug response, cellular plasticity et al.

In summary, our MADM-based tumor mouse model presents an important tool to study the biology of the initial phase of cancer development and to test efficacy of cancer-prevention candidate drugs, mainly due to the following strengths: 1) the sparse generation of GFP+ mutant cells precisely mimics human cancer initiation allows tracking of disease progression at cellular level; 2) the co-existence of equal number internal wildtype control cells provides G/R ratio change as a time-saving readout drug efficacy. Of note, this scheme of MADM-modeling of cancer in principle can be applied to any other tumor type as well as genetic diseases, and should promote the understanding of their initiation and promotion.

Materials and Methods

Animals

All animal procedures are approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Virginia in accordance with national guidelines to ensure humanity of all animal experiments. The following mouse lines were used to generate experimental and control MADM mice: *TG11ML* (stock NO. 022976 JAX), *GT11ML* (stock NO. 022977 JAX), *BRCA1*^{flox} (strain NO. 01XC9 NCI), *NF1*^{flox} (strain no. 01XM4; NCI), *TP53*^{KO} (stock no. 002101; JAX), *PAX8-rtTA* (stock NO. 007176 JAX), *TetO-Cre* (stock NO. 006234 JAX).

Doxycycline induction of Cre recombinase

Doxycycline was administered through the drinking water (2mg/ml) of mom and transmit to MADM pulps through milk (P1-P21).

5-ethynyl-2'–deoxyuridine (EdU) labeling

EdU (Invitrogen, Cat# A10044) was administered by i.p. injection (50mg/kg body weight) for short term labeling (24 hours). Mice were sacrificed after treatment and the fallopian tube was collected for analysis.

Immunostaining

Reproductive tracts of MADM mouse were harvested at a variety of ages then fixed overnight with cold 4% paraformaldehyde (PFA) at 4°C. Tissues were then washed with PBS to remove recessive PFA and soaked with 30% sucrose, embedded in optimal cutting temperature (OCT), and stored at -80°C. The tissues were then cryo-sectioned with 18μm thickness on glass slides. Slides with were warmed to room temperature followed by PBS wash to remove OCT. For staining, slides were incubated in Permeabilization/block buffer (0.3% Triton-X 100 in PBS

[PBT] plus 5% normal donkey serum) for 20 min then primary antibody (PAX8, 1:100 dilution ratio, company) diluted in Permeabilization/block buffer was added and incubated at 4°C overnight. Secondary antibody (Donkey anti-rat AF647 1:200 dilution ratio) incubation was performed for 1 hours at RT in PBT. To stain nuclei, slides were incubated in DAPI solution (1ug/mL in PBT) for 5 min before mounted with 70% glycerol and covered. Fluorescent images were acquired on Zeiss LSM 700/710 confocal microscope. Images were processed with Zen, Fiji and Adobe Photoshop.

Tissue clearing with CUBIC method

The PFA fixed ovary and fallopian tube were cleared for large-scale 3D imaging with the standard CUBIC method (26). Tissues were first immersed in 50% reagent-1 and shake at 110RPM 37°C for 6h, then transferred to 100% reagent-1with DAPI (1ug/ml) to shake for 2-3 days until the becoming transparent. After reagent-1. tissues were wash three times with PBST, 30mins each time with shaking to remove the reagent-1. Then tissues were immersed in 50% reagent-2 for 6h with shaking at 37°C, thereafter, exchange buffer to 100% reagent-2 to shake 24h.

3D imaging

The Zeiss Z.1 light sheet microscopy system was used for acquiring images. Tissues were embedded within 2% (w/w) agarose (LE quick dissolve agarose, GeneMate E-3119-500) gel, the agarose was dissolved in a modified reagent-2 (10% urea, 50% sucrose, 30% H_2O, hot stir at 80°C until fully dissolved, then add 10% Tri-ethanolamine). We aspirated warm agarose gel solution together with tissue into a 1ml syringe which has the neck cut off, then put the syringe on ice for a quick solidification. The syringe could be placed on the holder of the light-sheet microscope, and the tissue could be pushed out for imaging.

Clone size analysis

The clone size of GFP+ mutant epithelial cells in the FT (Fig. 6*B*, *C*) were through the fallowing method: fallopian tube tissues were continuously cut through with 25um thickness, all slides were fixed and the whole FT region were imaged with tile-scanning. Then GFP+ cells continuously showed at the same physical location were counted at a single clone, the number were counted manually.

Acknowledgments

We would like to thank Dr. Ammasi Periasamy and Dr. Ruofan Cao from the Keck Center for Cellular imaging for their assistance of using Zeiss Light Sheet microscopy system for 3D images; and the UVa Advanced Microscopy Facility for image acquisition and analysis (Dr. Stacey Criswell); Shelly Verling from the vivarium for animal care; Dr. Pat Pramoonjago from the Biorepository and Tissue Research Facility and Sheri Vanhoose from the Research Histology Core for assistance with the immuno-histochemistry staining. Joceylyn Ray for help with the art work of fallopian tube cartoon. This work was supported by grant from the Department of Defense (W81XWH-17-1-0174). Jianhao Zeng was supported by UVA cancer center training grant.

References

1. Vaughan S*, et al.* (2011) Rethinking ovarian cancer: recommendations for improving outcomes. *Nat Rev Cancer* 11(10):719-725.
2. Siegel RL, Miller KD, & Jemal A (2020) Cancer statistics, 2020. *CA Cancer J Clin* 70(1):7-30.
3. Hoskins WJ (1995) Prospective on ovarian cancer: why prevent? *Journal of cellular biochemistry. Supplement* 23:189-199.
4. Cheon DJ & Orsulic S (2011) Mouse models of cancer. *Annu Rev Pathol* 6:95-119.
5. Bowtell DD*, et al.* (2015) Rethinking ovarian cancer II: reducing mortality from high-grade serous ovarian cancer. *Nat Rev Cancer* 15(11):668-679.
6. Greaves M & Maley CC (2012) Clonal evolution in cancer. *Nature* 481(7381):306-313.
7. Zong H, Espinosa JS, Su HH, Muzumdar MD, & Luo L (2005) Mosaic analysis with double markers in mice. *Cell* 121(3):479-492.
8. Visvader JE (2011) Cells of origin in cancer. *Nature* 469(7330):314-322.
9. Walsh T CS, Lee MK, Pennil CC, Nord AS, Thornton AM, Roeb W, Agnew KJ, Stray SM, Wickramanayake A, Norquist B, Pennington KP, Garcia RL, King MC, Swisher EM. (2011) Mutations in 12 genes for inherited ovarian, fallopian tube, and peritoneal carcinoma identified by massively parallel sequencing. *Proc Nati Acad Sci USA* 108(44):18032-18037.
10. Anonymous (2011) Integrated genomic analyses of ovarian carcinoma. *Nature* 474(7353):609-615.
11. Ahmed AA*, et al.* (2010) Driver mutations in TP53 are ubiquitous in high grade serous carcinoma of the ovary. *J Pathol* 221(1):49-56.
12. Patch AM*, et al.* (2015) Whole-genome characterization of chemoresistant ovarian cancer. *Nature* (1476-4687 (Electronic)).
13. Angela Toss, Chiara Tomasello, 1 Elisabetta Razzaboni, 1 Giannina Contu, 2 Giovanni Grandi, 2 Angelo Cagnacci, 2 Russell J. Schilder, 3 and Laura Cortesi 1 (2015) Hereditary Ovarian Cancer: Not Only BRCA 1 and 2 Genes. *Biomed Res Int* (341723).
14. Perets R*, et al.* (2013) Transformation of the fallopian tube secretory epithelium leads to high-grade serous ovarian cancer in Brca;Tp53;Pten models. *Cancer Cell* 24(6):751-765.
15. Zhang S*, et al.* (2019) Both fallopian tube and ovarian surface epithelium are cells-of-origin for high-grade serous ovarian carcinoma. *Nat Commun* 10(1):5367.

16. Press JZ*, et al.* (2008) Ovarian carcinomas with genetic and epigenetic BRCA1 loss have distinct molecular abnormalities. *BMC Cancer* 8:17.
17. Barakat RR FM, Saigo PE, Robson ME, Offit K, Boyd J. (2000) Absence of premalignant histologic, molecular, or cell biologic alterations in prophylactic oophorectomy specimens from BRCA1 heterozygotes. *Cancer* 89(2):383-390.
18. Medeiros F MM, Lee Y, Elvin JA, Callahan MJ, Feltmate C, Garber JE, Cramer DW, Crum CP. (2006) The tubal fimbria is a preferred site for early adenocarcinoma in women with familial ovarian cancer syndrome.
. *Am J Surg Pathol.* 30(2):230-236.
19. Lee Y*, et al.* (2007) A candidate precursor to serous carcinoma that originates in the distal fallopian tube. *The Journal of pathology* 211(1):26-35.
20. Karnezis AN, Cho KR, Gilks CB, Pearce CL, & Huntsman DG (2017) The disparate origins of ovarian cancers: pathogenesis and prevention strategies. *Nat Rev Cancer* 17(1):65-74.
21. Traykova-Brauch M*, et al.* (2008) An efficient and versatile system for acute and chronic modulation of renal tubular function in transgenic mice. *Nat Med* 14(9):979-984.
22. Tasic B*, et al.* (2012) Extensions of MADM (Mosaic Analysis with Double Markers) in Mice. *PLoS One* 7(3):e33332.
23. Ghosh A, Syed SM, & Tanwar PS (2017) In vivo genetic cell lineage tracing reveals that oviductal secretory cells self-renew and give rise to ciliated cells. *Development* 144(17):3031-3041.
24. Moore K*, et al.* (2018) Maintenance Olaparib in Patients with Newly Diagnosed Advanced Ovarian Cancer. *N Engl J Med* 379(26):2495-2505.
25. Schoonen PM*, et al.* (2017) Progression through mitosis promotes PARP inhibitor-induced cytotoxicity in homologous recombination-deficient cancer cells. *Nature communications* 8:15981.
26. Susaki EA*, et al.* (2015) Advanced CUBIC protocols for whole-brain and whole-body clearing and imaging. *Nat Protoc* 10(11):1709-1727.

Figures and Tables

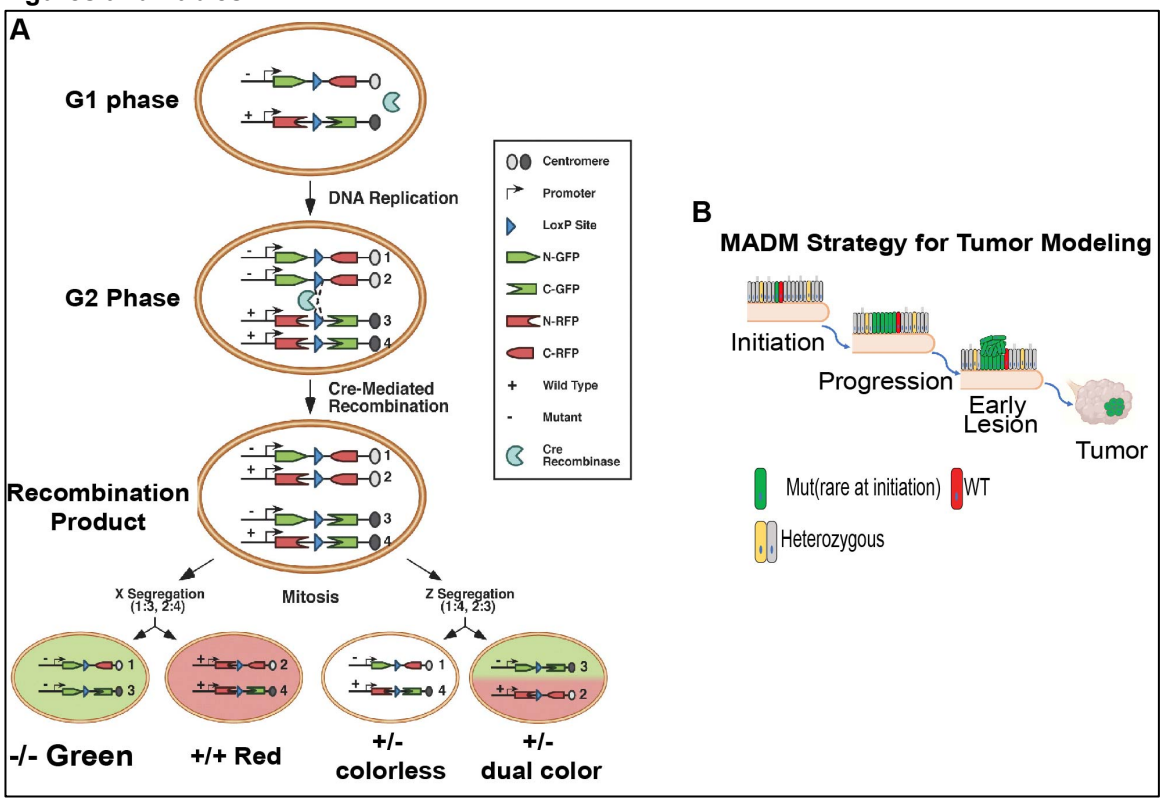


Figure 1. Scheme of the MADM strategy for modeling initial phase of cancer.

(A) MADM contains a pair of knock-in cassettes of chimeric GFP (green, G) and RFP (red, R) coding sequences, separated by a loxP-containing intron. Cre-mediated inter-chromosomal mitotic recombination followed by X segregation generates singly labeled cells that alter genotype if the original cell is heterozygous for a mutation of interest. G2 recombination followed by Z segregation generates either colorless or double-colored cells without altering genotype.

(B) MADM generates a small number of GFP+ homozygous mutant cells, which initiate the cancer development and gradually progress to pre-malignancy and tumor.

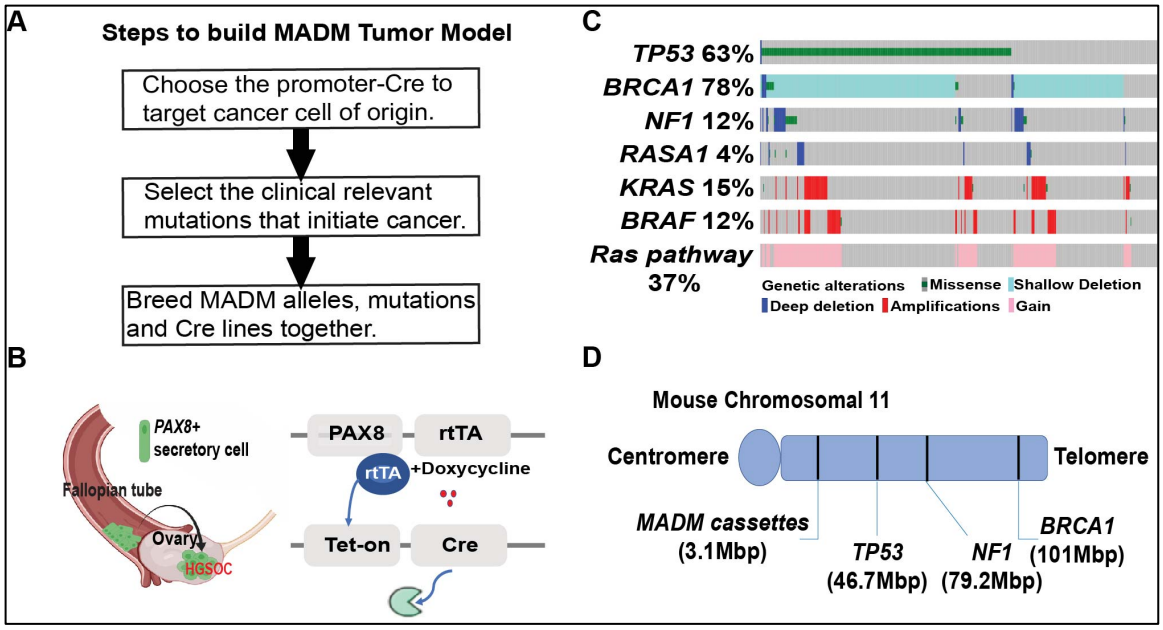


Figure 2. The establishment of a MADM-based HGSC mouse model.

(A) The diagram of how to build a MADM-based tumor model.

(B) Fallopian tube secretory epithelial cells as cancer cell of origin of HGSC and the Tet-On strategy.

(C) TCGA data-sets of HGSC demonstrates inactivation/deletion of *TP53*(96%), *BRCA1*/2 (50%) and aberrant activation of Ras-MAPK pathway (60%) including the loss of *NF1*.

(D) *TP53, BRCA1* and *NF1* genes are located in mouse chromosomal 11. MADM cassettes (TG and GT) were inserted into the Hipp11 locus also on chromosomal 11 to couple color labelling with specific genotype.

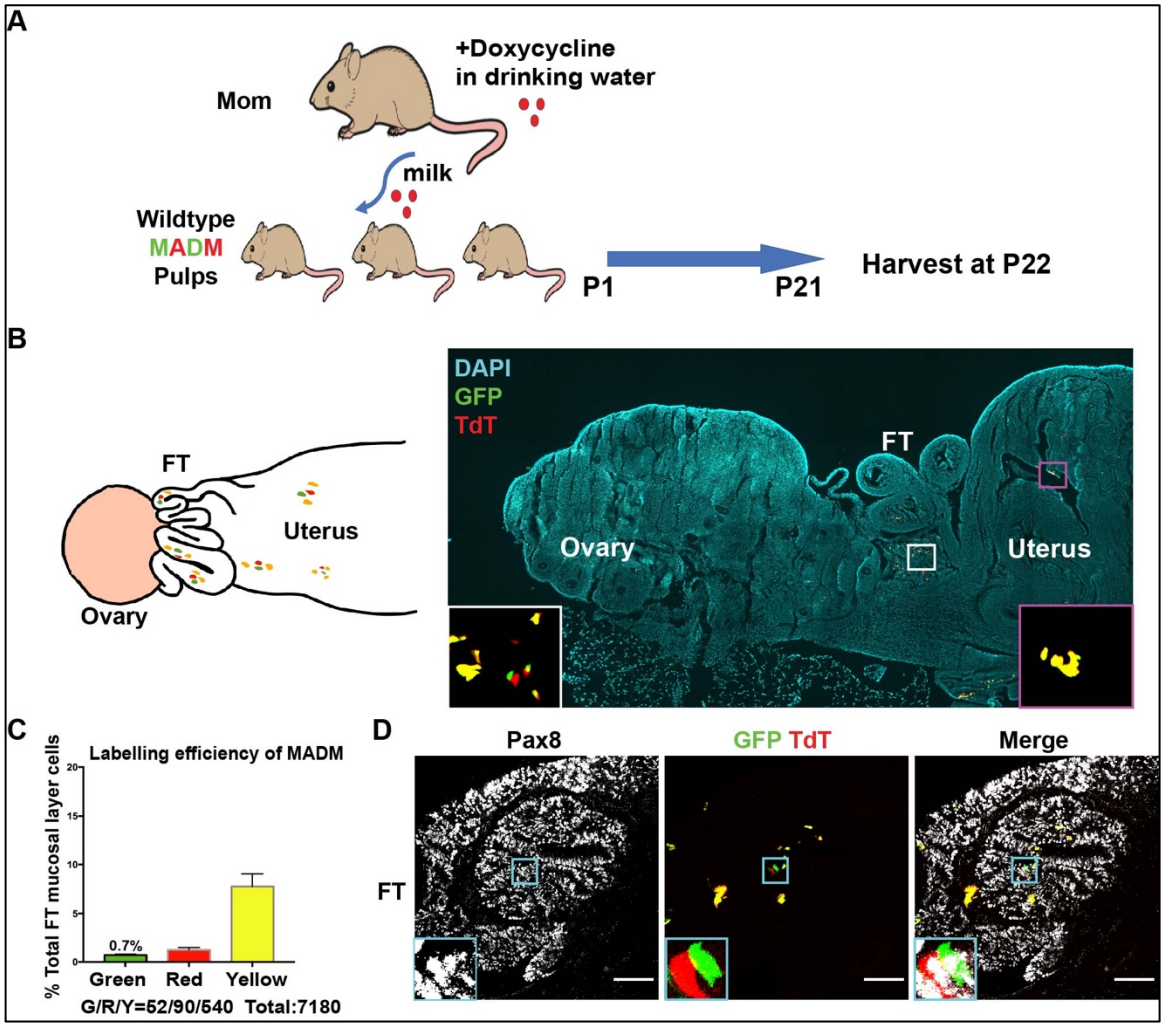


Figure 3. Validation of the sparseness of labeling in MADM mouse FT.

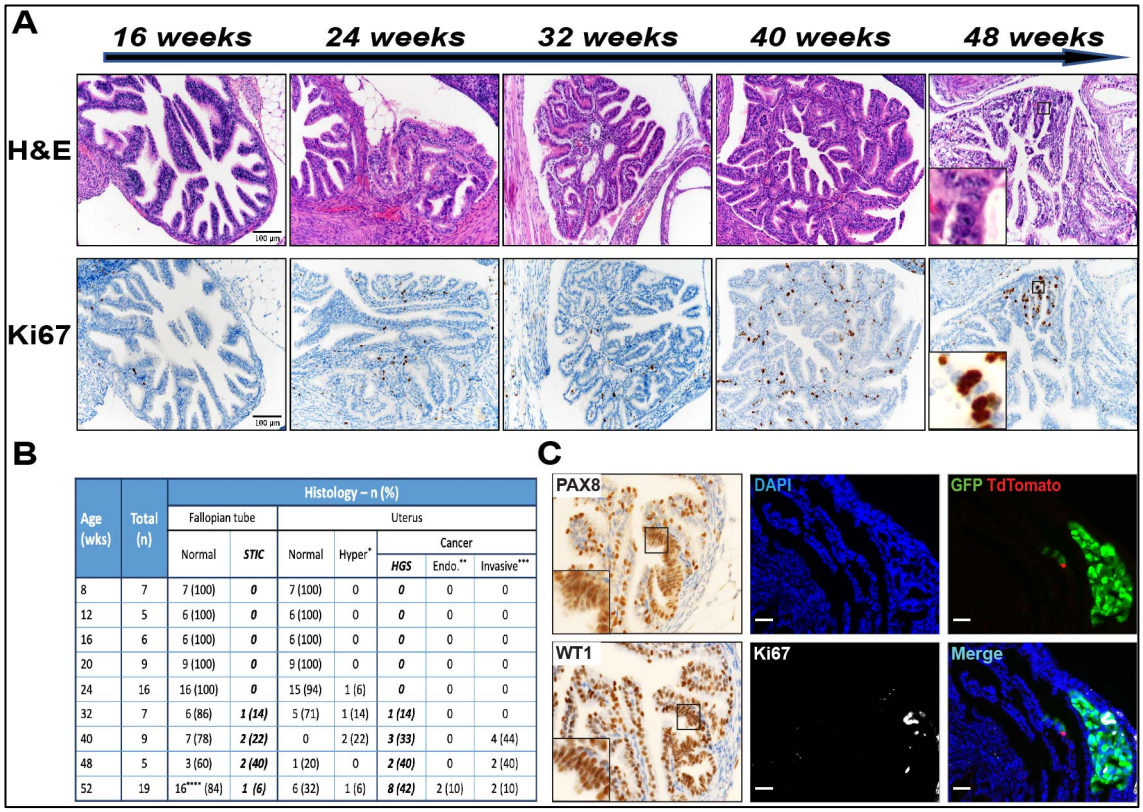
(A) Schematic of doxycycline administration to induce labeling in MADM mouse. 2mg/ml doxycycline was added to the drinking water of mom after giving birth and before wean (P1-P21), and transmit to MADM pulps through milk. Reproductive tracts of MADM mouse were harvested at 8 weeks age.

(B) Immunofluorescence analysis of cutted slides (20um thickness) of the reproductive tract from 8 weeks wildtype MADM mouse, representative of 3 mice.

(C) Quantification of the percentage of green, red and yellow cells in the FT mucosal

layer cells, data collected from three mice.

(D) Immunofluorescence staining of Pax8 of FT from 8 weeks wildtype MADM mouse, representative of 3 mice.



Age (wks)	Total (n)	Histology – n (%)						
		Fallopian tube		Uterus				
		Normal	*STIC*	Normal	Hyper*	Cancer		
						HGS	Endo.**	Invasive***
8	7	7 (100)	***0***	7 (100)	0	***0***	0	0
12	5	6 (100)	***0***	6 (100)	0	***0***	0	0
16	6	6 (100)	***0***	6 (100)	0	***0***	0	0
20	9	9 (100)	***0***	9 (100)	0	***0***	0	0
24	16	16 (100)	***0***	15 (94)	1 (6)	***0***	0	0
32	7	6 (86)	***1 (14)***	5 (71)	1 (14)	***1 (14)***	0	0
40	9	7 (78)	***2 (22)***	0	2 (22)	***3 (33)***	0	4 (44)
48	5	3 (60)	***2 (40)***	1 (20)	0	***2 (40)***	0	2 (40)
52	19	16**** (84)	***1 (6)***	6 (32)	1 (6)	***8 (42)***	2 (10)	2 (10)

Figure 4. Initial mutant cells gradually progressed to STIC lesions in *vivo*.

(A) Representative H&E and Ki67 staining(IHC) of fallopian tube from MADM mutant mouse at a variety of ages, 48 weeks FT were diagnosed with STIC by pathologist.

(B) Frequency of STICs lesions and Uterus cancer identified from a cohort of mice.

(C) Immunofluorescence staining of Ki67 with expanded green mutant cells in FT from 48 weeks mice.

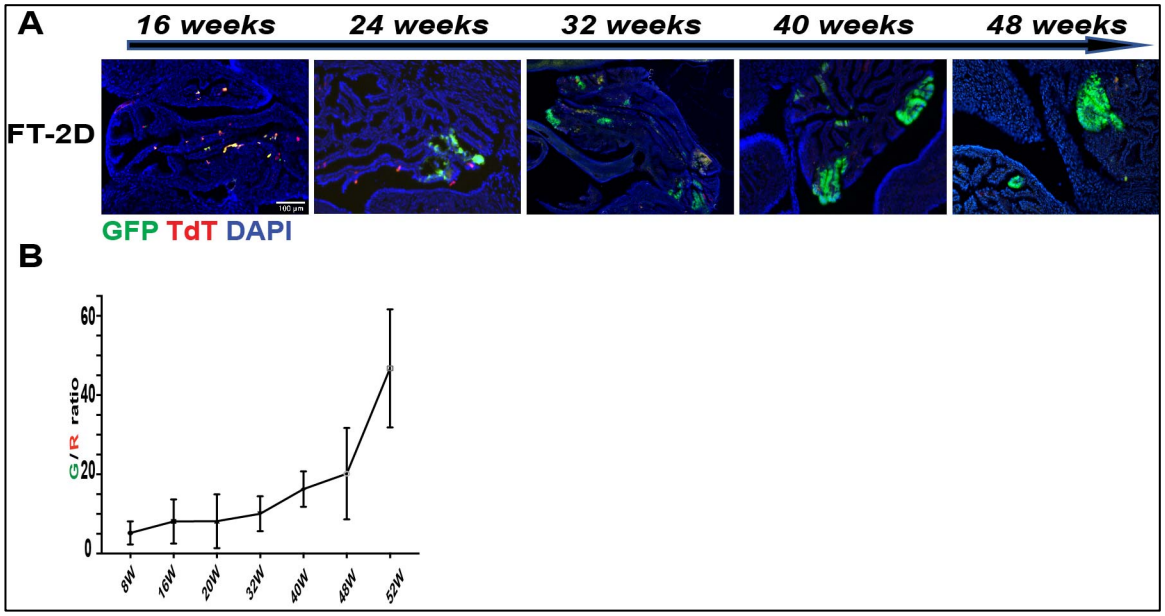


Figure 5. MADM-based model allows precise monitor of mutant cell expansion kinetics in early progression.

(A) Representative 2D images of FT slides from MADM mutant mouse at a variety of ages.

(B) The green mutant cell to red sibling wildtype cell ratio in FT from MADM mutant mouse at variety of ages quantified from tissue sections.

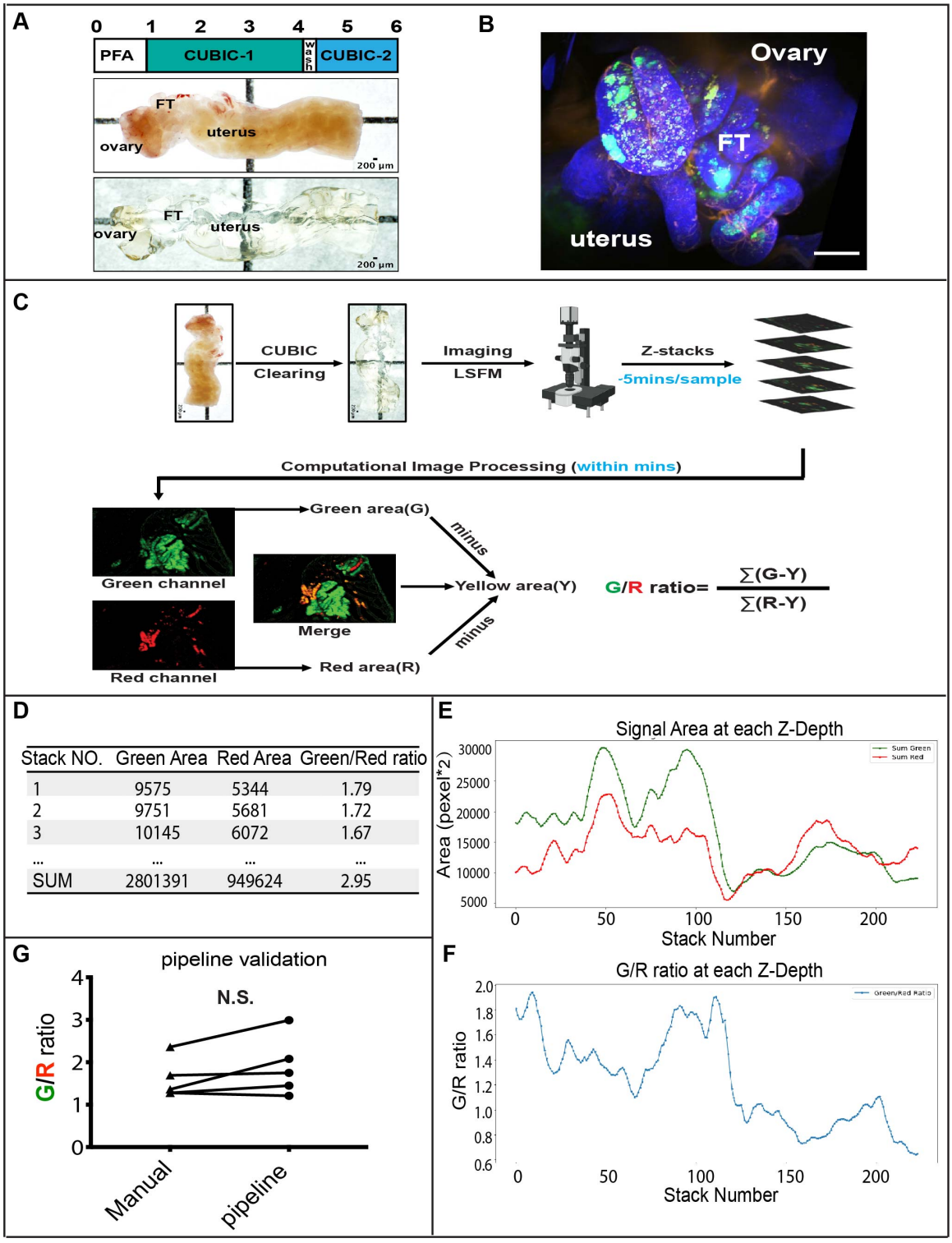


Figure 6. A computational image processing pipeline for analyzing G/R ratio.

(A) Procedure and timeline of tissue clearing of reproductive tracts from 8 weeks wildtype MADM mouse with CUBIC method (upper panel), bright-field images before (middle panel) and after (lower panel) tissue clearing are showed.

(B) Light-sheet imaging of cleared FT from 24 weeks mutant MADM mouse.

(C) Overview of the pipeline. Fixed reproductive tract were cleared though the CUBIC method, then rapid 3D imaging of the whole FTs were performed with LSFM. For the acquired image sequence, green (G) and red(R)signal area were automatically measured, then the double positive yellow area(Y) was calculated. Through subtraction of yellow area, we got true green area (G-Y) and true red area (R-Y). By adding up (G-Y) and (R-Y) of all images and dividing, we get the **G**/**R** ratio number.

(D) Summary of G/R ratio across the image stack.

(E) The curve of green area and red area at each Z-depth.

(F) The curve of G/R ratio at each Z-depth.

(G) Measuring of G/R ratio manually and with pipeline. 5 images from the same image stack (the 1st one of every other 50 images) were selected out for manual quantification of G/R ratio, and then compared with the value from the pipeline. The Wilcoxon matched-pairs signed rank test were used. P=0.18.

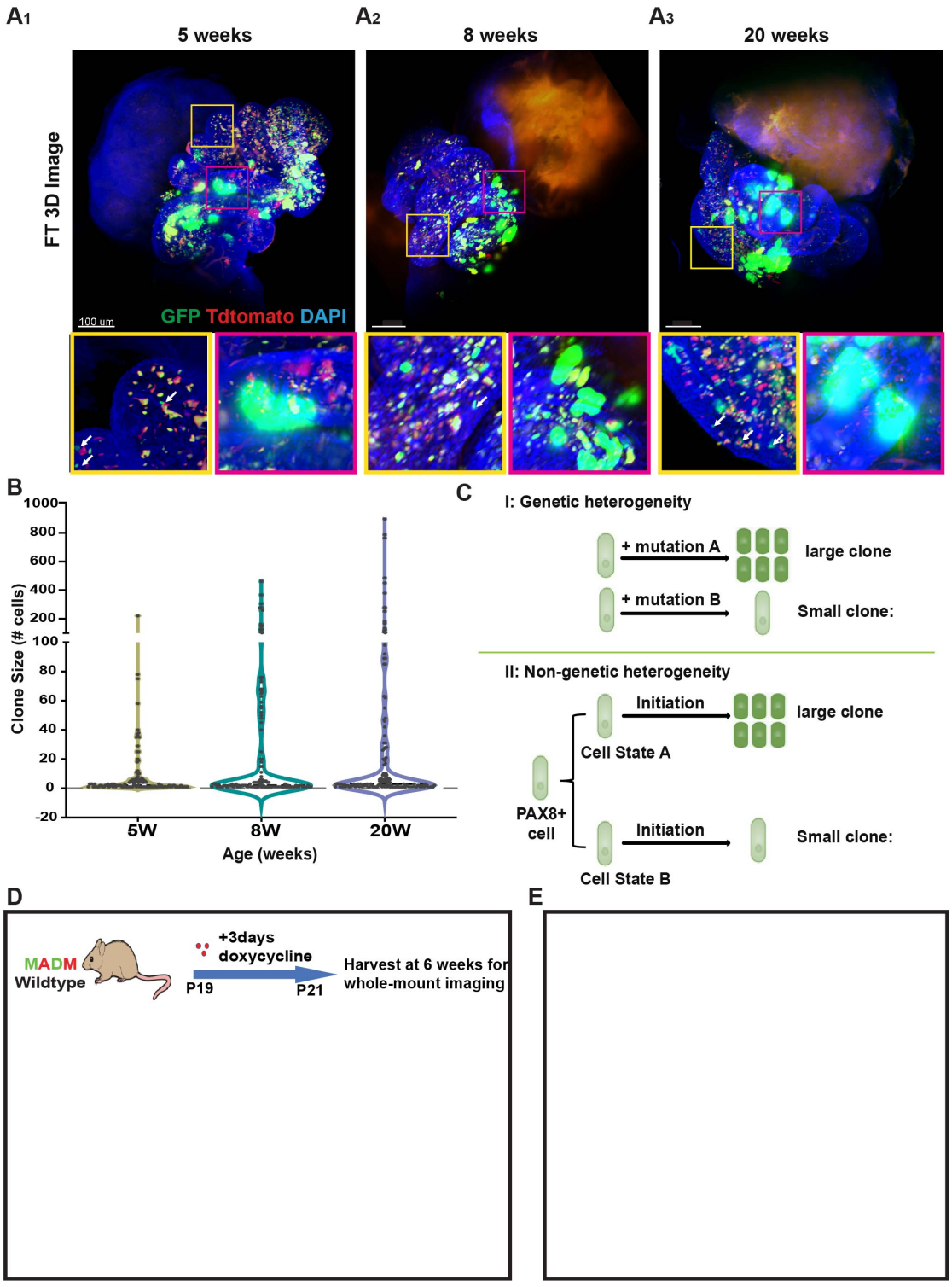


Figure 7. Heterogeneity of PAX8+ cells revealed by clonal analysis.

(A) Representative 3D images of green mutant clones from FT of 5 weeks, 8 weeks and 20 weeks MADM mutant mouse (N=4), yellow box shows single cell mutant clones, magenta box shows expanded mutant cell clones.

(B) Violin-plot of size distribution of mutant clones in 5 weeks, 8 weeks and 20 weeks FT from mutant mice. Clones quantified from 4 mice were plotted together.

(C) Hypothesis that can explain the clonal size heterogeneity.

(D) Whole-mount fluorescence images of fallopian tube from MADM-WT mice with 3 days doxycycline administration between P19-P21, tissues were harvested at 6 weeks old. Yellow box shows unexpanded clones, magenta box shows expanded clones. Representative image from three mice.

(E) The average clonal size of expanded clones from MADM-WT and MADM-Mut mice with 3 days doxycycline administration between P19-P21, tissues were harvested at 6 weeks old. N=3 for each group.

Supplementary figures

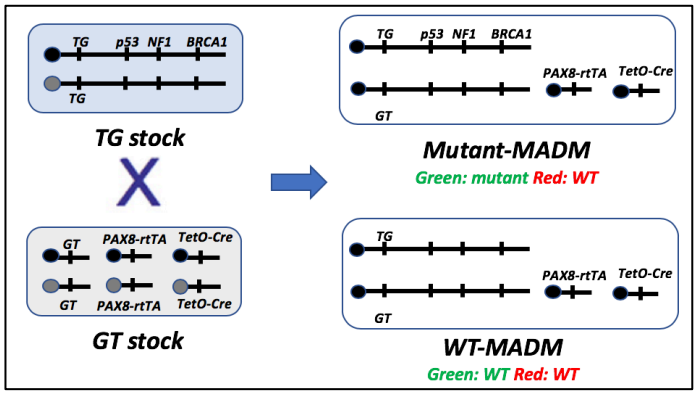


Figure 1: Breeding scheme of MADM-based HGSOC model.
The gene mutations were incorporated into the TG stock and the PAX8-rtTA and TetO-Cre transgene were incorporated into the GT stock, to breed mutant-MADM mice with genotype of *TP53*$^{+/-}$*, BRCA1*$^{+/-}$*, NF1*$^{+/-}$*; PAX8-rtTA, TetO-Cre*.

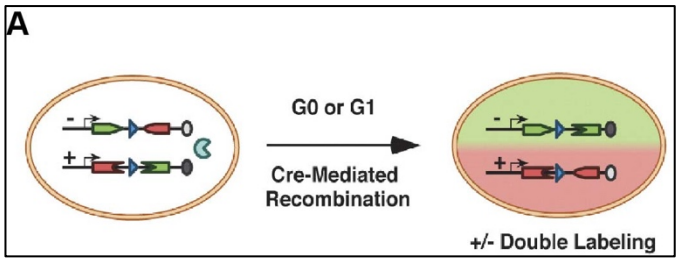


Figure 2: mitosis independent MADM labelling

Recombination occurring in G1 or post- mitotic cells (G0) generates double-colored cells without altering genotype.

Method: G/R-count software

Access to the software package: https://github.com/Mitogenie/GR_Count

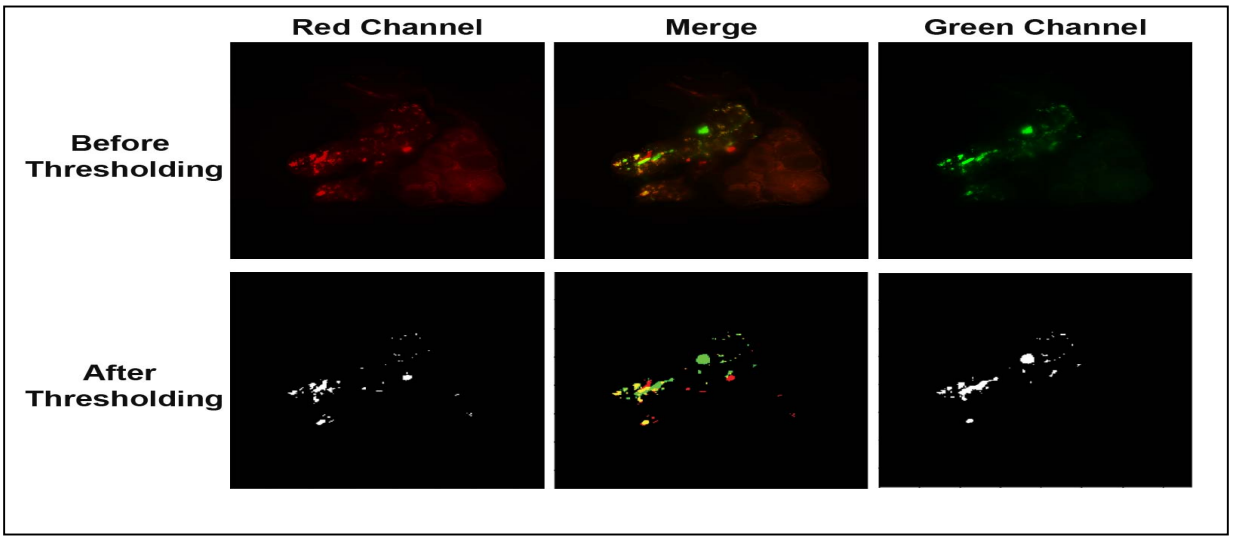
Input： the software uses Tagged Image File Format (.tiff) image stacks from whole-fallopian tube 3D light-sheet imaging. Each image slice contains GFP and RFP channels, thus we will have a GFP-channel image stack and a RFP-channel image stack saved under one parental folder as input.

Code and functional annotations

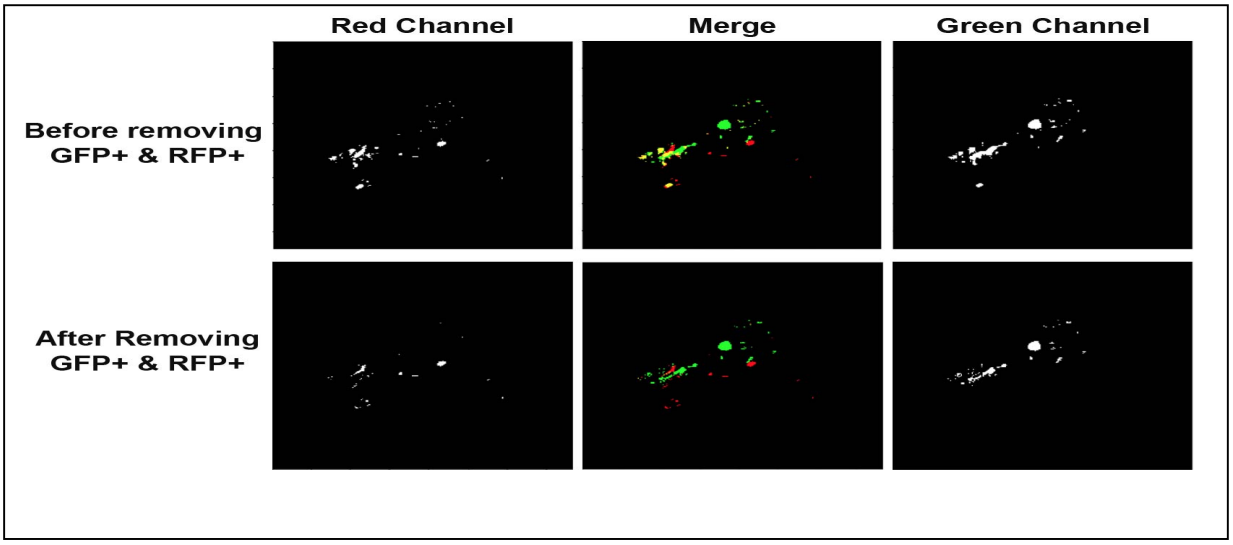
The G/R-cou package is developed in python 3.6, and to make it user friendly, we implemented the interface (UI) in Jupyter notebook. G/R-count offers a semi-graphical UI and provides controls through Jupyter Widgets for interactive data exploration.

Step 1-Read images: Image stacks are exported as separate TIFF. images, automatically named in the following format: "name_of_the_stack_z***_c***.tiff", where the z*** provides the depth info, and c*** indicates the channel info. The G/R-count software reads images stacks and extracts the channel and depth info from the filenames. After that, G/R-count creates an internal link between different channels at each depth.

Step 2-Set threshold to remove background noise: To decide the optimal threshold, we picked out the first one of every other 50 images from the stacks and select a threshold value that removes background for the most without losing the true signals (fig. *S3*). The threshold was then applied to all images. The user can use the provided slider to independently set a threshold for signal levels of either red or green channels, and remove the noise. This step creates a binary mask of the signal zones in each channel.



Step 3- Remove GFP & RFP double positive area to measure single positive area: As the individual green or red channel includes double positive area (yellow) which represent heterozygous cells, to remove that, the package merges the two channels and calculates the double-positive pixel area. A slider for setting the overlap threshold is provided: if the strength ratio of green signal to red signal is above the overlap threshold value, the pixel will be taken as a double positive. The optimal overlap threshold should be the largest number that fully remove yellow pixels, which means no yellow color showed in the lower middle panel of fig. s4. After removing the yellow pixels, the single positive ones are measured for calculating G/R ratio. The package repeats this process for images at each Z-depth.

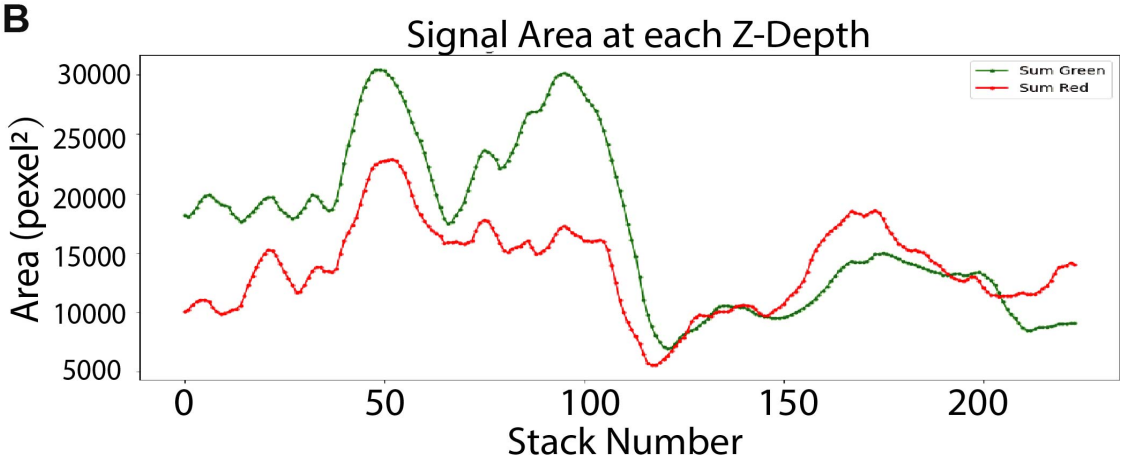


Output The package outputs are 1) a summary of green area and red area in images of each Z-depth, as well as the added up green area, red area, **G**/**R** ratio with a CSV file (fig. *S5*A).; 2) plotted **G**/**R** ratio at each individual Z-depth to reveal the spatial distribution patterns of mutant cells (fig. *S5* B, C).

A

Stack NO.	Green Area	Red Area	Green/Red ratio
1	9575	5344	1.79
2	9751	5681	1.72
3	10145	6072	1.67
...	...	...	...
SUM	2801391	949624	2.95

B



C

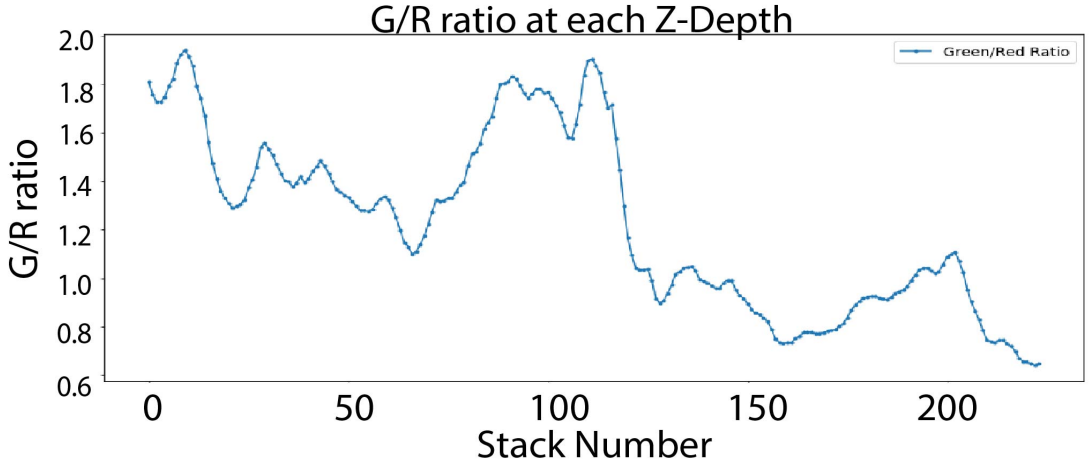


Fig. S5 （A）summary of G/R ratio across the image stack. (B) the green area and red area at each Z-depth.