Winter 12-22-2018

Modeling Human Cancer Therapy Response in Patient Derived Xenografts

Joan Malcolm
University of Maine, joan.malcolm@jax.org

MODELING HUMAN CANCER THERAPY RESPONSE

IN PATIENT DERIVED XENOGRAFTS

By

Joan Malcolm

B.S. University of Maine, 2007

A DISSERTATION

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

(in Biomedical Sciences and Engineering)

The Graduate School

The University of Maine

December 2018

Advisory Committee:

Carol Bult, Professor, The Jackson Laboratory, Advisor

Rosemary Smith, Professor, University of Maine, Co-Chair

Susie Airhart, Senior Director, The Jackson Laboratory

Joel Graber, Senior Staff Scientist, Mount Desert Island Biological Laboratory

Matthew Hibbs, Professor, Trinity College

MODELING HUMAN CANCER THERAPY RESPONSE

IN PATIENT DERIVED XENOGRAFTS

By Joan Malcolm

Dissertation Advisors: Dr. Carol Bult and Dr. Rosemary Smith

An Abstract of the Dissertation Presented
in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy
(in Biomedical Sciences and Engineering)
December 2018

Patient-derived xenografts (PDXs) generated by implanting human tumor tissue into a transplant compliant mouse host have been of increasingly importance to preclinical development and have been demonstrated to have advantages compared to cancer cell lines and cell-line xenografts (CLX) for modeling therapeutic responses in cancer. Nevertheless, many open questions remain regarding the relationship between study design factors and classification of treatment response and the molecular fidelity of tumors passaged in PDXs relative to the original patient tumor(s). The research described in this dissertation addresses both of these significant issues related to the use of PDXs as a tool for modeling human cancer therapy response.

This work investigated how study design factors (enrollment criteria, group size, study duration) and data analysis method influenced treatment responses in PDXs. Each of the methods evaluated consistently classified responsiveness which suggests it is feasible to

combine response data across studies that apply different methods assuming consistent response value thresholds have been used. The results demonstrated that a cohort size of three is sufficient for identifying the four highly responsive and nine highly non-responsive cisplatin treated tumors, suggesting that the use of a low cohort size to screen for chemotherapies that have a high degree of activity or models that are highly responsive is possible.

Mutational and expression profiles of engrafted tumors were compared with the original patient tumors and demonstrated that the molecular fidelity of PT and passaged tumors is high at a global level; the majority of variants with known clinical significance are preserved between PT and PDX tumors. However, individual markers are not invariably concordant and in some cases the molecular discrepancies observed between PDX and PT occurred in variants that are clinically relevant to either disease prognosis or to selection of therapy. Confirming the mutation status of donor tumors used to establish testing cohorts should be routine practice when PDXs are used as models for modeling treatment responses of individual patients.

DEDICATION

To Joshua

My life changed the moment you put your hand on my back.

Thank you for making my reality better than my dreams.

I love you.

ACKNOWLEDGEMENTS

I would first and foremost like to thank Joshua for his unwavering love, support and encouragement. I would also like to thank my family, particularly my sister, for helping to guide me to the person I am today. I would not have made it to the finish life without Dr. Carol Bult and Dr. Rosemary Smith. I am indebted to Carol for her patience, guidance, and insights into both science and life and Rosemary for her faith in my ability and support. To Susie Airhart, thank you for your honesty, clarity, and reality checks. To Joel Graber, thank you for instilling in me that the data leads the way. I cannot thank the many members of my "JAX family" enough for the support, friendships, and once-in-a-lifetime opportunities I experienced over years, particularly Chuck, Fitz, Shannon, Lori, and Rachel.

I am also indebted to The Jackson Laboratory PDX resource for generating and maintaining the PDX models used in this study and to Emily Jocoy and Margaret Bundy for their work to ensure the models meet high quality control standards. Special thanks to Tim Stearns, Grace Stafford, Vivek Philip, Xing Yi Woo and Al Simons for their PDX data crunching support. To the JAX Genomic Medicine Clinical Laboratory, thank you for guiding me through the complexities of genomics. To Honey Reddi, thank you for sharing your drive, strength and friendship with me. To Andrew Hesse, thank you for your thoughtful insights.

TABLE OF CONTENTS

TABLE OF CONTENTS - Continued

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

aCGH	Array Comparative Genomic Hybridization
BL1	Basal-like 1 subtype of Triple-negative breast cancer
BL2	Basal-like 2 subtype of Triple-negative breast cancer
CDX	Cell line-derived xenografts
CF	Cystic fibrosis
CNV	Copy number variation
EGFR	epidermal growth factor receptor
FFPE	Formalin fixed, paraffin embedded
GEMM	Genetically-engineered mouse models
GO	Gene Ontology
GW	GeneWeaver
HER2	human epidermal growth factor receptor 2
HSC	Hematopoietic stem cell
IHC	Immunohistochemistry
IM	immunomodulatory subtype of Triple-negative breast cancer
LAR	luminal androgen receptor subtype of triple-negative breast cancer
M	Mesenchymal subtype of Triple-negative breast cancer
MAF	Minor Allele Frequency
MSL	mesenchymal stem-cell like subtype of Triple-negative breast cancer
NCI	National Cancer Institute
NSCLC	Non-small cell lung cancer
NSG	NOD.Cg-Prkdcscid Il2rgtm1Wjl/SzJ mouse; JAX strain #5557
PBMC	Peripheral blood mononucleocytes
PCT	PDX clinical trial
pCR	Pathological complete response
PDMS	Polydimethylsiloxane
PDX	Patient-derived xenografts
PT	Patient tumor
RTC	Rate-based tumor to control
RTV	Relative tumor volume
SMA	Spinal muscular atrophy
SNV	Single nucleotide variant
TC	Tumor to Control ratio
VLAD	Visual annotation display

CHAPTER 1

INTRODUCTION

Ninety-five percent of clinical trials in oncology fail; over 30% of these failures are attributed to preclinical models that did not accurately predict patient response.[1,2] This highlights two areas of preclinical research that require improvement: 1) preclinical models that accurately reflect patient responses[3] and 2) experimental design that take in to account an understanding of factors that influence the reproducibility and reliability of data generated by preclinical model systems. The research described here focuses on the later in a specific model system: patient-derived xenografts (PDXs). PDXs are generated through the implantation of tumor material from a patient into an immuno-deficient mouse host. If the tumor successfully engrafts, it can be passaged multiple times to generate cohorts of tumor-bearing mice that can then be tested with different therapies (Figure 1-1).

Patient-derived xenografts (PDXs) have been of increasingly importance to preclinical development and have been demonstrated to offer an improvement over traditional cancer cell lines and cell-line xenografts (CLX), which are characterized by poor histological and phenotypic similarities to the original cancer, lack of native tumor microenvironment, and limited correlation of response to patients.[4,5] Patient-derived xenografts (PDXs) have proven to be powerful for investigating the complexities of cancer diseases, with published correlations concerning histological phenotypes, genomic signatures, and treatment responses.[6-10] They are widely used for identifying

molecular targets of compounds,[11,12] predicting patient response and survival,[9,10] and assessing the responses of tumors to compounds at both the individual patient level[13-15] and the population level, as demonstrated in a groundbreaking PDX clinical trial study[16].

Despite numerous documented benefits of PDXs for assessing therapeutic responses in cancer, questions remain regarding the impact of design factors the robustness and reproducibility of treatment responses in PDXs and also about the molecular fidelity of PDXs over different passages relative to patient tumors.[17] Interpreting treatment responses in PDXs and comparing response data across various studies is challenging due to variations in study designs and the lack of rigorous analysis regarding how differences in these elements influence study outcomes. Evaluating the influence of study design factors such as tumor enrollment volume, group size, study duration, and selection of response classification method are all critical to ensuring the reproducibility of results from PDX models. Regarding molecular fidelity, human tumor material used to generate a single PDX model is typically from one cancer lesion at one point in time. Cancer is known to evolve in patients; therefore, a PDX model may not represent the full complexity of the individual disease. Genetic changes have been proven to occur in PDX at rates which are intrinsic to tumor type;[18] this suggests that tumors from PDXs may not have the same genomic configurations as patient tumors.[17,19] In previous studies, the loss of specific amplifications such as *PHGDH*[16] and the disappearance of mutations in genes such as *CDKN2A* and *MDM4* have been reported.[16] The extent of genomic differences and their potential impact on treatment responses in PDXs has not been evaluated extensively to date.

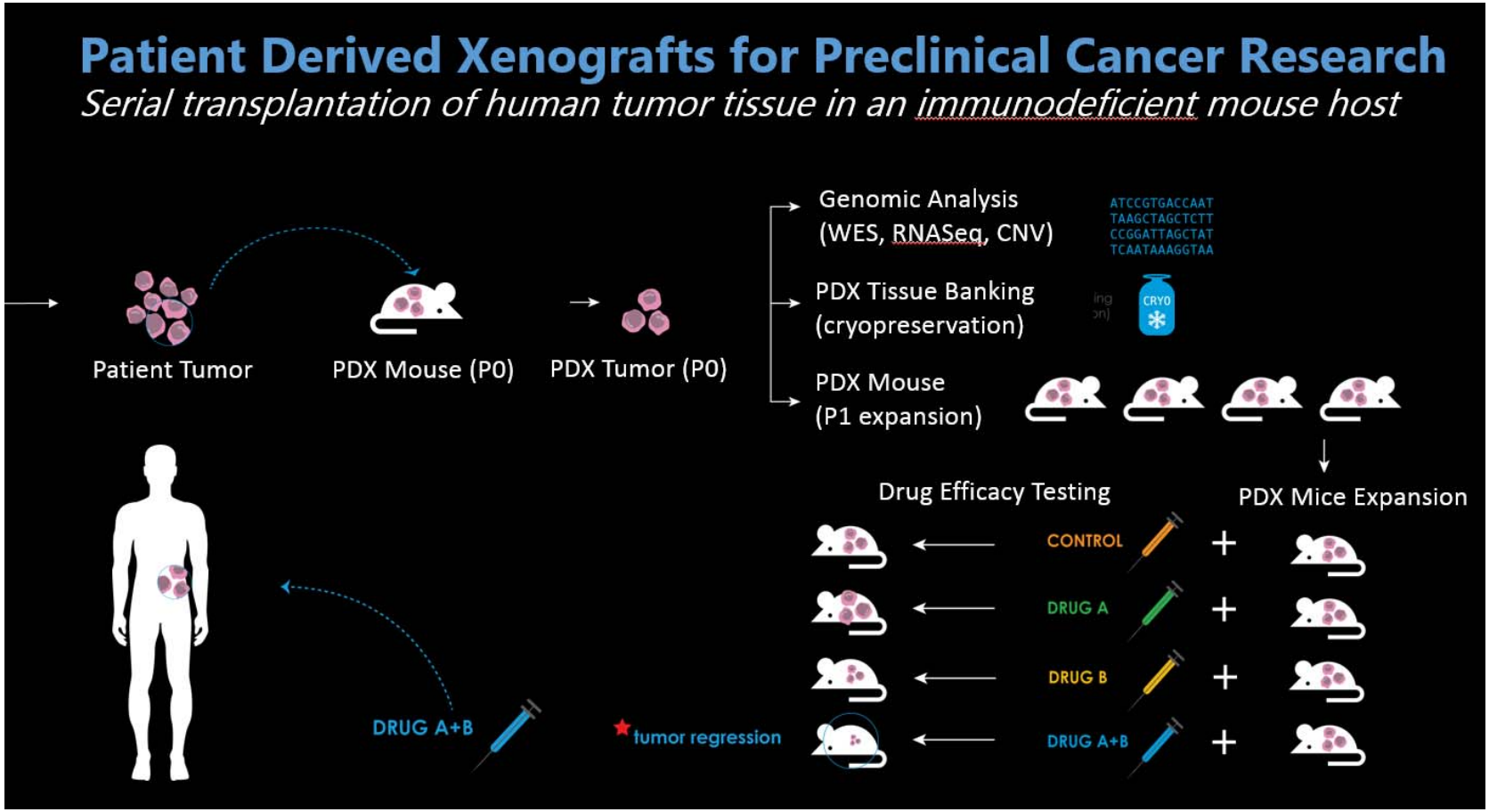

Figure 1-1 Diagram of Patient-Derived Xenograft Process: Patients with a solid cancer tumor undergo biopsy or surgery resection. Tumor specimens are implanted into NSG mice to establish PDXs, with any remaining tissue available for genomic testing or histology. When the first PDXs (Passage 0 or P0) are established, some PDX specimens are implanted to more NSG mice to establish more PDXs (P1) while other specimens are submitted for genomic analysis, including whole exome and transcriptome sequencing. Passaging can be repeated to expand the number of specimens available for testing and then treated grouped into treatment and control groups and treatment with a single compound or combination of compounds. The results of how the tumor responds to treatment can be used to further understand the compound or cancer, or in some cases, used to influence treatment selection for the individual patient. Image source: Dr. Charles Lee, The Jackson Laboratory.

The research described in this dissertation addresses two significant issues related to the use of PDXs as a tool for modeling human cancer therapy response; study design and molecular fidelity. The focus of analyses presented in Chapter 2 is on how different study design factors influence treatment response classification in PDXs. Chapter 3 focuses on molecular fidelity of the PDX model system by comparing the genomic characterization of engrafted tumors with the original patient tumors. The remainder of this Chapter provides background on model systems for modeling cancer therapy.

Growth of Oncology Drug Development and Precision Oncology

According to the American Cancer Society (ACS), two out of every five people will be diagnosed with cancer during their lifetimes.[20] While the death rate from cancer has declined steadily over the past two decades, the incidence rate has increased. Over 16 million people were living with cancer in 2016 in the United States alone, and this number is expected to increase to 19 million by 2024.[21] Along with this rise, US expenditures for cancer care will reach $165 billion per year by 2020. The cancer-drug pipeline and the number of new therapies being developed has also expanded; more than $50 billion is now invested in this annually.[22] The number of active compounds in oncology research and development has quadrupled since 1996 and nearly doubled between 2008 and 2016[16,22] to 836 cancer drugs and vaccines either in clinical trials or under review by the FDA. The landscape of new therapies is also evolving: Approximately 80% of these new compounds are considered first-in-class therapies, and

three quarters are classified as “personalized” medicine because they target a specific genomic aspect of the tumor.[23]

Within the drug development pipeline, there is an intensifying focus on reducing the attrition rates of compounds that have entered clinical trials. Retrospective studies have demonstrated a 5% to 20% success rate of cancer treatments entering clinical trials; over 30% of attrition is due to a lack of efficacy in late-stage clinical trials.[1,24,25] Providing preclinical models that mimic patient response and allow for early, reproducible efficacy evaluation of novel therapies continues to be a bottleneck in relation to accelerating new developments; this accounts for a significant portion of the attrition.[3] Recent improvements in the availability of preclinical models have reflected the various subtypes and complex profiles of cancer, but continued improvements of these models in terms of their ability to efficiently and reproducibly predict the response of cancer to a drug are necessary.

Over the past 100 years, treatments for cancers have advanced immensely; the only available option of was surgery in the 1900s, while radiation, hormone therapies, immune therapies, and combination treatments are also used today. Chemotherapies emerged in the 1950s and expanded through the screening of compounds to discover those that kill rapidly-dividing cells. In the 1970s, links to molecular genetics underlying the basis of cancer were first discovered. During this period, the first two critical families of genes that, if mutated, could precipitate this overgrowth were identified: oncogenes and tumor suppressor genes.[26] Molecular profiling of tumors revealed that

many cancers defined primarily by histology are, in fact, comprised of molecularly distinct subtypes. The third category of cancer-causing alterations was uncovered in early 2000 with the discovery of microRNAs (miRNAs). Alterations in these genes are most frequently somatic events, although germ-line mutations occur and contribute to roughly 15% of all cancers.[26,27] Understanding these molecular alterations and the biological mechanisms involved has produced new sub-classifications of cancers that can be studied differently in terms of prognosis and response to treatment.

Genetic and genomic characterization of cancers has ushered in an era of "precision medicine" where treatment options are driven by genomic properties of an individual patient's tumor. Precision medicine has the potential to revolutionize the practice of medicine by enabling healthcare providers to match treatment and prevention strategies to a patients' individual profiles by combining "-omics" signatures, microbiome composition, health history, environment, and lifestyle.[28,29] Throughout the past two decades, genetics has proven to be a key contributor and clear guidepost to some patient prognostic, diagnostic, and theranostic complexities, which indicates that distinct genetic variants can cause conditions with ostensibly identical symptoms but with drastically different responses to treatments.

Our enhanced understanding of the genetics and genomics that drive cancer has also yielded a new class of targeted therapies which can block essential biochemical pathways or mutant proteins of the tumor cell, resulting in tumor regression or arrest.[30] However, these targeted therapies are frequently associated with cancer recurrences,

which has been caused largely by the emergence of drug-resistant variants;[30] therefore, while targeted therapies offer promise, their limitations continue to indicate cytotoxic chemotherapeutics as the backbone of cancer treatment.

Molecular alterations in tumors may also be correlated with the prognosis of the disease. OncotypeDX is one example of a test that uses an assessment of 21 genes to estimate recurrence risk in breast cancer. The test is both prognostic and predictive and is used both to estimate risk of recurrence and to predict which patients will benefit from chemotherapy and which will benefit from hormone therapy only.[31,32] OncotypeDX has been incorporated into the guidelines of the National Comprehensive Cancer Center Network (NCCN) and the American Society of Clinical Oncology (ASCO) and widely adopted to make treatment decisions.[33] These discoveries, along with a rapid penetration of next-generation sequencing (NGS) technologies, have allowed drug development and clinician treatment strategies to be influenced by individuals' cancers' genomic profiles.

Models of Human Cancer Used in Drug Development

Models of human cancer are used throughout early and late-stage preclinical drug development and can also be used to further expand the clinical use of approved drugs. In early stages of development, they can be used to narrow libraries of hundreds of compounds down to several that will advance to late-stage development; in late stages, they can be used to pinpoint molecular, histological, or phenotypic signatures of tumors that respond to the drug. This can therefore more precisely focus the design of the

clinical trial on a specific patient population. However, experimental models designed to assimilate known disease complexity have too often proven ineffective, only to highlight the limitations of our common understanding of the disease or the model system. Table 1-1 displays commonly used models in oncology preclinical development pipelines. In many cases, results are combined from models such as cell lines, genetically engineered mice, and PDX, since cumulative information can help to understand the subtleties of cancer formation and response to therapy in greater detail. Complexities may include tumor microenvironment influence,[34] intra- and inter-tumor molecular and biological heterogeneity,[35] systemic and tumor immune[36] and metabolic response,[37] and the ability of drug-resistant stem-like cancer-initiating cells to repopulate treated cancers.[38]

In vitro models

Several *in vitro* systems have been optimized for high-throughput efficacy screening of compounds, thereby allowing entire libraries of potential pharmacologically relevant compounds to be screened for different types of cell signals which are relevant to tissue damage or therapeutic goals. These models consist of cell lines, spheroids, and organoids. Cell lines have been used in *in vitro* screening for over four decades; the most commonly used lines can be found in a resource maintained by the NCI called the NCI-60, a panel of 60 human cancer cell lines grown in culture. Since 1990, industry and academia have screened over 100,000 compounds using the NCI-60.[39]

Cultured cell lines produce a relatively homogenous population of cells; while it has historically been the most broadly used *in vitro* cancer model, it presents known

limitations and, therefore, alternatives have been developed. For example, tumor spheroids are a non-clonal 3D structure of tumor cells grown in culture (derived from a mixture of cells); they have been used since the late 1970s. Because they are comprised of different cell types from a tumor, they are designed to represent some of the intrinsic heterogeneity observed in solid tumors. Tumor organoids, in contrast with spheroids, are a clonal 3D structure of tumor cells grown in culture (derived from a single progenitor cell); they are more difficult to culture because they develop from a disaggregated single cell, but they allow the opportunity to study 3D structures of genetically identical cells.

These models are used in early-stage, high-throughput screening of compounds and have been shown to be highly reproducible and time- and cost-effective. However, as a result of limited histological and phenotypic similarities with primary cancer, lack of a native tumor microenvironment, and limited correlation of response to patients, more advanced models are utilized for later stages of compound development to confirm findings.[40] An NCI announcement in the Spring of 2016 provided evidence of this shift away from cell lines, which stated that it would be retiring the NCI-60 repository and replacing it with a repository of PDX models, curated with relevant clinical information of the originating cancer.[41]

Lab on a chip technology

Lab-on-a-chip devices are designed to integrate multiple function on a single integrated chip. With sizes typically ranging from millimeters to a few square

centimeters, they aim to enable automation and high-throughput screening, handling small fluid volumes down to less than pico-liters. Preclinical models that utilize human cancer tissue for testing often require that the tissue is amplified either in a biological system, such as by expanding human tumors in a mouse, or in an *in vitro* culture. While a degree of genetic stability is evident, particularly in aggressive driver gene mutations, there are concerns of genetic and phenotypic fidelity of amplified tumors.[10,42] Drug screening methods for primary cancer tissue remain a focus among researchers who are developing lab-on-a-chip technologies. To circumvent the sample input problem, drug screening platforms based on nano- and micro- volume fluidic channels have increasingly been explored throughout the past decade.

One system developed by Wong et al. uses as few as 16,000 cells obtained within 24 hours after tumor resection from cancer patients. This provides the potential for a powerful tool for rapid, low-input drug screening of primary cancer tissue.[43] Adaptation of the assay to cancer cell lines suggests its application in potentially all types of cancers. It enables the opportunity to observe and quantify cellular drug response on the single-cell level and to achieve population analysis through statistical analysis of multiple droplets. While this technology offers promise for a future tool for rapid, low-cost screening, it is an emerging technology with limited data regarding the reproducibility and predictiveness in the clinic.

Genetically-engineered mouse models

Genetically-engineered mouse models (GEMMs) are generated through the introduction of targeted genetic mutations associated with specific human malignancies. Mutations may be either gain-of-function oncogenes or loss-of-function tumor suppressor alleles that are constitutively or conditionally expressed in mouse models.

To date, GEMMs have been developed for many common tumor types including lung, prostate, breast, colon, and pancreatic cancers, which have been reviewed previously.[44] Several advantages are offered by GEMMs; for instance, the murine tumors are exposed to a functional immune system and intact tumor microenvironment, specific genetic abnormalities that are present in human tumors can be induced at specific ages in the tissue-type of origin, and the stages of tumor progression can be studied over time, which allows several therapeutic approaches to be explored at various stages of tumor development. Genetically-engineered mouse models (GEMMs) provide the only opportunity to evaluate drug delivery, therapeutic response, and biomarker expression for cancers evolving within their natural microenvironment.

However, not all GEMMs are designed and created with the same technical standards, and various criteria and guidelines have recently been suggested; these call for more precise gene mutations in GEMMs when used in human cancer modeling.

Table 1-1 Characteristics and applications of preclinical cancer models used to assess therapeutic responsiveness

Model	Features	Utility	Strength	Caveats	Clinical Predictability
in vitro cell lines, spheroids, organoids	Culture of human cancer tissue	-Early-stage discovery screening -Rapid, high-throughput screening of novel compounds for anti-tumor efficacy	-Reproducible -Cost- and time-effective	-Limited histological and phenotypic similarities with primary cancer -Loss of tumor heterogeneity -Lack of native tumor microenvironment	-Limited predictiveness for most cancers
Transgenic and GEMMs syngeneic	Spontaneous, autochthonous development of tumors in immuno-competent rodent	-Mid- or late-stage discovery screening -Evaluates effects on immune surveillance, invasion, tumor microenvironment, tumor angiogenesis	-Phenotypic, histological, and genetic similarities to primary cancers -Can evaluate effect on tumor microenvironment, immune system	-Asynchronous tumor development -Can be outbred mice with non-uniform genetic background Lack of tissue-specific promoters -Imaging modalities necessary for assessment of tumor development	-Predictive in therapeutics' responsiveness and resistance
Cancer cell line-derived xenografts (CDX)	Human cell line engrafted in immuno-deficient rodent	-Early- or mid-stage discovery screening -Rapid screening of compounds for anti-tumor efficacy	-Reproducible -Applicable to many tumor types	-Limited histological and phenotypic similarities to primary cancer -Loss of tumor heterogeneity -Lack of native tumor microenvironment	-Limited predictiveness for most cancers with noted exceptions (i.e., oncogenic driver mutations)
Patient-derived xenografts (PDX)	Human-derived, low-passage tumor engrafted in immuno-deficient rodent	-Mid- or late-stage discovery screening -Evaluates novel compounds and therapy strategies on human-derived tumors	-Preserves genetic, histological, and phenotypic features and heterogeneity of patient tumor -Maintains stroma and tumor microenvironment (although not native) -Assesses novel biomarker signatures of therapy response	-Access to PDX models can be limited -High front-end costs and labor-intensive preparation -Limited engraftment rates and long time to model development	-Shown to be predictive for individual cases -Can be used for "clinical trials in a mouse"
Lab-on-a-chip	Micro- and nano-volume fluidic channels requiring low input of human cancer cells	-Early-stage discovery screening -High-throughput screening of novel compounds for anti-tumor efficacy	-Requires very low input of tumor material, allowing for numerous compounds to be tested without needing to amplify (culture, passage, etc.) tumor material	-Typically uses dissociated cells with lack of native tumor microenvironment -Emerging technologies with limited data	-Poorly understood

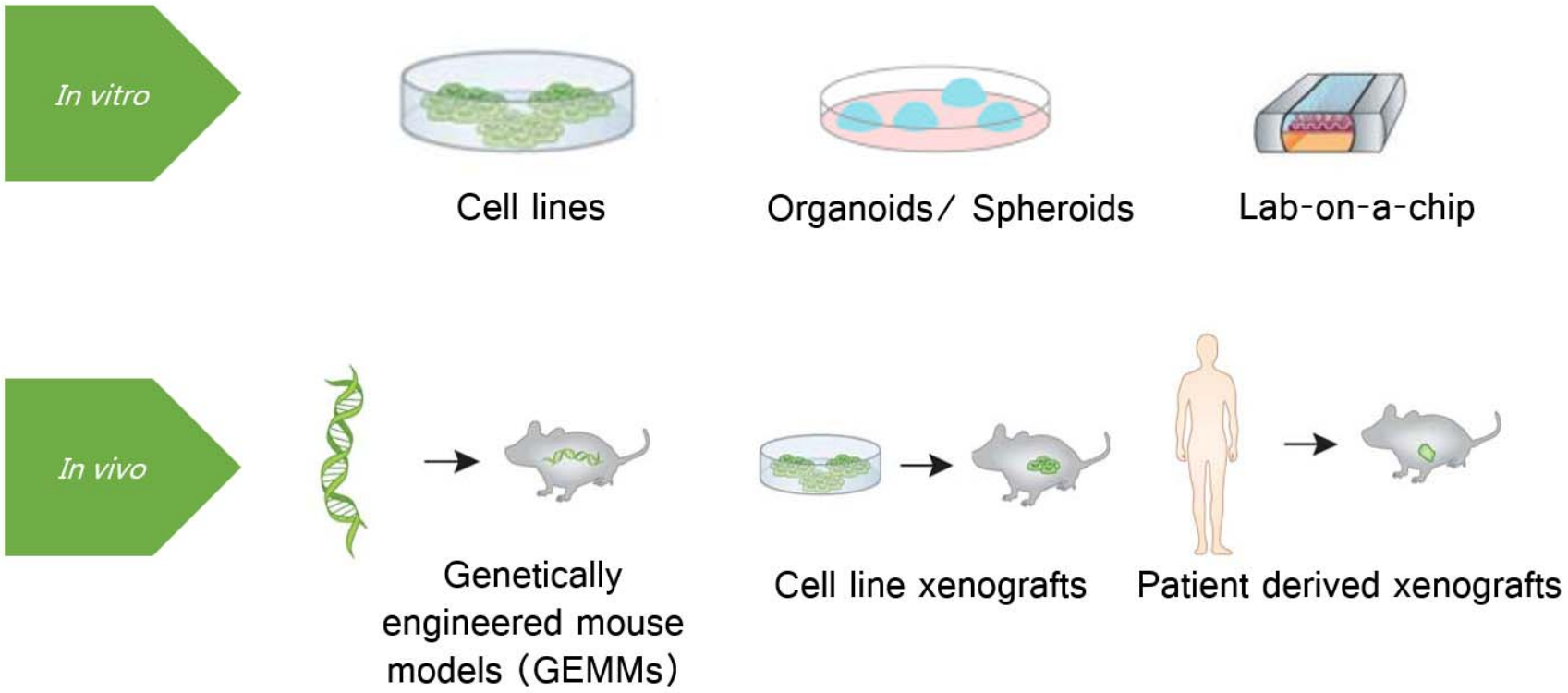

Figure 1-2 Models of human therapy response. Adapted from Nature Biotechnology 30, 648–657 (2012)

For example, (1) mice must carry the same mutation that occurs in human tumors; (2) mutations should be engineered within the endogenous locus and not expressed as a transgene; (3) mutated genes should be silent during embryogenesis and early postnatal development (except when studying inherited pediatric tumors); and (4) mutations should be within the specific target tissues in selected cell types.[24] The precision required for the creation of GEMMs can be complex and technically challenging and may involve advanced genetic-engineering techniques that can result in "off-target" effects of gene editing with unintended consequences[45]. The time to create a new model that contains the genetics of interest can require up to a year of gene editing and subsequent breeding. Additionally, crucial studies must be performed to determine whether GEMMs are more effective than xenograft models in predicting the efficacy of cancer therapies.

Patient-Derived Xenograft Model System: History and Applications

The use of PDXs has been explored since the 1960s but more rapidly expanding in the early 1980s, when immunologically-compromised mice that could support human tumor growth became widely available.[46] Patient-derived xenografts have been addressed in many reviews [8,12,18,47] and have proven to be a powerful model to investigate complexities of cancer diseases; correlations related to histological phenotypes, genomic signatures, and treatment responses have been published.[6-10] Some of the advantages of PDXs include the following: (1) ease of model generation; (2) therapeutic assessment occurring in human-derived tumor tissue with the complex

genetic and epigenetic abnormalities in the human tumor are intact; (3) multiple therapies can be tested from a single tumor biopsy; (4) molecular sequencing data can be readily obtained from the xenograft tissue, before and after drug therapy; (5) stroma is present, although human stroma is replaced by mouse stroma as the model develops; and (6) identification of individualized molecular therapeutic approaches[48,49] is possible through the representation of a wider range of cancers.

Limitations of PDX include the relative cost and long lead time of utilizing the model, relative to cell-line models. The time required to passage PDX model system to develop study animals with tumors at a treatable volume can take up to a year for some slower growing models. While PDX are a powerful tool some researchers, particularly those under fast-paced drug development timelines, avoid using PDXs because of the slow growth rate and longer study duration time[47] compared to other available models such as cell-lines or syngeneic GEMM tumors.. PDX are also heterogenous tumor fragments which provides a key advantage to the model, but how the tumors are divided and engrafted can also incorporate sampling bias across animal cohorts on study. There is also evidence for a lack of molecular fidelity of tumors from PDX models as compared to the original patient tumor(s). This has been observed through a study of the aggregation of data across multiple datasets,[7] monitoring the genomic signatures of a PDX over multiple passages,[17] and through an evaluation of a bank of single tumor-type.[50] The full extent of molecular changes in PDX tumors and their impact on the utility of these models to reflect patient responses to drug treatment is still an open question.

While PDXs have been widely adopted for preclinical studies, standardized methodologies to support reproducibility and common treatment response readouts have not been widely adopted across the research community. In a survey of 50 published PDX studies that reported drug response data there was a wide range in study designs across experiments reported including volume of tumor to treatment enrollment (range: 50-400mm^3), duration of response monitored (range: 12-42 days), size of treatment and control groups (range: 3-20 tumors), and highly variable methods used to classify whether and to what degree a tumor responded to a therapy. The diversity of study designs used for PDX studies limits the ability to compare findings and aggregate data across different laboratories and consortia which, in turn, represents a major challenge to the success of multi-institutional initiatives aimed a preclinical development of PDXs such as NCI's PDX Net and the EurOPDX Consortium.[51]

CHAPTER 2

STUDY DESIGN IN PATIENT-DERIVED XENOGRAFT MODELS

The number of active compounds in oncology has quadrupled since 1996 and nearly doubled between 2008 and 2016[22] to over 800 cancer drugs and vaccines in clinical trials or under review by the FDA. While growth in the number of compounds is key to improving therapy options for patients, retrospective studies have demonstrated a meager 5% to 20% success rate of cancer treatments entering clinical trials;[1,2,24] the remaining 80% to 95% of those compounds, which involved investment of significant resources to advance to a clinical trial, do not achieve approval or provide benefit to patients. Over 30% of the attrition has been attributed to a lack of efficacy uncovered in late-stage clinical trials.[2] Preclinical models that reproducibly and accurately reflect drug efficacy and predict patient responses are also needed to eliminate ineffective compounds earlier in the development process.[3]

In addition to better preclinical models, standardization of study designs and robust endpoints for determining drug efficacy are also important for reproducibility and accurate interpretation of results. Guidelines such as ARRIVE (Animal Research: Reporting of In Vivo Experiments), were developed to improve the design, analysis and reporting of research using animals, however, these guidelines do not address specifically the issues surrounding the use of PDXs for preclinical oncology applications.[52]

For PDXs to deliver on their promise as an effective platform for preclinical studies for cancer therapeutics, it is important to understand how different study design factors influence the reproducibility of treatment response classifications in this model system. Different methods of classifying treatment responses in xenografts have been proposed, but many have been evaluated only in cell lines and CDXs which lack the heterogeneity and subsequent growth variability supported by PDXs.[53-56] For example, Hather et al. compared a tumor growth rate based method (RTC) to a traditional tumor volume treatment to control method (TC) using 219 CDXs and reported that RTC required fewer animals to achieve the same effectiveness as TC.[56] They also determined that a cohort size of four animals achieved over 95% statistical power. To date, there have been no published studies demonstrating the relevance of the Hather study to PDXs.

Beyond the standard cohort design (multiple tumor-bearing animals from the same model per treatment arm), the use of PDX for population-level drug response assessment to mimic a clinical trial setting has grown in use during the past five years. This approach was introduced by Migliardi et al.[57] to indicate a response to a targeted therapy but is best known because of a large PDX clinical trial reported by Novartis.[16] In the Novartis study drug efficiacy was determined using a clinical trial like study design: one mouse, one drug for many models (also known as a PDX clinical trial, PCT).

In the study reported here we evaluated the impact of initial tumor volume, rate of tumor growth, cohort size, study duration, and analysis method on treatment

response classifications in PDX models from The Jackson Laboratory (JAX) PDX Resource. The dataset used for this study consisted of cisplatin treatment response data for 70 PDX models representing ten cancer types with up to 28-day study duration using cohorts of 3-10 tumor-bearing mice. The analysis was repeated using docetaxel treatment response data for 40 PDX models.

Methods

PDX Models

The data for this study were obtained from the JAX PDX Resource, under protocol 12027 approved by The Jackson Laboratory Institutional Animal Care and Use Committee (IACUC) before study initiation. All animal studies followed the IACUC guidelines. The models were generated using subcutaneous tumor implantation in female NOD.Cg-Prkdc^{scid} Il2rg^{tm1Wjl}/SzJ (NSG; JAX strain 005557) host mice. Detailed methods for model generation are described elsewhere [58]. Information and data for PDX models from the JAX PDX Resource are publicly available from the PDX Portal hosted by Mouse Tumor Biology Database (MTB; http://tumor.informatics.jax.org/mtbwi.pdxSearch.do).[59]

PDX Dosing Study Design

To assess the impact of study design factors on treatment response classification to chemotherapy agents, we analyzed drug response data from 70 cisplatin-treated PDX

models representing ten cancer types (appendix A-1) and 40 docetaxel-treated PDX models representing ten cancer types (appendix A-2) from the JAX PDX Resource.

Cisplatin and docetaxel are commonly used chemotherapeutic drugs used for treatment of numerous human cancers including bladder, head and neck, lung, ovarian, and testicular. Cisplatin and docetaxel interfere with DNA replication and repair mechanisms leading to apoptosis in dividing cells, but they do so via different mechanisms. Cisplatin is an alkylating agent and docetaxel is a taxane that interferes with microtubules.

For dosing studies, human tumor fragments (3-5 mm^3) were implanted subcutaneously using a trocar into the right hind flank of NSG mice. Animals were monitored until the engrafted tumor reached approximately 200mm^3. Mice were then randomly categorized into study groups and treated with either 2.0 mg/kg IV cisplatin once per week for up to three weeks, 10.0 mg/kg IV docetaxel once per week for up to three weeks, or a vehicle control. The compound and dosing schedule for vehicle controls varied across studies because of other treatment groups that were run simultaneously with cisplatin or docetaxel. Most models (66) used 5.0 ml/kg IV 5% dextrose in water; the remaining used one of the following alternatives: 1) 10.0 ml/kg PO 0.5% CMC daily (3 models), 2) 5.0 ml/kg PO 2.5% DMSO in PBS daily (1 model), 3) 5.0 ml/kg IV combination of 5% dextrose in water and 0.5% CMC (1 model).

Clinical observations, body weights, and digital caliper tumor volume measurements were performed at least twice weekly for up to 28 days after study

enrollment. Volume was calculated by using the following formula: volume = (length) x $(width)^2$ / 2, in which the length is the larger of the two perpendicular axes and width is the smaller of two perpendicular tumor axes. Following the IACUC protocol, animals were removed from the study if tumors became visibly necrotic or if the animals demonstrated greater than 20 percent weight loss.

PDX Dosing Study Data

All models involved one treatment group and one control group evaluated for a minimum of 14 days, with 67 models enrolled for a minimum of three weeks and 53 for four weeks. Initial individual tumor volume ranged from 50 to 367mm^3 across all studies. The number of mice per cohort ranged from N=3 to N=11; 55 of the models had a treatment group with a minimum of N=8. While the response was analyzed at four time points, the animal and tumor volume monitoring occurred two to three times per week for the duration of the study. Not all measurements used to calculate treatment responses at the end of a dosing study were available precisely on days 0, 14, 21, and 28. If the timepoint did not have corresponding data, then the value that fell within the following range was used: day -1 to 1 grouped as day 0 or baseline volumes, 11 to 17 as day 14, 18 to 24 as day 21, and 25 to 31 as day 28. If multiple points fell within that range, then the closest day to 14, 21, or 28 was used. The study day of the control group follow-up was matched to the study day of the corresponding treatment group follow-up. Outliers were assessed using JMP v13.2.1; minor outliers were defined as 1.5 times the mean volume interquartile range, and major outliers were defined as 3 times

the interquartile range, for both the treated and control groups on a specific measurement day. Outliers were included in all subsequent analyses, as there was no indication that they were the result of a technical error.

Calculations of Treatment Response

Table 2-1 lists the data analysis methods for treatment response that were evaluated in this study. For the study reported here, treatment response was assessed at three time points (14, 21, and 28 days, as available). Each of these methods incorporates a calculation of the ratio of treatment to control measures. The most frequently used method is based on the ratio between the average volume of the treatment group (V_T) and the average volume of the control group (V_C) on a specific day during the study, referred to as the tumor-to-control method (TC). The relative tumor volume (RTV)[60-62,] measure accounts for variation in initial volume of a tumor. Either the mean or median of the RTV of the treatment group (RTV_T) is divided by the mean/median of the control group (RTV_c). Neither the TC nor the RTV accounts for the rate of growth or measurements of volume beyond the final day of the study and the first day of the study. Hather et al. addressed this potential shortcoming by developing a rate-based method (RTC) that accounts for all the individual volume measurements throughout a study.[56]

Based on the tumor volume ratio of treated and controls calculated using the three methods described above, a response classification of “highly responsive” (ratio less than 0.1), “responsive” (ratio between 0.1 and 0.42), or “nonresponsive” (ratio

greater than 0.42) was assigned. These criteria were adapted from the Drug Evaluation Branch of Cancer Treatment, NCI, which defines a T/C <=0.42 as an active drug and T/C <=0.10 as a highly active drug.[63-65] The change in tumor volume was used as an indicator of effectiveness of the drug and as an indicator of drug activity. Figure 2-1 illustrates the distribution of T/C response at day 28 across 53 models treated with single-agent cisplatin. Of 53 models with a study length of 28 days, four PDX models were highly responsive to cisplatin, 11 were responsive, and 38 were nonresponsive.

To evaluate the impact of the number of replicates used to assign a response classification, we applied a bootstrapping[66] technique, which randomly resamples the data with replacement to estimate parameters of a population empirically. By selecting subsets of both the treatment and control groups and using these to compute the response values of each of the three methods compared, we aimed to demonstrate how changes in study length and number of animals influenced the precision of the response values and classifications.

Utilizing only those models with the number of replicates per group (N) of greater than or equal to 8 for the duration of the study, 1,000 bootstrap replicates were performed for each subset from N=3 to N=7 relative to N=all, in which “all” refers to the maximum number of subjects available (8-10). The output determined the probability that the response of the subset N differed by: (1) >|0.05| of the N=all response value; (2) >|0.10| of the N=all response value; (3) change from N=all response classification.

This method allows for replication of realistic sub-samples that may have been encountered if a smaller subset had been used in the study.

Table 2-1 Methods used to calculate PDX response values; V_T is average volume of treatment cohort, V_C is average volume of control tumors, V_x is final individual tumor volume of study, V_0 is initial individual tumor volume of study, u_T is average rate of treatment growth of treatment cohort, u_c is average rate of control growth of treatment cohort, V_t is average tumor volume at end of study.

Method	Abbreviation	Equation
Traditional T/C	(TC)	$\frac{V_T}{V_C}$
Relative tumor volume T/C	(RTV)	$\frac{\left({}^{V_x}/_{V_0}\right)_T}{\left({}^{V_x}/_{V_0}\right)_c}$
Rate-based T/C	(RTC)	$10^{(\mu_T - \mu_c)\ x\ t}$
PDX Clinical Trial	(Gao Method)	$\Delta Vol_t = (\frac{V_t - V_{initial}}{V_{initial}})$

Table 2-2 Thresholds for classifying tumor responses to therapy based on response values

Individual-Level Drug Response				
		Highly Responsive	**Responsive**	**Nonresponsive**
Traditional T/C				
Relative tumor volume T/C		<0.1	0.1-0.42	>0.42
Rate-based T/C				
Population-Level Drug Response				
PDX Clinical Trial	**Complete Response**	**Partial Response**	**Stable Disease**	**Progressive Disease**
Best Response	<-0.95	<-0.50	<0.35	not otherwise categorized
Best Average Response	<-0.40	<-0.20	<0.30	

Statistical power analyses were run using the pwr package (version 1.2-2) in R (version 3.3.1). Effect sizes were determined based on the TC/RTC/RTV (mean) /RTV (median) value using all available samples that that were nonresponsive (value of 0.42 or greater), responsive (value of 0.10 to 0.42) or highly responsive (value of less than 0.10). One thousand re-samplings were used to calculate new TC/RTC/RTV (mean) /RTV (median) values based on a subset of the original samples with replacement. The effect size was divided by the standard deviation of the resampled values. The sample size was determined as the number of subset samples used during the re-sampling process. An alpha value of 0.05 was established. Treatment 2-2 outlines the thresholds for the classifications of response that was based on the calculated TC/RTC/RTV (mean)/RTV (median) values, 0.42 and 0.10 for both directions (greater and less than) individually. The population-level Gao assessment response classifications were based on the calculated response value of the best response at a given time in the study as well as the response over the average of the study.

Results

Broad distribution of model type and cisplatin response across dataset

Figure 2-1 illustrates the distribution of T/C response at day 28 across 53 models treated with single-agent cisplatin. Of 53 models with study lengths of 28 days, four PDX models were highly responsive to cisplatin, 11 were responsive, and 38 were nonresponsive.

Growth rate may impact treatment response classification more than initial tumor volume

Spearman correlation coefficients of other variables with treatment response values were assessed. We observed that tumor volume of control samples at the end of the study had a moderate correlation value (R= 0.48) with response values across all methods which was most pronounced with the RTC data analysis. Figure 2-2 plots the control tumor volumes from study Day 21 of the cisplatin dosing studies. The average volume of the control tumors of the models that were classified as responsive were larger than those that were classified as nonresponsive indicating that the tumors that grew larger when unchallenged also were to be the ones that had a greater degree of response to the compound ($R = 0.36$, $p = 0.01$). The growth rate of the control tumors over the study duration was also moderately correlated ($R = 0.35$) to the response values. No other strong correlations ($R > |0.30|$) between the variables were identified.

Calculating treatment response from mean tumor volume was more reproducible than from the median tumor volume

Figure 2-3 shows the comparison between response values that were calculated from the mean and median tumor volumes across the study groups. The power to detect the response (<0.42) was higher for mean compared to median for all three methods; Table 2-3 shows the power in detecting response versus the number of animals per group for RTV. Ten percent of the models assessed (5 of 50) had a change in the response classification when calculating response with mean as opposed to

median. Of the five models that differ in classifications, the mean was more sensitive for one model (TM00185), while the median was more sensitive for the other four. Two of these five models had major outliers (TM00185 in control group, TM00302 in cisplatin group); one model (TM00185) had four of eight control tumors not exceeding 500mm^3 over the duration of the study.

Treatment response classification at 21 days is concordant with classification at 28 days

Figure 2-4 shows the response classifications on study days 14, 21, and 28 for the three treatment response methods. All models classified as highly responsive or responsive on day 21 were consistent with those classified on day 28 with the exception of one model. While models that were classified as highly responsive on day 28 were observed to be responsive to cisplatin on day 14, a 14-day study duration would not have distinguished highly responsive and responsive models. No PDX models demonstrated a response to cisplatin by day seven.

Treatment response classification does not vary by method of calculation

Table 2-4 displays a comparison of the treatment response values that were calculated for the three methods outlined in Table 2-1. Mean and median response values across all 53 models assessed were within 0.01 (1.7%) of each other for all methods. The standard deviation across models was minimal, but slightly higher in RTC. Figure 2-5 illustrates the classified response values determined by each of the methods

using N=all at day 21. Only 1 (1.8%) of the models had a different response call among the methods.

Figure 2-1 Distribution of cisplatin responses across 53 PDX models treated with single-agent cisplatin; these exhibited various response (TC) values at 28 days. Of 53 models with study length of 28 days, 4 PDX models were highly responsive (TC <0.1) to cisplatin, 11 were responsive (0.1 – 0.42), and 38 were nonresponsive (>0.42).

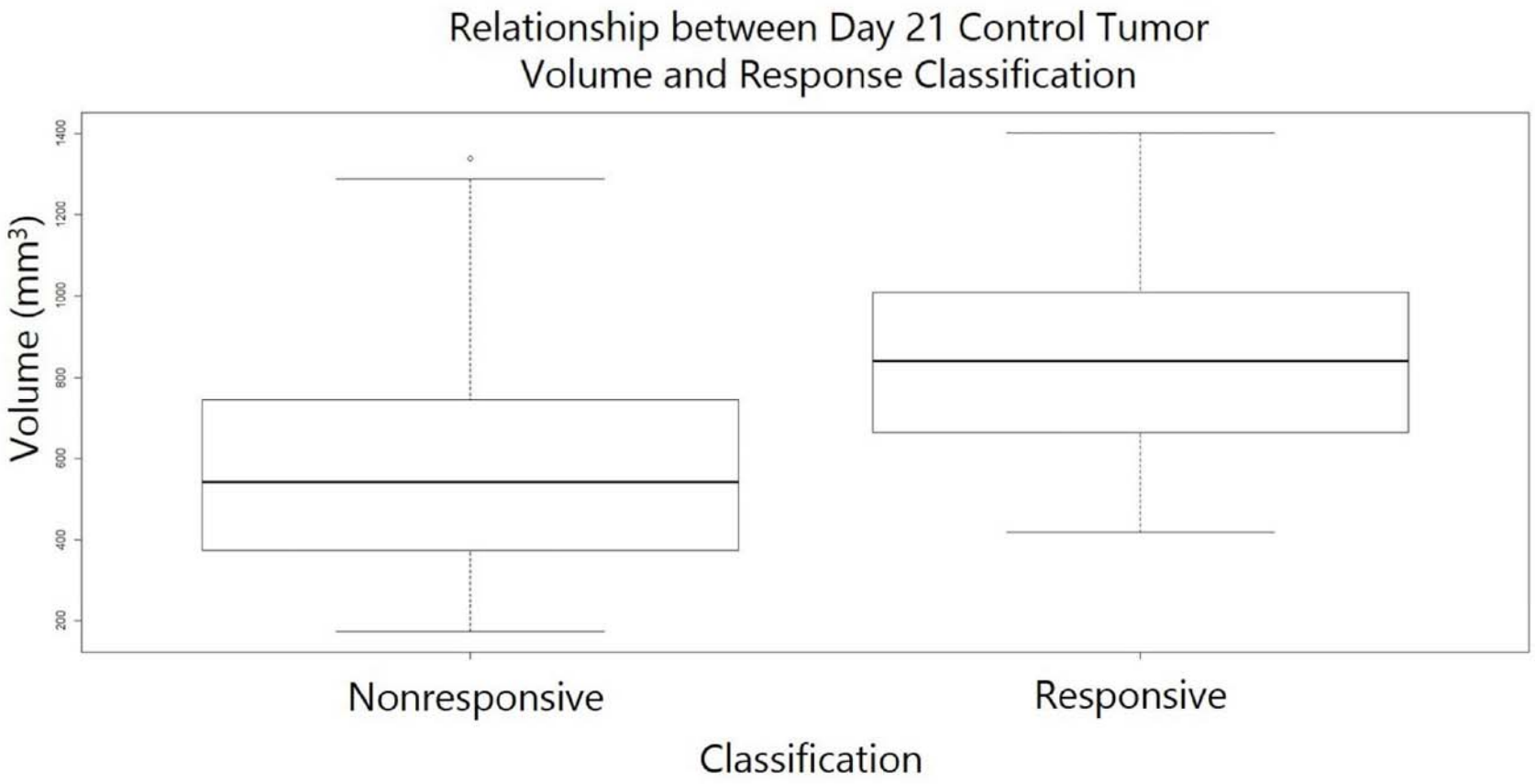

Figure 2-2 Day 21 control tumor volumes by response classified by RTC. Tumors that grew larger during the study were more likely to result in a response classification.

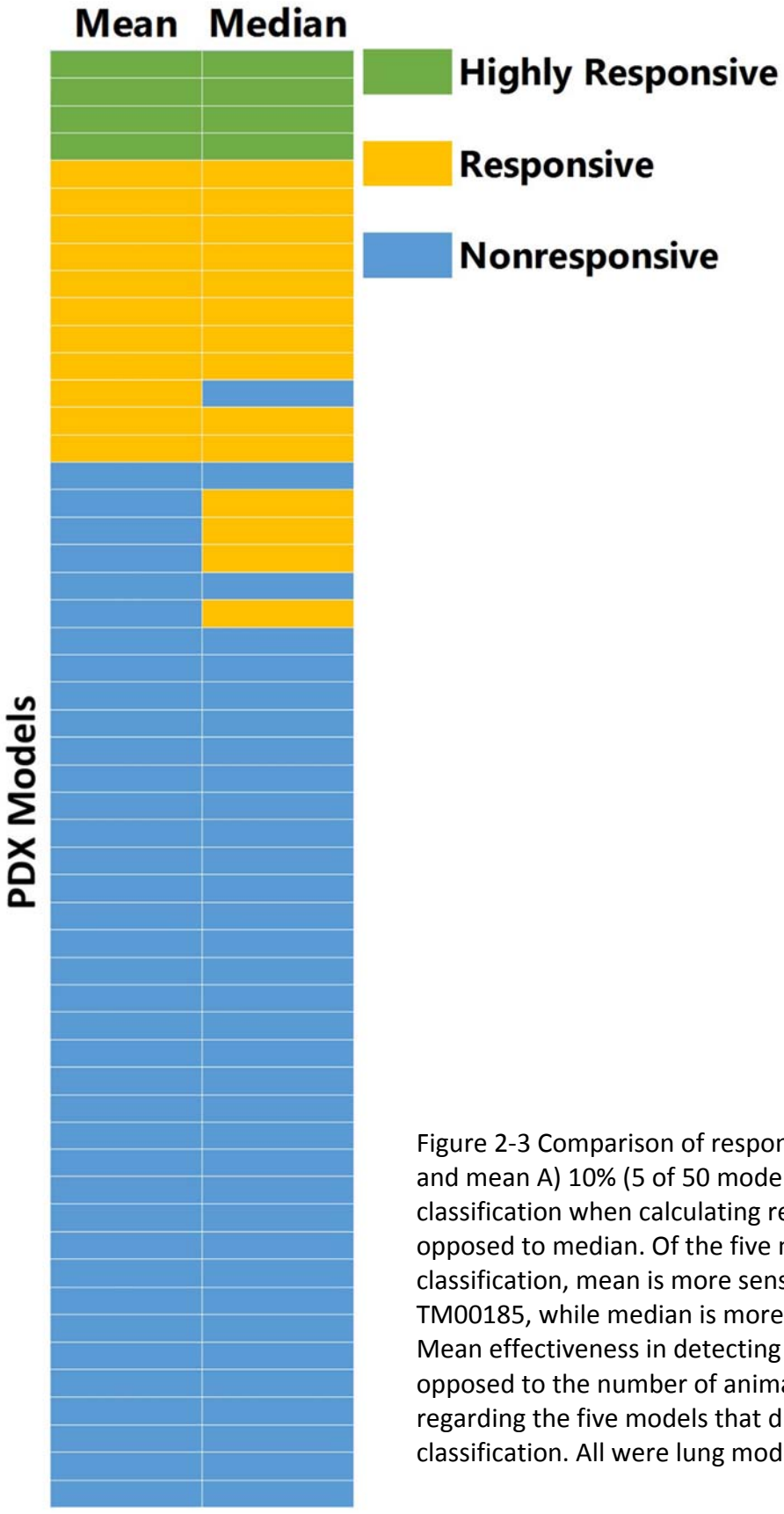

Figure 2-3 Comparison of response values for median and mean A) 10% (5 of 50 models) display a change in classification when calculating response with mean as opposed to median. Of the five models that differ in classification, mean is more sensitive for one model, TM00185, while median is more sensitive for four. B) Mean effectiveness in detecting responses (<0.42) as opposed to the number of animals per group. C) Details regarding the five models that did indicate a change in classification. All were lung models with passage ≤ 5.

Table 2-3 Mean power of classification versus the number of tumors per group. Calculation of RTV is shown becuase it is representative of the trend in power across number per group of both TC and RTC methods.

Number of Animals	Power to detect response by using mean volume	Power to detect response by using median volume
3	0.770	0.716
4	0.826	0.796
5	0.869	0.811
6	0.881	0.852
7	0.963	0.903
8	1.000	0.986

Table 2-4 Comparison of individual-level response values by method at day 21, N=all

	TC	RTC	RTV
Average	0.59	0.59	0.58
Median	0.62	0.63	0.63
Standard Deviation	0.27	0.28	0.27

Models with greatest and least responsiveness demonstrated the lowest probability of changing classifications at a lower N.

A probability analysis was performed to determine the frequency in changes in response classifications when a smaller N is used (figure 2-6). This assessment addressed fifty-three models with PDX response data on day 21 with an N of 8 or greater. Overall, the most highly responsive and least responsive models correspond with the lowest probability of changes in classification at lower N, which suggests that the ability to use a smaller cohort size depends upon the degree of response that is of interest. The RTV method consistently classified models as highly responsive to cisplatin and indicated a 0% to 5% probability that the classification would differ from N = all when using N = 3. The TC method consistently classified models as highly responsive to cisplatin (0% to 10% probability of the classification differing from N = all) for subset cohort sizes as low as N = 3. Overall, the highly responsive and least responsive models have the lowest probability of a change in classification at lower N, suggesting that the ability to use a smaller cohort size depends on the degree of response that is of interest.

A cohort size of seven is required to achieve high statistical power

The RTV method consistently classified models as highly responsive to cisplatin, with a 0% to 5% probability that the classification would differ from N of all when using an N of 3. The TC method consistently classified models as highly responsive to cisplatin (0% to 10% probability of the classification differing from N=all) for subset cohort sizes as low as N = 3. The highly responsive and least responsive models demonstrated the lowest probability of changing classifications at lower N.

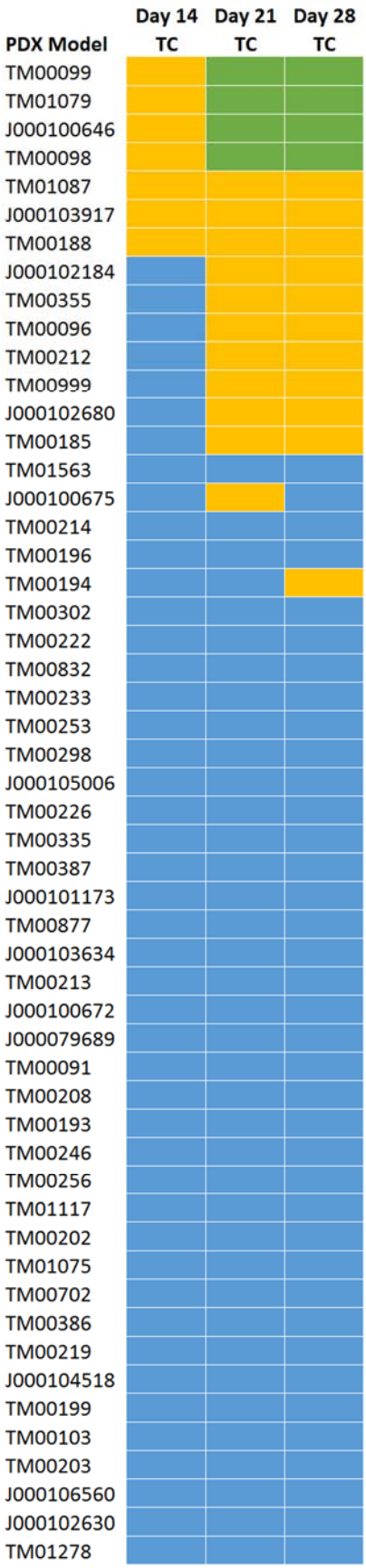

Figure 2-4 Response classification at study day 14, 21, and 28

Figure 2-6 displays the mean effectiveness (as conveyed by the statistical power of the test) in classifying the model response. In general, the effectiveness decreases as the number of tumors per group decreases.

Figure 2-7A demonstrates that the RTC method is the least effective for determining this classification, with effectiveness declining at an N of3 (power of 0.534) relative to TC (0.791) and RTV (0.835). Both TC and RTV have an effectiveness of over 0.99 when an N of 8 is used; RTC's effectiveness is also high, at 0.97. Figure 2-7B displays the effectiveness of identifying a response for the highly responsive models, which is 1.0 across all methods; this indicates that a highly responsive model would consistently be recognized as at least responsive with an N of 3. The statistical power in classifying nonresponsive models for TC, RTC, and RTV are 0.585, 0.692 and 0.661 respectively. The power in classifying highly nonresponsive (greater than 0.80) models as nonresponsive are 0.830, 0.901, and 0.878 respectively. Aside from RTC for nonresponsive classifications, the RTV method is consistently higher in effectiveness across categories and number of animals used. Correlations between power and growth rate, initial tumor volumes, and control tumor volumes were run. Only weak correlations ($R < 0.30$) were found; the largest correlations detected indicated that the standard deviation of the initial control tumor volume is negatively correlated with the mean effectiveness of nonresponse classifications (Spearman coefficient of -0.41).

While the highly responsive and nonresponsive models were assessed using an N of 3, a greater cohort size may be necessary to interpret response in models other than those with extreme response or nonresponse. The RTV was the most effective for classifying any

responsive models with N=3 at 77.0%, as illustrated in Figure 2-7C, but an N of 7 was required to achieve above 90% statistical power (the chance of calling a treatment response when there is a treatment response to be detected), and an N of 8 was required to achieve an effectiveness rate of over 99.0%.

Findings from studies of docetaxel produced similar results to those of cisplatin

Like the findings for cisplatin, the three methods demonstrated consistent docetaxel response classifications in 96% of models (22 of 23) and consistently classified the same models as highly responsive to docetaxel (figure 2-10). Unlike cisplatin where no models demonstrated responses by day seven, three models had a docetaxel response classification computed for at least one of the three methods. Like cisplatin, the models highly responsive to docetaxel have minimal (less than 5%) probability of changing classifications with a cohort size of three animals. The RTV median method was the least effective in classifying docetaxel responses for 60% of models (12 of 20); in contrast, it was the most effective method for detecting cisplatin responses. The RTC method classified the lowest response values (was most sensitive) for docetaxel in 65% of models (13 of 20), suggesting that the sensitivity of the methods can be therapy-dependent.

No strong correlation was determined for all other variables assessed

Figure 2-8A illustrates the distribution of volumes enrolled across individual mice and enrollment arms. Average enrolled volume is 173.4mm^3 (SD 69.0), with a median of 163.8 mm^3 and a range of 53.1-326.9 mm^3. Figure 2-8B indicates the starting volume average across all

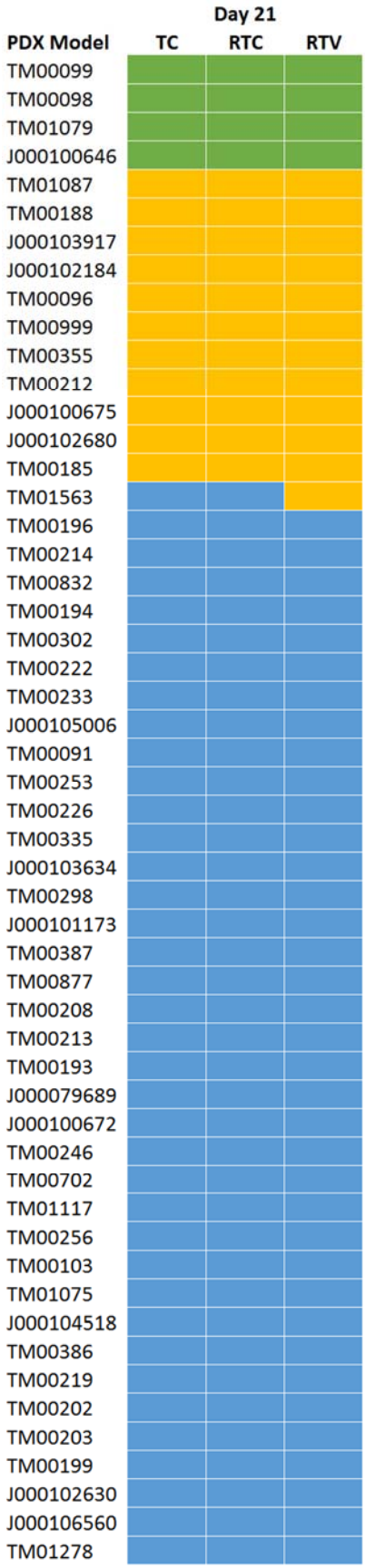

Figure 2-5 Response classification at day 21 across three individual-level response methods

PDX Model	Response Value (TC)	Probability of response classification change (%)		
		TC	RTC	RTV
TM00099	0.05			
TM00098	0.05			
TM01079	0.06			
J000100646	0.07			
TM01087	0.14			
TM00188	0.20			
J000102184	0.28			
TM00096	0.30			
TM00999	0.32			
TM00355	0.33			
TM00212	0.33			
J000104256	0.34			
J000100675	0.38			
J000102680	0.38			
TM01563	0.44			
TM00214	0.46			
J000100674	0.48			
TM00832	0.49			
TM00194	0.49			
TM00302	0.52			
TM00222	0.55			
J000105006	0.57			
TM00091	0.58			
TM00253	0.58			
TM00107	0.60			
TM00226	0.62			
J000080739	0.63			
TM00335	0.65			
J000103634	0.66			
TM00298	0.67			
J000101173	0.70			
TM00387	0.72			
TM00877	0.72			
TM00208	0.74			
TM00193	0.75			
TM00702	0.78			
TM01117	0.80			
TM00256	0.84			
TM00103	0.84			
TM01075	0.87			
J000104518	0.88			
TM00386	0.88			
TM00219	0.89			
TM00192	0.90			
TM00203	0.91			
TM00199	0.92			
J000106560	1.06			
J000104944	1.07			
TM01278	1.09			
TM01149	1.19			

Figure 2-6 Probability of a change in classification when using N = 3 relative to using N =< 8

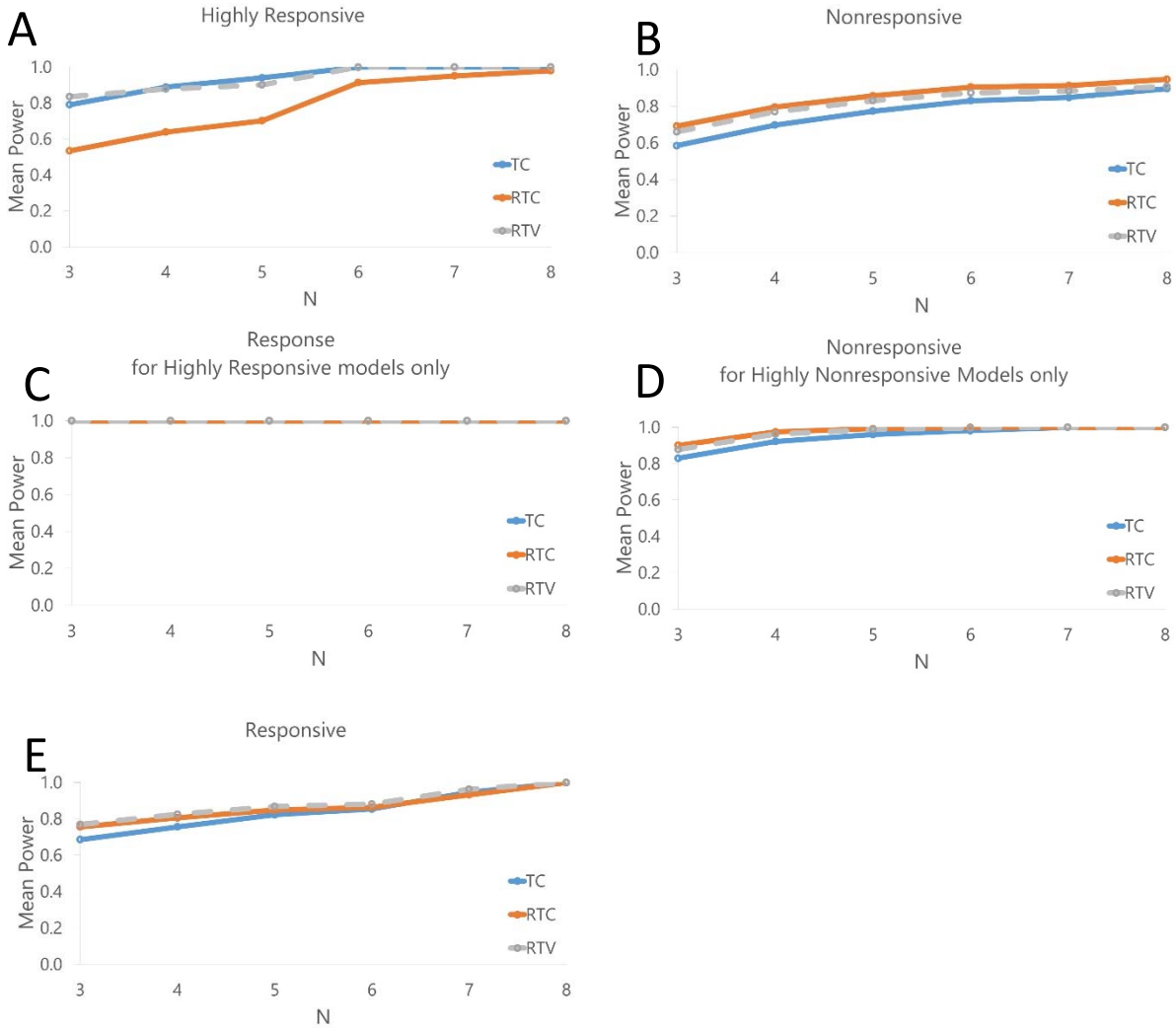

Figure 2-7 Mean power to detect cisplatin response classifications by number of animals used per group for each response method; A) mean power for highly responsive classifications; B) mean power for calling a response in models that are highly responsive (< 0.10), which demonstrates that highly responsive models would be classified as responsive with 100% power; C) mean power for responsive models; D) nonresponsive models; E) mean power for nonresponsive classifications of only those models that were highly nonresponsive (> 0.80).

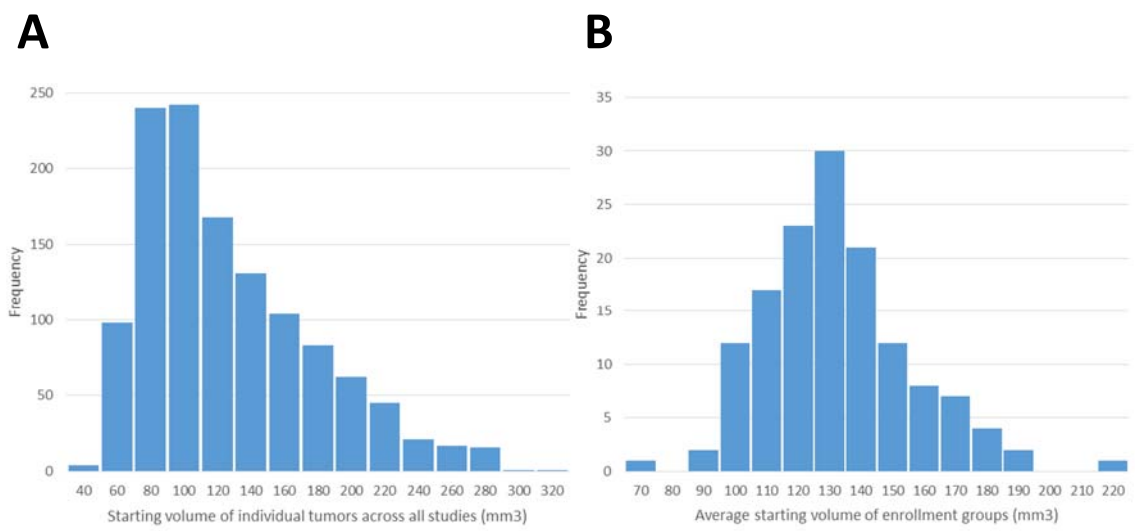

Figure 2-8 Distribution of tumor volume ranges at study enrollment across 70 PDX models: A) Distribution of starting volume of all individual tumors enrolled; B) average starting volume across all control and treatment groups of studies in the dataset.

enrollment groups. The average initial group volume is 137.2mm^3 (SD 23.2), with a median of 136.2 mm^3 and a range of 75.8-220.6 mm^3. While there was a range of both initial tumor volumes across groups (figure 2-8) and sizes within groups, within the dataset used, no correlation between initial tumor volumes and response values (TC) at day 21 or day 28 exceeded an R^2 value of 0.10. Study design variables including tumor growth dynamics (growth rate, doubling time, acceleration of growth), tumor volume, and volume range of treatment and control groups were assessed in terms of whether they influenced response values. Figure 2-9 displays the random distribution of the response values in relation to the control arm average as an example of what was observed. No other strong correlations ($R > 0.30$) between the variables and response were identified.

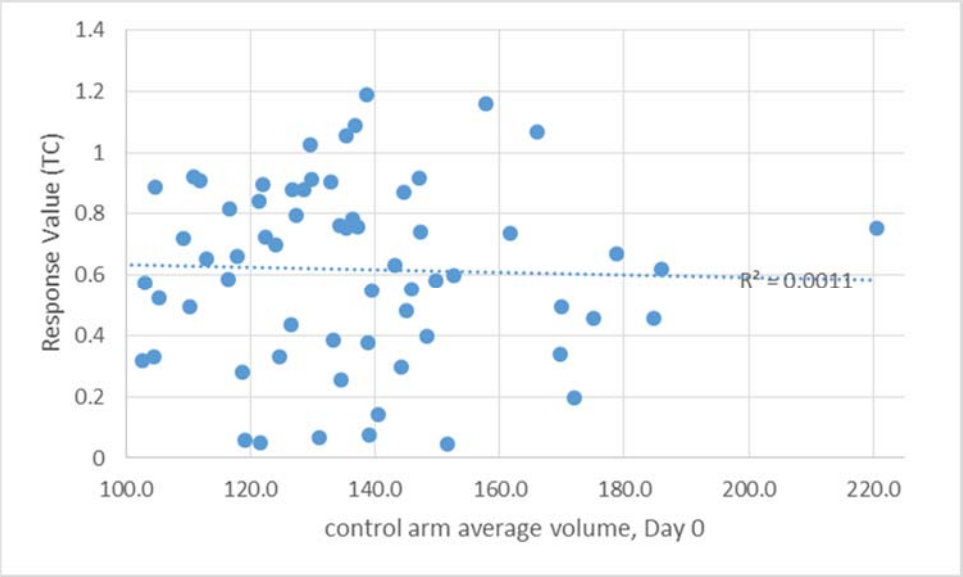

Figure 2-9 No relationship between control arm initial tumor volume and response value

Individual-level as opposed to population level methods

Figure 2-11 illustrates the cisplatin classifications of both the individual-level method (TC) and a population-level assessment. In a standard PDX Clinical Trial CT study, one mouse per treatment arm would be used. However, to demonstrate the reproducibility of this system, the mRECIST classification was calculated for each individual mouse used in the study group.

Twenty-three percent of PDX models (14 of 51) corresponded with a consistent mRECIST classification for all tumor replicates, and all of these were classified as progressive disease. The most frequently called mRECIST classification across a group was described as the "likely mRECIST" for that model. Additionally, this figure indicates the individual-level response classification called for each tumor. There are various discrepancies between the individual model responses; for instance, the six models that were demonstrated to classify the response

as progressive disease (mPD), which indicates limited tumor response to the drug, were classified as responsive by the individual-level method. However, there are similarities between the overall response when the response rates are observed across the panel of tumors. Figure 2-10B illustrates the population-level breakdown of the individual and mRECIST classifications; it displays similar rates between mPR and highly responsive (5% and 8%, respectively), mSD and responsive (18% and 24%, respectively), and mPD and nonresponsive (71% and 74%, respectively).

Discussion

In this study, we examined three commonly used methods for classifying treatment responses in PDX models and the impact of several study design factors on the consistency of those classifications. We showed that a cohort size of three is sufficient for identifying the four highly responsive and nine highly non-responsive cisplatin treated tumors, suggesting that the use of a low cohort size to screen for chemotherapies that have a high degree of activity or models that are highly responsive is possible. However, it is important to note that cohort size depends on the study endpoints; distinguishing between the anticancer activity agents with similar tumor responses is not always possible with a low N. Achieving a balance between ensuring reproducible results and reducing cohort sizes and study duration will minimize the number of animals required and enable savings in laboratory resources.

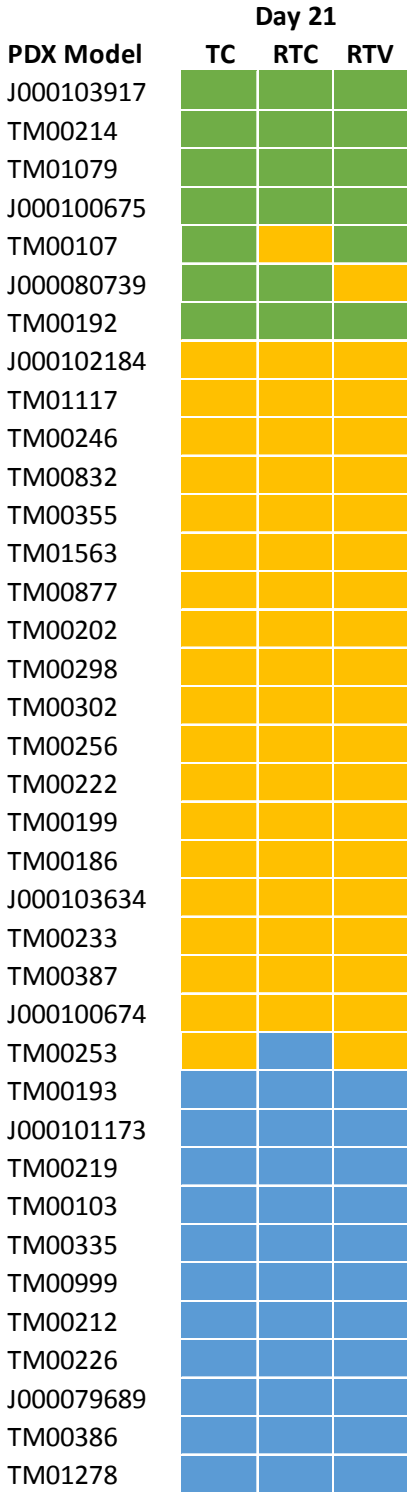

Figure 2-10 Docetaxel response classifications by model for each classification method. Data is displayed for day 21, N = all.

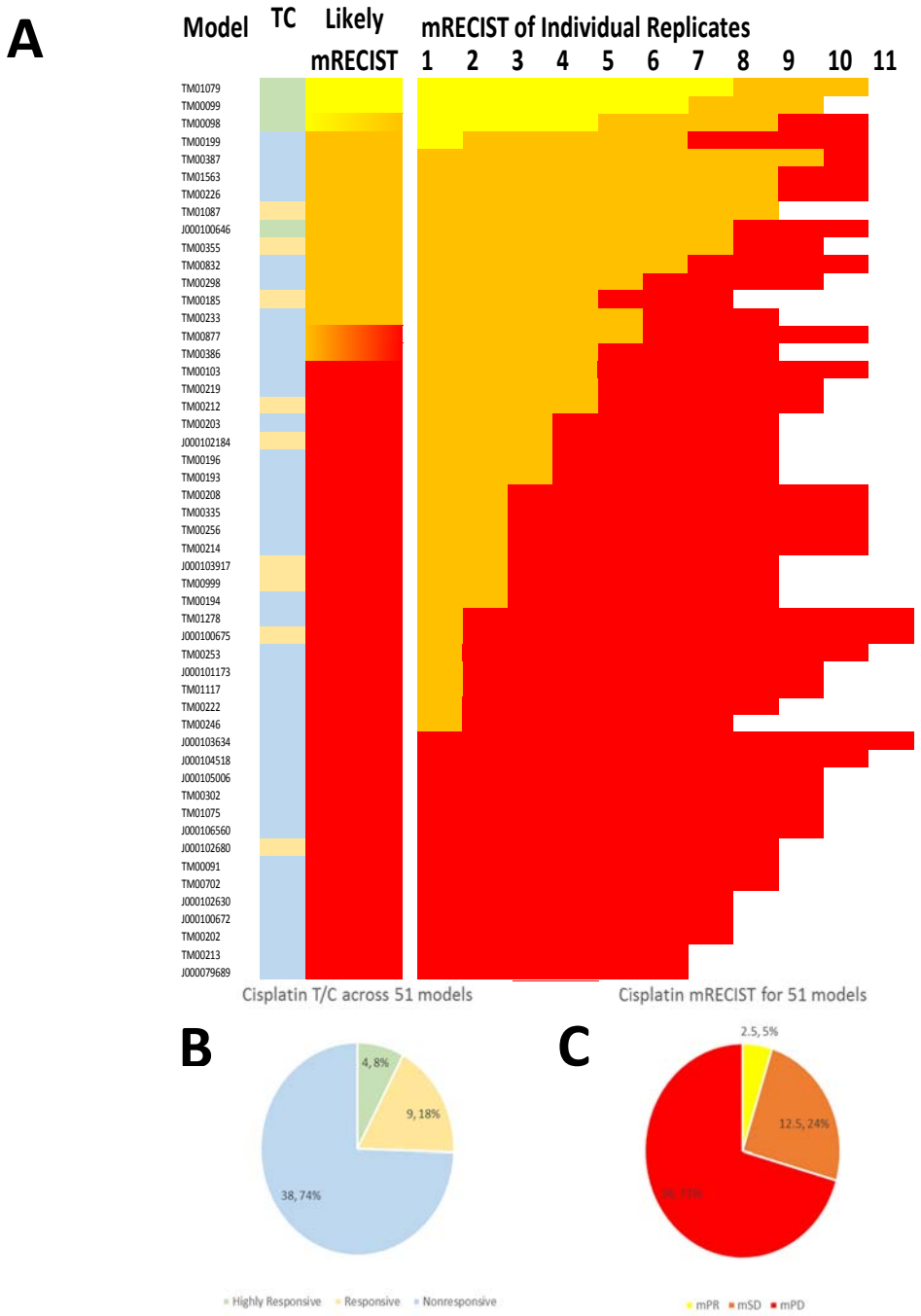

Figure 2-11 Individual and population-level response in cisplatin compared to likely C call A) The TC call of highly responsive (green) responsive (orange) and nonresponsive (blue) are shown based on the average of the groups. The modified RECIST (mRECIST) is shown as complete response (yellow), partial response (orange) stable disease (red) and progressive disease (red). mRECIST is shown for each tumor on far right. The mRECIST classification that was found most frequently in the individual samples was called as the "Likely mRECIST". No direct correlation between individual model responses using TC versus mRECIST was found. B) and C) show that the total number of responsive calls across the model bank are similar between the methods

We found that the average volume of the control groups of the models that were classified as responsive were larger than those that were classified as nonresponsive indicating that the models that grew larger when unchallenged also had a greater degree of response to the compound. While this makes intuitive sense as a larger denominator in the treatment to control ratio would indicate a higher response, it may also suggest that a minimum tumor volume threshold is needed to assess response. This also highlights the importance of utilizing a control as results may be dependent upon the behavior of the tumor under normal/vehicle conditions as well as on the treatment.

Our results showed that a 21-day study duration is sufficient for reliable treatment response classification in studies focused on estimating drug efficacy. We observed that the same classifications that were discovered with a 28-day study duration could also be discovered for only a 21-day study duration for the drugs tested; this could enable faster distribution of data and greater savings of valuable lab resources. This finding is not consistent with the 14 day optimal duration shown by Hather, and we hypothesize that the lack of concordance is due to the differences in model system (CDXs vs PDXs). We note that as with cohort size, study duration is very dependent on endpoint criteria. For example, monitoring for acquired resistance or duration of response, will require longer study duration while dosing studies aimed at understanding infiltration of a compound into a tumor may require the tumors to be harvested and evaluated within hours or days after drug exposure.

Finally, we found that each of the methods consistently classifies responsiveness which suggests it is feasible to combine response data across studies that apply different methods

assuming consistent response value thresholds have been used. The frequency of outliers in treatment response in PDX tumors has caused many researchers to consider the median to reduce the impact of outlying tumor measurements on the response classification[46,67,68], but our work shows that using the mean provides a more reproducible classification if no major outliers exist. Various enrollment criteria have also been described in the literature [61,69], but most frequently, an initial tumor volume group average of 200 mm^3 is used[49,70-74], while maintaining a small tumor volume distribution within a study group. We discovered that there is a non-statistically significant correlation between the standard deviation of the control tumor volumes at the beginning of the study and the effectiveness for nonresponsive classifications. Additional mice per cohort should be considered if tumor volumes vary within a study group.

An important caveat to the work described in this manuscript is that our analyses were performed on less than a hundred PDX models and were limited to two chemotherapy agents; the findings may not extend to other drug classes. Other study design factors and analysis methods that could influence treatment responses in PDXs that were beyond the scope of this analysis and should be evaluated in future work. Orthotopically implanted PDXs, for examples, were not considered but may be a factor in both engraftment success[75] and clinically relevant treatment responses to chemotherapy.[74] Co-engraftment of fibroblasts and matrigel is also a method that varies across PDX studies, as does the use of hormone supplements such as estradiol for hormone-dependent cancers; the influence of these factors on study outcomes has not been investigated. The dosage and route of treatment agents were not considered here but could also be important factors in the consistency of treatment responses. With respect to methodology for calculating treatment response, we focused here on methods that compare

treatment to control arms. We were specifically interested in determining if the findings described in Hather et al performed on CDXs would apply to PDXs. We did not investigate response classification methods that are based on tumor regression or starting and ending values of the treated tumors only.

Ultimately, the value of PDXs as a preclinical platform will depend on how robustly treatment responses in these models translate to responses in patients. The results of our analyses do not speak directly to this important issue but they do support the robustness of treatment responses calculated using a typical PDX treatment study design and suggest that treatment responses can reasonably be compared across different studies that use an experimental design similar to the one described in this manuscript.

CHAPTER 3

MOLECULAR FIDELITY OF PATIENT-DERIVED XENOGRAFT MODELS

While PDXs have been demonstrated to capture the cellular and molecular characteristics of human cancer better than cell line-based models and are highly correlated with original patient tumors,[76-81] molecular differences in patient and PDX tumors have also been shown to occur,[82] which raises concerns about the ability of this model system to accurately reflect the patient's tumor's response to treatment.[83] Evidence suggests that the passaging tumors can result in altered molecular profiles that diverge from the original tumor.[77] These changes can occur at rates intrinsic to the tumor type[18] and include genomic instability in PDXs,[17,19] loss of amplifications,[16] or changes in the stroma and tumor microenvironment.[84] It has also been reported that most of these changes, particularly related to the stromal and microenvironment interactions, occur during the first two tumor passages.[77,85,86]

Here, we report on our comparison of genomic data from patient tumors (PT) and their corresponding PDX models using data available from the JAX PDX Resource and evaluate the potential for differences to impact utility of PDXs for accurate modeling of treatment responses. Because the models in The Jackson Laboratory (JAX) PDX Resource were generated using common engraftment methods and hosts, the resource represents a unique source of data to evaluate molecular fidelity between PDX tumors compared to the original patient tumor and also over different passages. Since its inception in 2009, over 1,970 individual PT specimens have been collected by the JAX PDX Repository; with over 400 models across 42 cancer types generated. These highly characterized models include de-identified clinical and demographic

annotations from the patient as well as immunohistochemistry (IHC), targeted exome sequencing, whole transcriptome expression analysis, and whole genome copy number variation assessment of the engrafted tumors. A total of seventeen models in the repository had genomic data for both the PT and the corresponding PDX model (Table 3-1). In some cases results of genetic testing for a limited number of markers were available in the de-identified clinical annotations associated provided along with a patient's tumor material..

Methods

PDX model generation

Methods for PDX model generation are described in Chapter 2.

Molecular Profiling

Somatic mutations. The JAX Cancer Treatment Profile, a 358-gene deep sequencing targeted exome panel assay,[87] and the JAX bioinformatics workflows for genomic analysis[88] were used to identify somatic mutations in both PT and PDX tumors. Variants detected in the patient tumor and corresponding passaged PDX tumor were compared. Any confirmation testing in cases where CTP data for a PDX tumor was found to have failed to validate a marker reported in the clinical annotation of a patient was performed with BioRad Droplet Digital PCR (ddPCR).

Transcriptomics. Whole transcriptome expression profiling was performed to identify putative translocations, RNA splicing abnormalities, and fusion transcript events using mRNA sequencing of the Illumina® TruSeq methodology. Prior to 2016, non-stranded libraries were

constructed using TruSeq RNA Library Prep Kit v2 (Illumina). Starting in 2016, stranded libraries were prepared by the Genome Technologies core facility at The Jackson Laboratory using the KAPA mRNA HyperPrep Kit (KAPA Biosystems), according to the manufacturer's instructions. Briefly, the protocol entails isolation of polyA containing mRNA using oligo-dT magnetic beads, RNA fragmentation, first and second strand cDNA synthesis, ligation of Illumina-specific adapters containing a unique barcode sequence for each library, and PCR amplification. Libraries were checked for quality and concentration using the DNA 1000 assay (Agilent Technologies) and quantitative PCR (KAPA Biosystems), according to the manufacturers' instructions. Libraries were pooled and sequenced 75 bp paired-end on the NextSeq 500 (Illumina) using NextSeq High Output Kit v2 reagents (Illumina), or 100 bp paired-end on the HiSeq2500 (Illumina) using TruSeq SBS v3 reagents (Illumina).Data were analyzed using principal component analysis (PCA) in JMP version 13.2.1.

Copy Number Variants. Allele-specific copy number alteration profiles for PDX and PT tumors were obtained from Affymetrix SNP 6.0 arrays using copy number analysis workflows described in detail in Woo et al.[88] A low level copy number change threshold ($\log_2$(CN/ploidy) = $\pm$0.25) was used to detect sub-clonal copy number gains and losses. For each sample, we identified the regions of gains and losses and calculated their proportions relative to the entire genome. In order to quantify the similarity of the gain and loss regions between PDX and PT tumors, we calculated the Jaccard Index (JI) based on the common or unique gain or loss regions for each PDX and PT pair.

$$JI_{Gain} = \frac{Intersect_{Gain}(bp)}{Intersect_{Gain}(bp) + PT\_unique_{Gain}(bp) + PDX\_unique_{Gain}(bp)}$$

$$JI_{Loss} = \frac{Intersect_{Loss}(bp)}{Intersect_{Loss}(bp) + PT\_unique_{Loss}(bp) + PDX\_unique_{Loss}(bp)}$$

Determining the clinical impact of variants

Mutation calls that differed between patient tumors and engrafted tumors were identified through comparisons of genomic coordinates. All matching entries were filtered into a list of mutations which are likely to exert a clinical impact, thereby removing all non-coding (UTR, pseudogene, intronic features, etc.) and synonymous single nucleotide substitutions from the analysis. The remaining candidate mutations were evaluated using PolyPhen-2[89] and SIFT[90] for in-silico damage predictions, Ensembl for evolutionary orthology[91], InterPro[92] for protein sequence feature analysis, and HGMD[93], ClinVar[94], CKB[95], and PubMed for literature. An assessment to determine a therapeutic relationship was performed for each mutation through a manual review of peer-reviewed scientific literature with information collected from CKB, CIViC, OncoKB, PudMed, clinicaltrials.gov, FDA.gov, and NCCN clinical practice guidelines. Truncating mutations (frameshift, nonsense) and variants that altered canonical splicing sequences within tumor suppressors were classified as “loss of function (inferred)” in the absence of direct studies demonstrating their effects and when these mutations were consistent with established mechanisms of pathogenesis.

Results

The majority of variants with known clinical significance are preserved between PT and PDX tumors

A comparative analysis of the mutation calls obtained for all seventeen models with matched PT and PDX genomics (table 3-1) was performed. The mean depth of coverage was 1059X (1120X for PT, 998X for PDX). All 17 models had known pathogenic mutations and variants of unknown significance (VUS). There were 130 mutations maintained between PT/PDX pairs (average of 8.5 per model, range of 3-26), 227 mutations were detected in the PT but not in the PDX in at least one PDX/PT pair, and 112 mutations were undetected in the PT but detected in the PDX in at least one PDX/PT pair. These mutations were then evaluated to determine molecular function and whether there was any relationship to a therapy either in development or currently on the market. Sixteen of 17 models had clinically significant mutations preserved, 5 of 17 models lost as least one clinically significant mutation, and 4 of 17 models had at least one clinically significant mutation gained.

A total of 51 (tables 3-2 and 3-3) clinically significant variants were retained in PDX/PT pairs, averaging 3 per model (Range: 0-36; Median: 3) across 29 unique genes. The majority of these (80%; 36/45) were truncating loss of function mutations in known tumor suppressors. Included in the preserved mutation cohort were an additional 81 variants of unknown significance from 60 unique genes. The breakdown of types of mutations retained by each of the 17 models is shown in figure 3-1.

Conversely, there were 297 mutations that were identified in PT but not found in matched PDX (figure 3-2). Eleven of those mutations were clinically significant variants found to have a change in presence between PDX/PT pairs (table 3-4 through 3-7) that were involved in an active clinical trial enrollment strategy for indications other than the cancer type of the tumor in which they were identified. These were all determined to be either well-known hotspot mutations or inferred tumorigenic based on evidence found in the biomedical literature (see methods). The majority (7 of 11) of these clinically relevant alterations occur in established tumor suppressor genes, which have recently been pursued as therapeutic targets in preclinical studies and are involved in active clinical trial enrollment strategies. No significant

Table 3-1 Models included in analysis with tissue for PDX and PT matched pairs.

PDX ID	Clinical Diagnosis	Primary Site	Collection Site	Primary/Metastatic
TM01448	lung squamous cell carcinoma	Lung	Lung	Primary
TM01014	adenocarcinoma	Rectum	Chest wall	Metastatic
TM01231	adenocarcinoma	Colon	Liver	Metastatic
TM00926	colon adenocarcinoma	Colon	Colon	Primary
TM01239	colon adenocarcinoma	Colon	Liver	Metastatic
TM00980	colon carcinoma	Colon	Colon	Not Specified
TM01029	invasive bladder transitional cell carcinoma	Bladder	Bladder	Primary
TM01031	adenosquamous carcinoma	Lung	Lung	Primary
TM00113	invasive ductal carcinoma	Breast	Breast	Primary
TM01117	invasive ductal carcinoma	Breast	Breast	Primary
TM01302	leiomyosarcoma	Bone	Bone	Primary
TM01137	melanoma	Skin	Other	Metastatic
TM01149	melanoma	Skin	Lymph node	Metastatic
TM01264	melanoma	Skin	Colon	Metastatic
TM01423	renal cell carcinoma	Kidney	Bone	Metastatic
TM01053	rhabdomyosarcoma	Soft tissue	Retroperitoneum	Metastatic
TM01085	sarcoma	Soft tissue	Lung	Metastatic

Table 3-2 Variants maintained between PDX/ PT pairs that have a relationship to an experimental or approved oncology drug.

PDX ID	Tumor Type	Genomic Marker	Molecular Function	Potential Drug Classes ✓ Approved o Experimental	FDA Approved Therapy (Indicated))	FDA Approved Therapy (Off label)	Guideline resistance
TM00113	Breast cancer	PALB2 p.L253fs c.758_759insT	GOF	✓ PARP Inhibitor		Talazoparib	
TM00113	Breast cancer	PIK3R1 p.T576_L581del c.1728_1742delGA GAGACCAATACTT	LOF (inferred)	✓ mTOR Inhibitor o AKT Inhibitor		Temsirolimus	
TM00113	Breast cancer	TP53 p.V173L c.517G>T	LOF (inferred)	o WEE1 Inhibitor			
TM00926	Colon adenocarcinoma	MLH1 p.S252* c.755C>G	LOF (inferred)	✓ Anti-PD-L1 mAb		Atezolizumab	
TM00926	Colon adenocarcinoma	PIK3R1 p.K575_T576del c.1725_1727delGAC	LOF (inferred)	✓ mTOR Inhibitor o AKT Inhibitor		Temsirolimus	
TM00926	Colon adenocarcinoma	RB1 p.R857H c.2570G>A	LOF (inferred)	o CHK1 Inhibitor			
TM00980	Colon cancer	BRAF p.V600E c.1799T>A	GOF	✓ BRAF Inhibitor ✓ MEK Inhibitor ✓ Multikinase inhibitor ✓ RAF Inhibitor o ERK inhibitor		Cobimetinib Dabrafenib Regorafenib Trametinib Vemurafenib	Cetuximab Panitumumab
TM00980	Colon cancer	TP53 p.R213* c.637C>T	LOF (inferred)	o WEE1 Inhibitor			
TM01014	Rectum adenocarcinoma	FBXW7 p.R278* c.832C>T	LOF (inferred)	✓ mTOR Inhibitor o CHK1 inhibitor		Everolimus Sirolimus Temsirolimus Metformin	
TM01014	Rectum adenocarcinoma	KRAS p.G12C	GOF	✓ MEK Inhibitor ✓ CDK4/6 Inhibitor		Abemaciclib Binimetinib	Cetuximab Panitumumab

		c.34G>T		o o o o o	Anti-CD47 mAb Anti-gpA33/CD3 mAb Anti-KRAS ASO ERK Inhibitor WEE1 Inhibitor	Cobimetinib Palbociclib Ribociclib Trametinib
TM01014	Rectum adenocarcinoma	TP53 p.R248W c.742C>T	LOF	o	WEE1 Inhibitor	
TM01029	invasive bladder transitional cell carcinoma	AKT2 p.E17K c.49G>A	GOF	o	AKT Inhibitor	
TM01029	invasive bladder transitional cell carcinoma	AKT2 p.D324N c.970G>A	GOF (inferred)	o	AKT Inhibitor	
TM01029	invasive bladder transitional cell carcinoma	ERBB3 p.E43* c.127G>T	LOF (inferred)	✓	HER Inhibitor	Afatinib Neratinib
TM01029	invasive bladder transitional cell carcinoma	MSH6 p.T1085fs c.3254_3255insC	LOF (inferred)	✓	Anti-PD-L1 mAb	Pembrolizumab
TM01085	Soft tissue sarcoma	TP53 p.V73fs c.217delG	LOF (inferred)	o	WEE1 Inhibitor	
TM01117	Breast invasive ductal carcinoma	TP53 p.C176R c.526T>C	LOF (inferred)	o	WEE1 Inhibitor	
TM01137	Melanoma	KIT p.L576P c.1727T>C	GOF	✓	Multikinase Inhibitor	Axitinib Dasatinib Nilotinib Regorafenib Sunitinib
TM01149	Melanoma	NF1 p.R1362* c.4084C>T	LOF (inferred)	✓ ✓	MEK Inhibitor mTOR Inhibitor	Cobimetinib Everolimus Sirolimus Temsirolimus Trametinib

TM01231	Colon adenocarcinoma	TP53 p.R273C c.817C>T	LOF (inferred)	o WEE1 Inhibitor			
TM01264	Melanoma	BRCA2 p.A75P c.223G>C	LOF	✓ PARP Inhibitor ✓ Anti-PD1 mAb o CHK1 Inhibitor			Niraparib Olaparib Rucaparib Pembrolizumab
TM01264	Melanoma	NRAS p.Q61R c.182A>G	GOF	✓ MEK Inhibitor o ERK Inhibitor			Binimetinib Trametinib
TM01423	Renal cancer	VHL p.L129fs c.387_399delGGTT AACCAAACTinsTG	LOF (inferred)	✓ Anti-VEGF mAb ✓ Anti-VEGFR mAb ✓ mTOR Inhibitor ✓ VEGFR Inhibitor	Axitinib Bevacizumab Cabozantinib Everolimus Lenvatinib Pazopanib Sorafenib Sunitinib Temsirolimus		Nintedanib Ramucirumab Regorafenib Sirolimus Vandetanib
TM01423	Renal cancer	VHL p.V130fs c.389_399delTTAAC CAAACT	LOF (inferred)	✓ Anti-VEGF mAb ✓ Anti-VEGFR mAb ✓ mTOR Inhibitor ✓ VEGFR Inhibitor	Axitinib Bevacizumab Cabozantinib Everolimus Lenvatinib Pazopanib Sorafenib Sunitinib Temsirolimus		Nintedanib Ramucirumab Regorafenib Sirolimus Vandetanib
TM01448	Lung squamous	PTEN p.S287* c.860C>G		✓ PARP Inhibitor o PIK3CB Inhibitor o AKT Inhibitor o PI3K/mTOR Inhibitor			Niraparib Olaparib Rucaparib
TM01448	Lung squamous	TP53 p.I255_T256del c.764_766delTCA	LOF (inferred)	o WEE1 Inhibitor			
TM01239	Colon adenocarcinoma	TP53 p.R273H c.818G>A	LOF (inferred)	o WEE1 Inhibitor			

Table 3-3 Variants maintained between PDX/ PT pairs that have a relationship to a prognostic indication.

PDX ID	Tumor Type	Genomic Marker	Molecular Function
TM00926	Colon adeno	B2M p.L13fs c.37_38delCT	LOF (inferred)
TM00926	Colon adeno	B2M p.F82fs c.245_246delTC	LOF (inferred)
TM00926	Colon adeno	DNMT1 p.R228* c.682C>T	LOF (inferred)
TM00926	Colon adeno	NOTCH1 p.P2485fs c.7455delC	LOF (inferred)
TM00926	Colon adeno	SOX9 p.S199fs c.596_597insC	LOF (inferred)
TM00926	Colon adeno	SOX9 p.P304fs c.912delG	LOF (inferred)
TM00980	Colon cancer	CDKN1B p.F87fs c.259_262delTTCT	LOF (inferred)
TM01014	Rectum adeno	APC p.E1408* c.4222G>T	LOF (inferred)
TM01014	Rectum adeno	APC p.S837* c.2510C>G	LOF (inferred)
TM01014	Rectum adeno	CREBBP p.L653fs c.1959delA	LOF (inferred)
TM01085	Soft tissue sarcoma	ATRX p.R808* c.2422C>T	LOF (inferred)
TM01137	melanoma	APOBEC3A p.L78fs c.234delG	LOF (inferred)
TM01137	melanoma	AR p.Q58L c.173A>T	GOF(inferred)
TM01137	melanoma	SF3B1 p.R625H c.1874G>A	GOF(inferred)
TM01149	melanoma	APC p.R405* c.1213C>T	LOF (inferred)
TM01231	Colon adeno	APC p.L1488fs c.4463_4467delTATTA	LOF (inferred)
TM01231	Colon adeno	APOBEC3A p.L78fs c.234delG	LOF (inferred)
TM01231	Colon adeno	CREBBP p.G1479* c.4435G>T	LOF (inferred)
TM01239	Colon adeno	APC p.R1114* c.3340C>T	LOF (inferred)
TM01239	Colon adeno	SMAD4 p.L535fs c.1605_1612delCCTAGACG	LOF (inferred)
TM01239	Colon adeno	ROR2 p.P160fs c.479delC	LOF (inferred)
TM01302	Bone sarcoma	YAP1 amp	GOF(inferred)
TM01423	renal cancer	CDKN2C p.E120fs c.358_377delGAGTTCCTGGTGAAGCACAC	LOF (inferred)

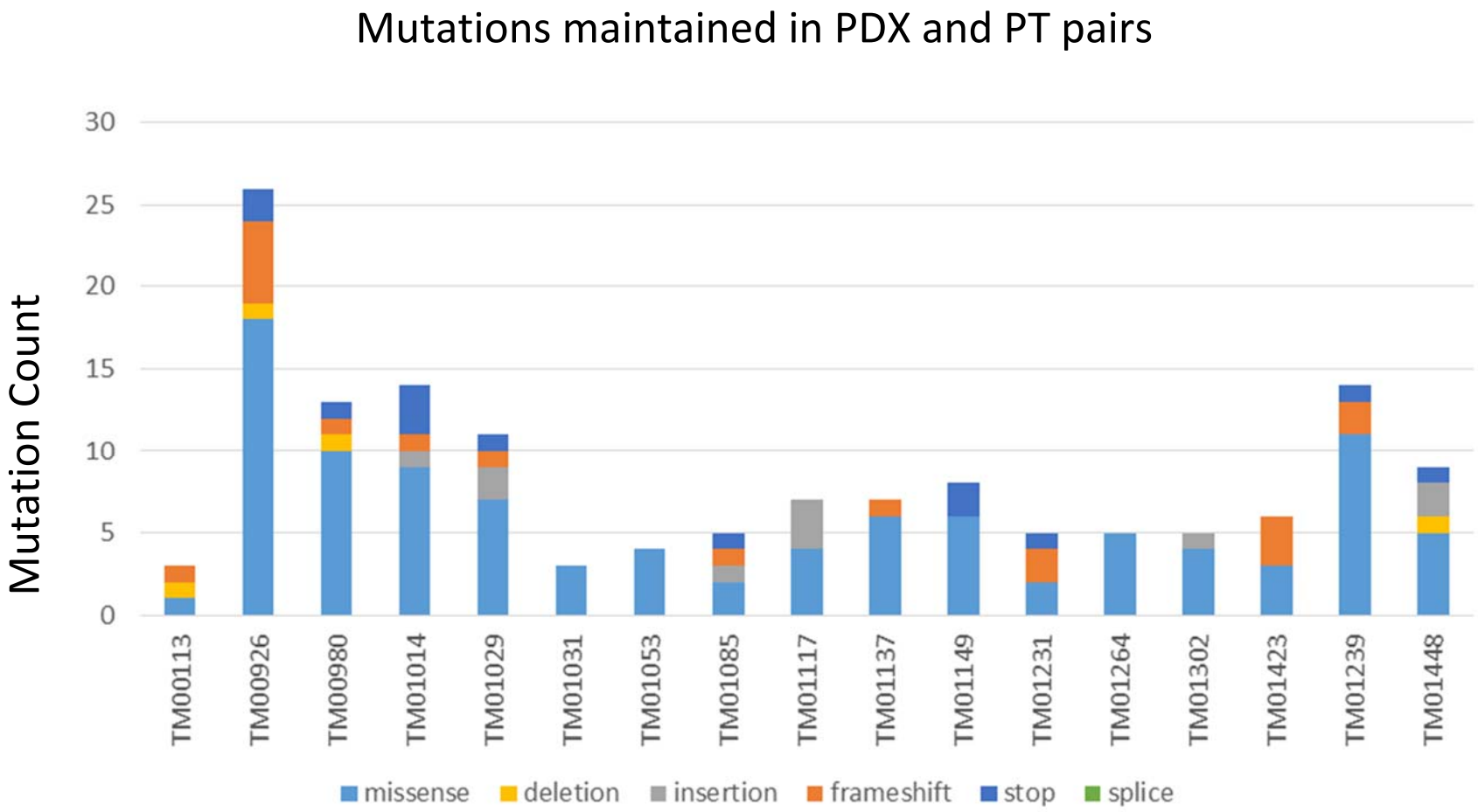

Figure 3-1 Mutations maintained in PDX and PT pairs.

Table 3-4 Mutations with therapeutic relevance identified in PT but not found in matched PDX

Model ID	Tumor Type	Genomic Marker	Molecular Function	Potential Drug Classes ✓ Approved, o Experimental	FDA Approved Therapies
TM01014	rectum adenocarcinoma	**TP53** p.Arg248Trp c.742C>T	Gain of function	op53 gene therapy oWEE1 Inhibitor	NONE
TM01029	bladder carcinoma	**TP53** p.Met243fs c.726_727delCAinsT	Loss of function (inferred)	op53 gene therapy oWEE1 Inhibitor	NONE

Table 3-5 Prognostic markers detected in PDX but not in matched PT

Model ID	Tumor Type	Genomic Marker	Molecular Function
TM00926	colon adenocarcinoma	**EP300** p.Gln895fs c.2683dupC	Loss of function (inferred)
TM01231	colon adenocarcinoma	**APC** p.Arg554* c.1660C>T	Loss of function
TM01117	breast invasive ductal carcinoma	**TGFBR2** p.Lys128fs c.383delA	Loss of function (inferred)
TM00926	colon adenocarcinoma	**MTOR** p.? c.2649+2T>C	Loss of function (inferred)

Table 3-6 Mutations with therapeutic relevance were detected in PDX but not in matched PT

Case ID	Tumor Type	Genomic Marker	Molecular Function	Potential Drug Classes ✓ Approved, o Experimental	FDA Approved Therapy *(All in indications other than tumor type)*
TM01423	Renal cell carcinoma	**TSC1** p.Gly596Arg c.1786G>C	Loss of function (inferred)	✓mTOR Inhibitors	Everolimus Metformin Sirolimus Temsirolimus
TM00926	colon adenocarcinoma	**ARID1A** p.Met1613fs c.4837delA	Loss of function (inferred)	✓PARP Inhibitor oAKT Inhibitor oBRD4 Inhibitor	Niraparib Olaparib Rucaparib
TM01031	Lung adenosquamous carcinoma	**BRAF** p.Gly596Arg c.1786G>C	Gain of function	✓MEK Inhibitor ✓Multikinase Inhibitor oERK Inhibitor	Cobimetinib Sorafenib Trametinib
TM01031	Lung adenosquamous carcinoma	**MAP2K1** p.Lys57Glu c.169A>G	Gain of function	✓MEK Inhibitor oERK Inhibitor oRAF Inhibitor	Trametinib

Table 3-7 Other clinically relevant mutations found in PDX but not in matched PT

Model ID	Tumor Type	Genomic Marker	Molecular Function	Relevance
TM00926	colon adenocarcinoma	**JAK1** p.Lys142fs c.425delA	Loss of function (inferred)	Resistance to Anti-PD-1/ PD-L1 therapy
TM01029	bladder carcinoma	**SF3B1** p.Trp1216* c.3648G>A	Loss of function (inferred)	Prognostic Indicator

trends were identified between variants as a function of their presence (detected or not detected), type of mutation (truncating or point mutations), or the nucleotide conversion (transitions or transversions).

Identified cases of discrepancies between PDX and reported clinical patient tumor information

For many of the models in the PDX resource, there was not sufficient PT to perform comprehensive genome testing. In a handful of cases there were results available of genetic tests on a limited number of markers included with the clinical annotations of the patient sample. In all but two cases, the markers and histology reported in the clinical annotations were also identified in the PDX tumor. In one case a lung PDX model (TM01244) appeared initially to have lost the critical EGFR T790M mutation observed in the patient based on CTP data for the P0 passaged tumor. However, digital droplet PCR performed on the P0 tumor did demonstrate the presence of the expected mutation but at a frequency (0.06%) below the limit of detection of the CTP assay. In another case, the clinical diagnosis of the patient tumor was adenosquamous carcinoma (figure 3-3A). When the tumor was passaged however the resulting PDX tumor retained only the squamous properties (figures 3-3B and 3-3C).

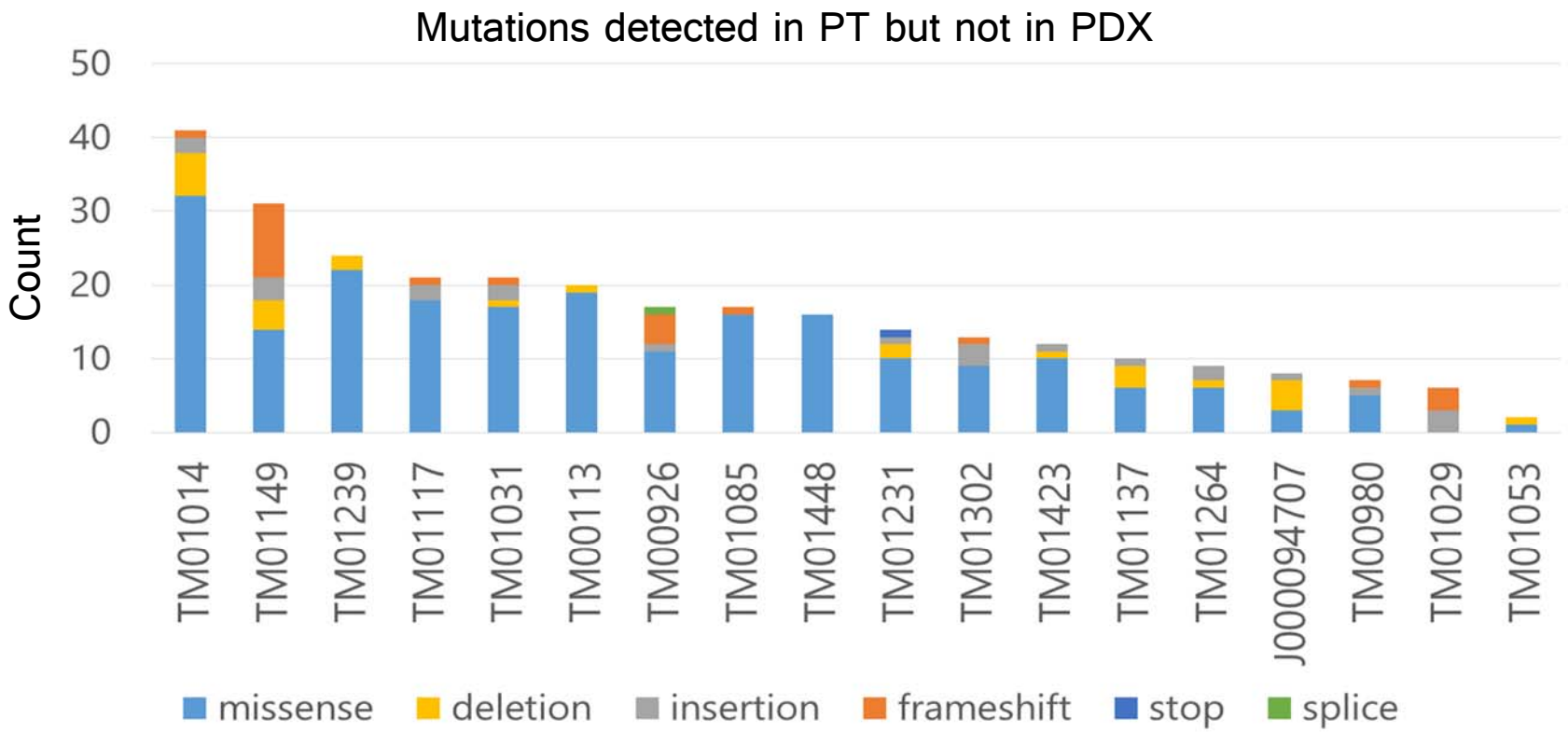

Figure 3-2 Mutations detected in PT but not found in matched PDX

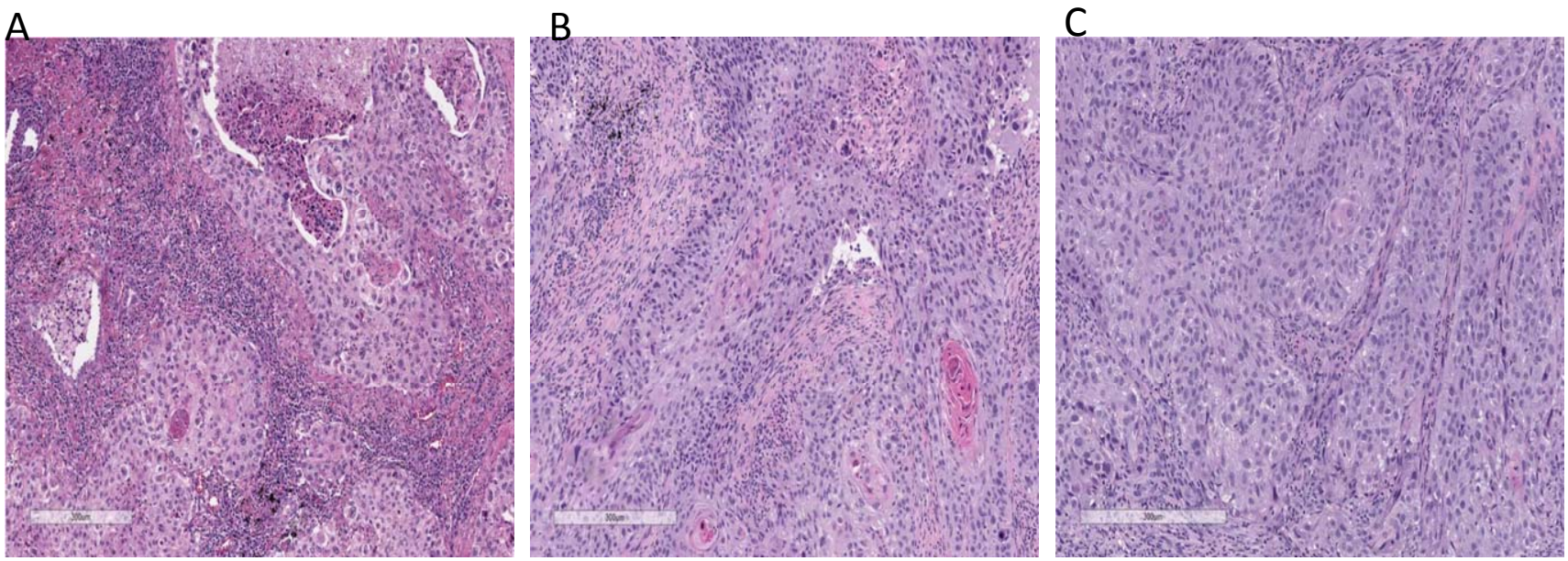

Figure 3-3 Histology of model TM01031, derived from a lung adnenosquamous: A) Patient tumor of model TM01031. Mixed features of gland forming tumor and sheets with squamous differentiation B) PDX Passage 0 showing only squamous component present, in sheets of large cells with keratinzation. C) PDX Passage 1 shows only the squamous component and no recovery of the gland forming tumor cells.

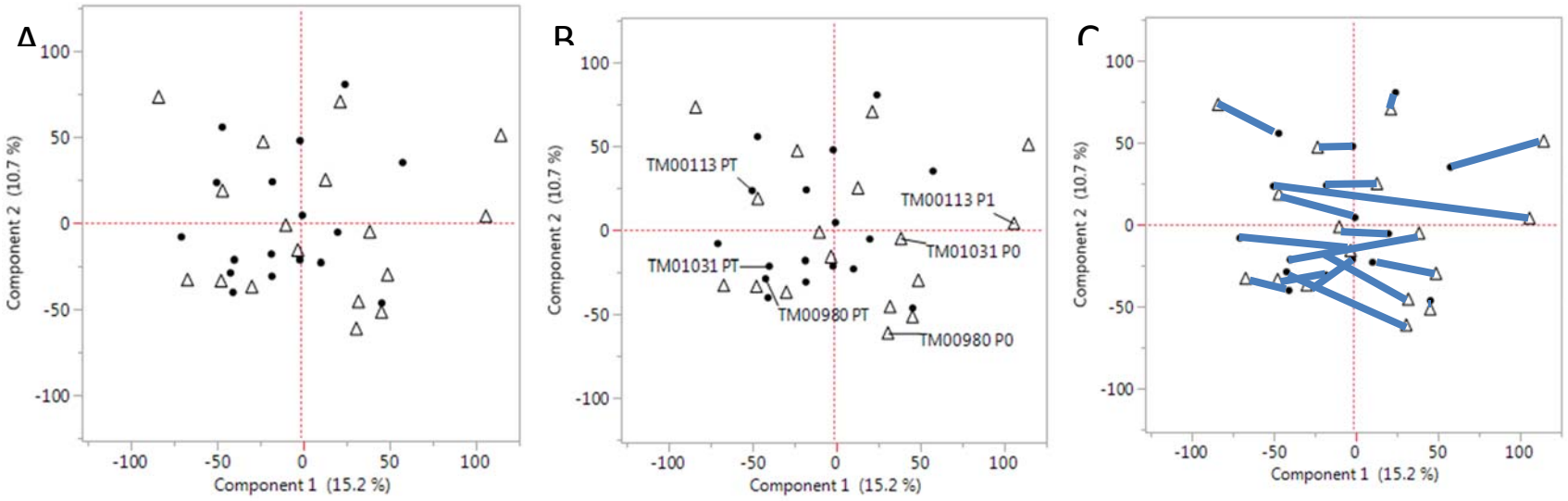

Figure 3-4 Principle component analysis of RNASeq transcripts per million (TPM) data. A) The principle component 1 is 16.9%. B) The three models that did not cluster together (labeled) are separated by PC1. C) Paired PDX/PT pairs.

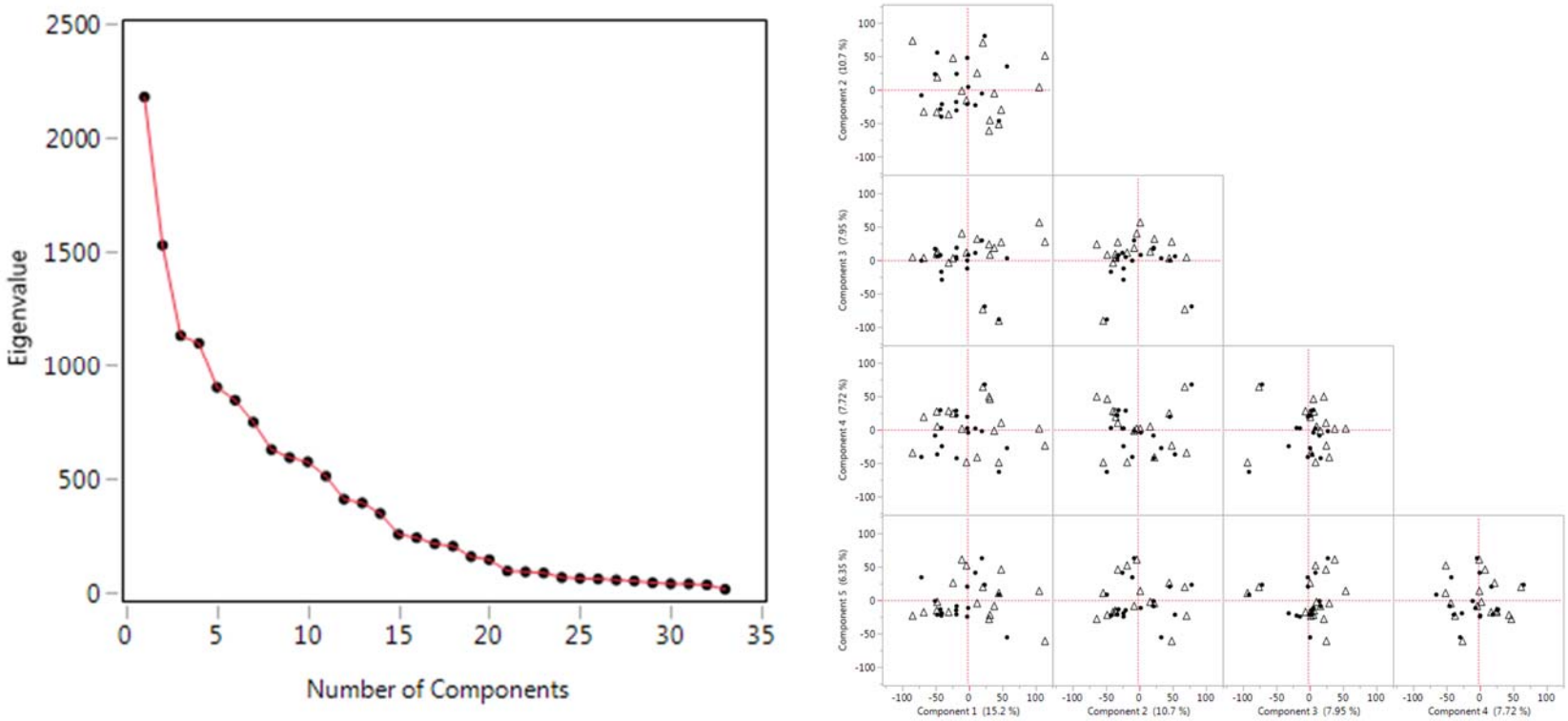

Figure 3-5 Principle component analysis a) Scree plot showing the variation each PC captures from the data b) Biplot with triangles representing that patient derived xenografts (PDX) and circles representing patient tumors (PT). Five principle components are shown. There is principle component that separates PDX and PT, indicating that the signals are driven by other variables amongst the cancers.

PDX and matched PT tumors cluster together based on gene expression

Principal component analysis (PCA) was performed on the gene expression dataset (figure 3-4). All principal components over 5% variance were reviewed to determine whether there was a separation of PT and PDX samples, (figure 3-5), but the samples clustered together indicating the samples are highly similar to one another with respect to global gene expression. Six principal components accounted for the majority (54.0%) of the variation in expression; PC1 accounted for more than 10% variability. PC2 distinguished carcinomas from other cancer types in the sample population.

Unsupervised hierarchical clustering of gene expression data for all matched PT and PDX datasets (figure 3-5) revealed that 15 of seventeen PT or PDX pairs clustered together, which confirmed similarity in global expression among tumor pairs. PT and PDX pairs for three models TM00113 (breast-invasive ductal carcinoma), TM00980 (colorectal adenocarcinoma), TM01031 (lung adenosquamous) did not cluster directly together. Principal component analysis of these datasets demonstrated that the three non-clustering models were distinguishable from the other samples by PC1. The biological pathways associated with genes that were highly correlated with PC1 (>.70) exhibit biological processes were related to regulation of transcription and cell-cell adhesion. Determining if these enriched processes have any explanatory power will require additional data.

Copy number variation between PT and PDX

The proportion of gain and loss across all PDX and PT tumors analyzed range from 0 – 44% and 0 – 58% (figure 3-7 and 3-8). The overlap of the copy number regions, as reflected by the Jaccard Index, varies from 0 – 0.97 (mean=0.72, median=0.83) for gains and 0 – 1 (mean=0.69, median=0.81) for loss. The lowest Jaccard Index of the models was seen in TM01031 which was 0 for both gains and losses, and a proportional PDX gain of 12.4% and PDX loss of 46.9%. Note that this model is also highlighted in figure 3, an adenosquamous carcinoma that lost the glandular (adeno) component of the tissue after passage in PDX. Model TM00926 also showed a low Jaccard Index of 0.29 for loss. There is a low proportion of gain and loss for this model is particularly low at <0.01 across both PDX and PT tumors. All other models had a Jaccard Index >0.45 for gains. Losses were showed less similarity across the models with 6 models falling below a 0.45 Jaccard Index.

Discussion

The results of our assessment of mutation, gene expression, and copy number variant data for models from the JAX PDX Resource indicates a high degree of correspondence between PDX tumors and their corresponding patient tumor. Most mutations with clinical significance were maintained between PT and PDX. Cluster analysis of global gene expression demonstrated that, in most cases, the PT and PDX samples clustered together. Principal component analysis of the available expression data yielded similar results (i.e., paired PDX and PT samples group together).

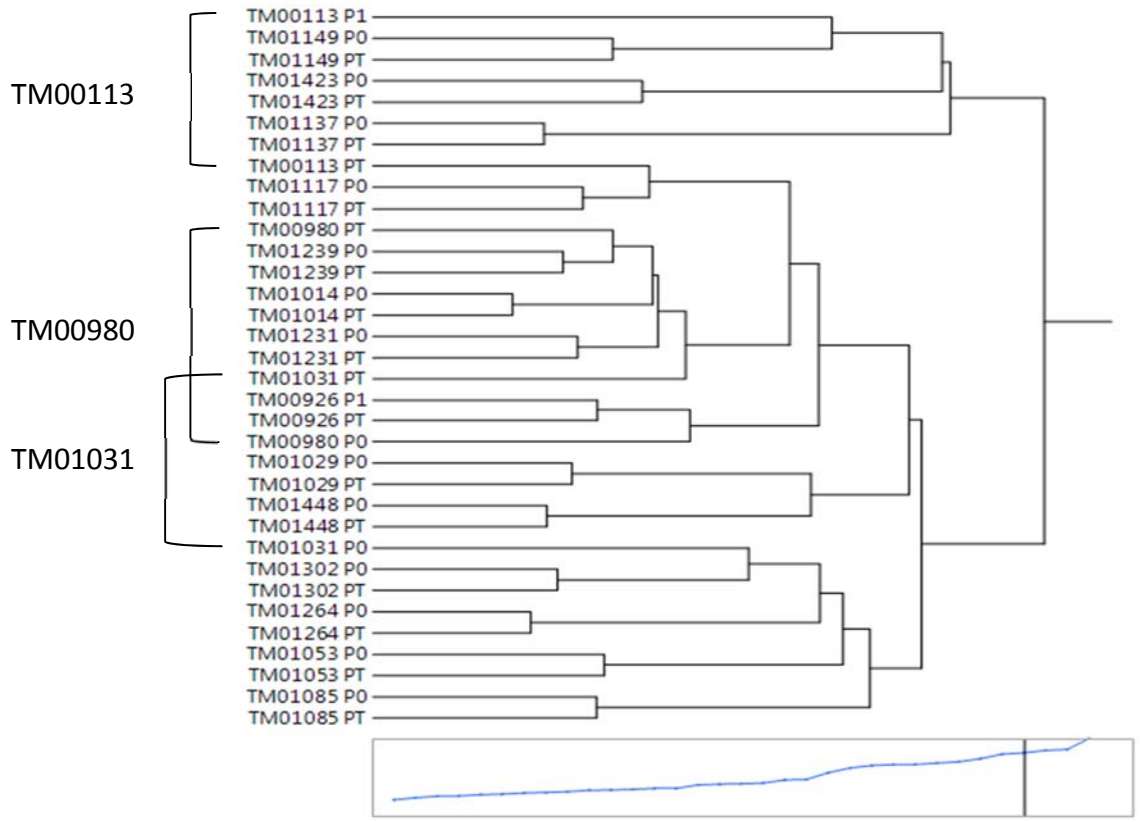

Figure 3-6 Clustering of RNASeq datasets shows all but three PDX/PT pairs cluster together.

Despite observing overall similarity, we did identify examples of molecular differences between the patient tumor and the PDX for some markers, including those associated with cancer drugs currently in development. In all cases, however, the associated drugs were for cancer types other than the samples being compared. That said, since PDX models are potential testbeds for evaluating drugs for indications other than those they are currently approved for, the differences observed in this analysis strongly supports the practice of confirming the presence of key markers prior to placing models on study.

The tumor heterogeneity and gain or loss of certain mutations through the mouse tumor microenvironment could potentially account for the differences in therapeutic response

relative to what is observed in the patient, since passaging can cause the emergence of a more aggressive tumor.[96] It is possible that variants detected in PDX but not in PT (and vice versa) are not genuinely mutation gains or losses but changes in allele frequencies as a result of selective pressures, loss of human stroma, or chance alone during the process of passaging.

The data presented here for PT and PDX tumors were based on comparing tumors from early passages (first and second). In the literature, tumors with greater than 10 passages have been reported.[69,97] The potential for differences in variants that would impact dosing study responses indicates that genomic characterization of tumors used to establish cohorts for dosing studies is advisable. For example, we noted that PDX model TM01423 appeared to gain a *TSC1* gene defect, which was not detected in the PT. Rather than reflecting a true loss, this observation could have resulted from chance changes in the clonal populations during the engraftment process. If this model was selected for preclinical testing of an mTOR inhibitor, but the passage used in the study was a separate lineage of the one originally sequenced, then the *TSC1* gene variant may not be present in the tissue used in the study. This could result in significant resource and time lost in evaluating a model that does not have the intended target.

We conclude that the molecular fidelity of PT and passaged tumors is high at a global level but that individual markers are not invariably concordant. We identified clear case examples of molecular differences in clinically relevant mutations. Confirming the mutation status of donor tumors used to establish testing cohorts should be a component of routine practice for using PDXs along with other typical quality control measures for ensuring tissue provenance.

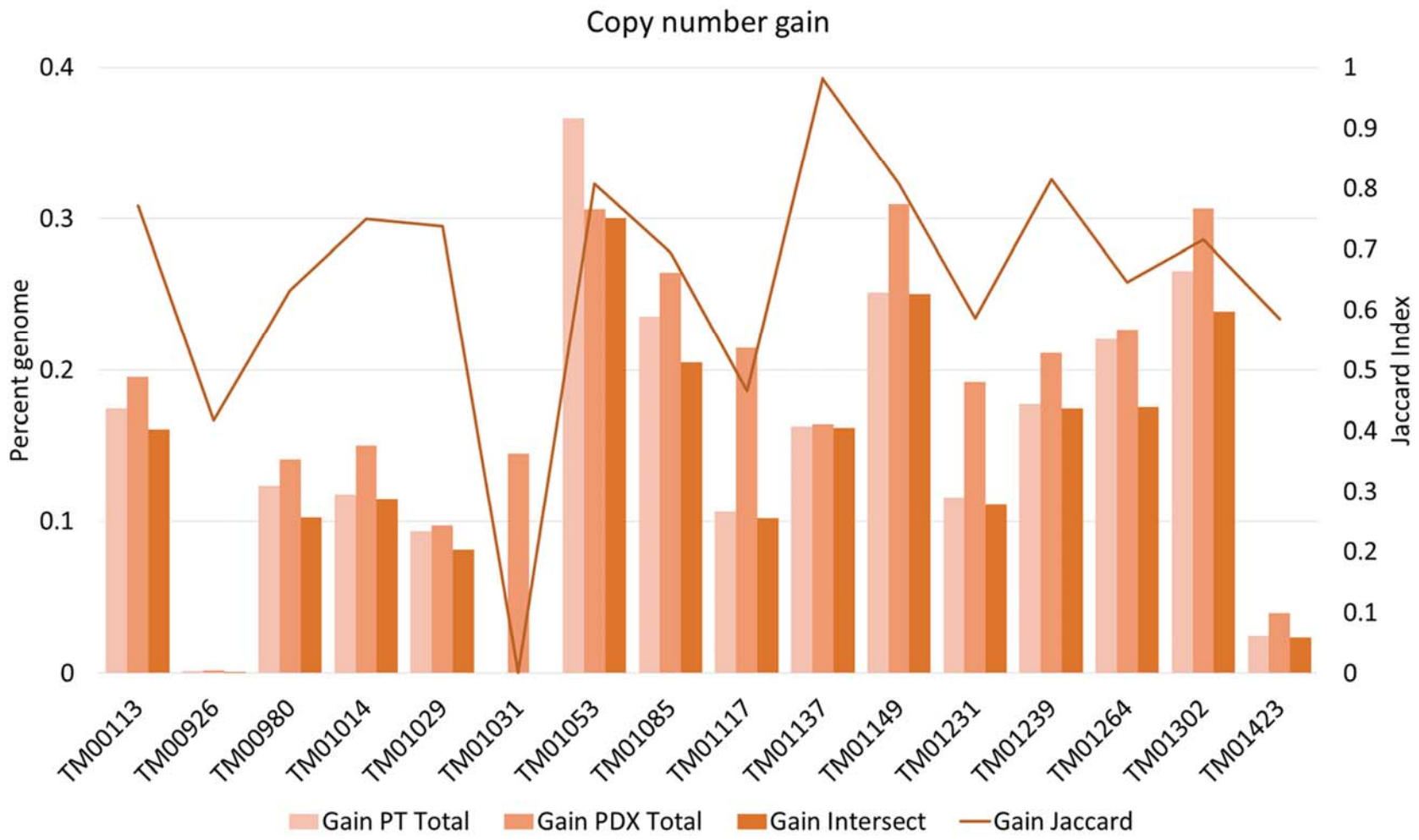

Figure 3-7 Proportion of gain and loss and Jaccard index for each PDX and PT pair.

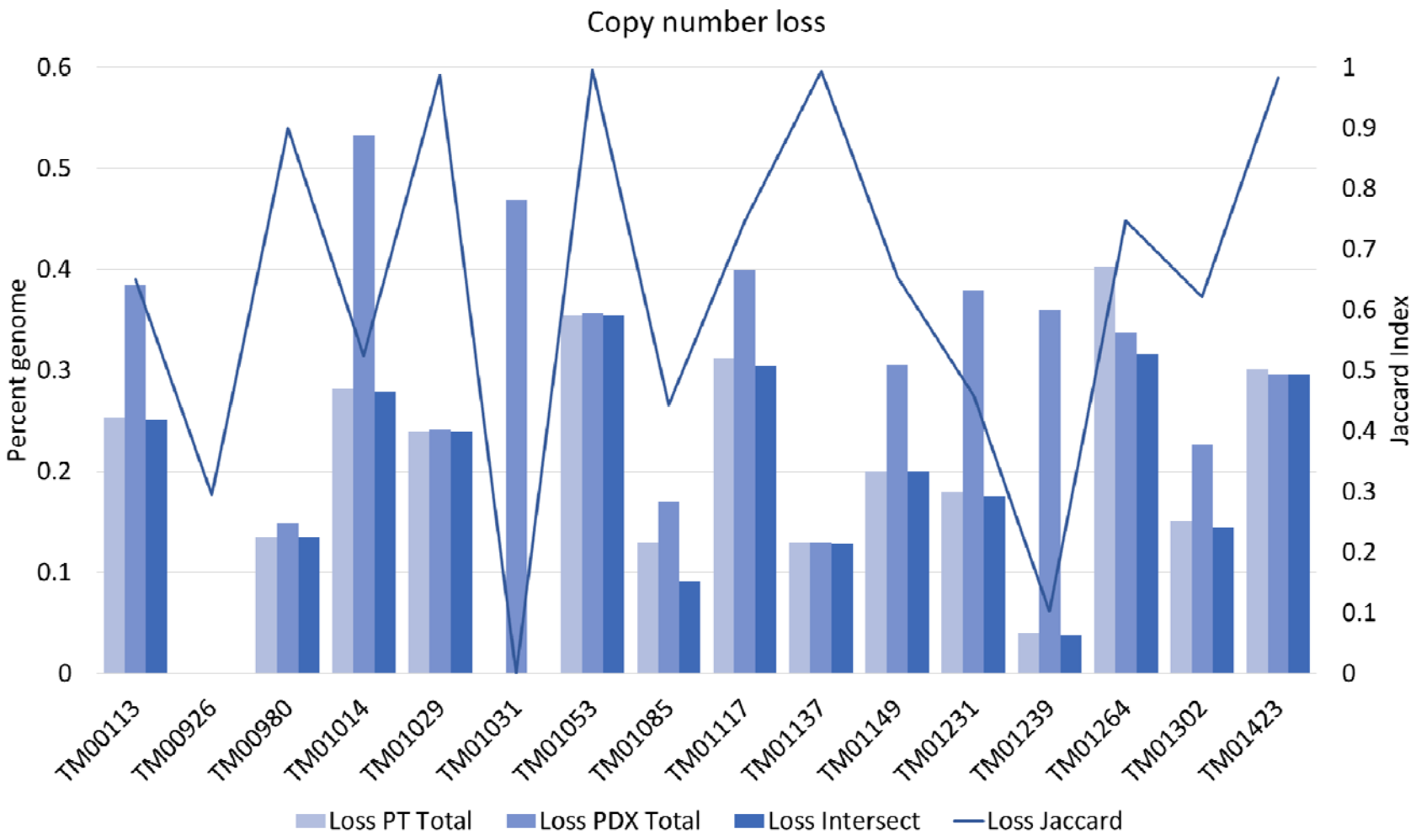

Figure 3-8 Proportion of gain and loss and Jaccard index for each PDX and PT pair.

CHAPTER 4

CONCLUSIONS AND FUTURE DIRECTIONS

The research described in this dissertation addressed fundamental questions about Patient Derived Xenograft models as a reliable preclinical platform for modeling cancer therapy.

The work represents the first assessment of the impact of study design factors of initial tumor volume, rate of tumor growth, cohort size, study duration, and analysis method on the robustness of treatment classifications in a repository of PDX models represented a wide diversity of cancer types. We examined three commonly used methods for classifying treatment responses in PDX models and the impact of several study design factors on the consistency of those classifications. We found that each of the methods consistently classifies responsiveness which suggests it is feasible to combine response data across studies that apply different methods assuming consistent response value thresholds have been used. We showed that a cohort size of three is sufficient for identifying the four highly responsive and nine highly non-responsive cisplatin treated tumors, suggesting that the use of a low cohort size to screen for chemotherapies that have a high degree of activity or models that are highly responsive is possible. However, it is important to note that cohort size depends on the study endpoints; distinguishing between the anticancer activity agents with similar tumor responses is not always possible with a low N. Achieving a balance between ensuring reproducible results and reducing cohort sizes and study duration will minimize the number of animals required and enable savings in laboratory resources.

We found that the average volume of the control groups of the models that were classified as responsive were larger than those that were classified as nonresponsive indicating that the models that grew larger when unchallenged also had a greater degree of response to the compound. While this makes intuitive sense as a larger denominator in the treatment to control ratio would indicate a higher response, it may also suggest that a minimum tumor volume threshold is needed to assess response. This also highlights the importance of utilizing a control as results may be dependent upon the behavior of the tumor under normal/vehicle conditions as well as on the treatment.

Our results showed that a 21-day study duration is sufficient for reliable treatment response classification in studies focused on estimating drug efficacy. We observed that the same classifications that were discovered with a 28-day study duration could also be discovered for only a 21-day study duration for the drugs tested; this could enable faster distribution of data and greater savings of valuable lab resources. We note that as with cohort size, study duration is very dependent on endpoint criteria.

Although our results demonstrate that PDXs can be a statistically robust platform and that the molecular fidelity of the system is high, the true test of validity depends upon the ability to translate drug response results in PDXs to outcomes in patients. This work has built on the published studies that provide anecdotal evidence of clinical validity, and finds that the molecular fidelity of PT and passaged tumors is high at a global level but that individual markers are not invariably concordant. Confirming the mutation status of donor tumors used to establish testing cohorts should be a component of routine practice for using PDXs along with

other typical quality control measures for ensuring tissue provenance. We have provided clear case examples of molecular shifts in clinically relevant mutations, and we recommend that since PDX models are used in preclinical oncology research, at least evaluation of the status of genomic variants at the tissue at the passage used in studies should be performed.

The results suggest that the platform can generate treatment response endpoints that are reproducible and that do not depend on the classification method used to determine the endpoints. Our work provides data-driven evidence for selecting minimal cohort sizes and study duration needed for reproducible treatment response classifications in PDXs. We also demonstrated the overall maintenance of molecular fidelity of engrafted tumors with patient tumors. That said, specific differences were noted; however, the evaluation of these differences was beyond the scope of the research and awaits future investigation.

Future Research Directions for the PDX platform

While PDXs have been widely adopted as a preclinical model system there remain open questions about what changes are happening in the engrafted tumor and how this impacts the utility of the model system. Work also continues on enhancing the mouse host so that it provides a more "humanized" environment for the tumor cells. Below we discuss some of the future directions for investigating, credentialing, and improving PDXs as a preclinical platform for evaluating cancer therapeutics.

Exploring the selective pressures of PDX on the clonal composition

Engraftment of human tumor cells into a foreign mouse environment exerts a selection pressure, which changes clonal composition of a tumor over time and may lead to the enrichment of particular clonal subpopulations in PDXs that differ from the original patient tumor. The nature of this selection is poorly understood. Some reports suggest that the stronger the initial selection, the more stable the clonal populations are in subsequent passages.[19,47] Aparicio and Hidalgo suggest that this indicates that the clonal changes are therefore the selection of pre-existing clones rather than the evolution of new clones. They also illustrated that when they replicated the engraftment of the same tumor cell population, the same clones with identical genomic aberrations emerged. This proves that these changes may not be stochastic but driven by specific selective pressure.[47] Investigating and monitoring clonal dynamics and the mechanisms of the selection pressure in PDXs should improve understanding of the model system and may provide insights into how targeted treatment of clonal populations can enhance short and long term therapy response.

Credential endpoints beyond tumor burden

For evaluating therapy response in PDX the endpoints used are most frequently tumor burden (the focus of this dissertation) or the modulation of molecular targets. But the ultimate goal of cancer treatment is not merely an initial tumor response but rather decreased cancer morbidity and improved quality of life. Therefore, the rate and duration of response, including time to acquire resistance and tumor growth delay, are also important study design factors that require evaluation along the lines of the work described in Chapter 2.

Development of models with human stromal elements

PDXs do not support the maintenance of the human tumor microenvironment. Human stroma replaced by murine components as the tumor grows within the host. Interaction of the human tumor with infiltrating mouse stromal microenvironment effect the tumor's paracrine signaling which may confound results. Kupperwasser et al. have "humanized mammary fat pads" by injecting immortalized human stromal cells and breast epithelial organoids into the cleared murine mammary fat pads in an attempt to preserve the human stroma with PDX development,[98] but results are early and would need additional evaluation. Combining tumor tissue with cancer-associated fibroblasts or mesenchymal stem cells is an emerging aiming to develop a PDX microenvironment that is more representative of human tumors for growth, and metastasis.

PDX models with human immune systems

Standard PDX models are not relevant for the testing of emerging immunotherapies because the mouse hosts lack functional immune systems. Establishing specific components of the human immune system in PDX hosts can be achieved through transgenesis.[99] Others are investigating the reconstitution of a human immune system through engraftment of peripheral blood mononuclear cells (PBMCs), hematopoietic stem cells (HSCs) from bone marrow or cord blood, or fetal-derived tissue. These models have been proven successful in demonstrating anti-tumor activity in a class of T-cell immunotherapies known as chimeric antigen receptor (CAR) T-cell therapy.[58,100] However, these models present limitations; for instance, only a subset of hematopoietic lineages are populated and there are immature lymphoid structures, and further

development of these models toward more comprehensive immune systems that replicate engagement with tumors is critical for advancing the development of new immunotherapies.

Combining PDX with high-throughput screening models

Even if therapy response results of the PDX platform always accurately reflected patient responses the cost and time of screening drugs and drug combinations would remain a major disadvantage of this platform for modeling cancer therapy. Among the most promising advances in preclinical platform development are those that combine large-scale high-throughput compound screening *in vitro* with validation using *in vivo* models.[101]

Of particular promise are 3D (organoid) *in vitro* systems that have been optimized for high-throughput efficacy screening of compounds, thereby allowing entire libraries of potential pharmacologically-relevant compounds to be screened.[101] These models are used in early-stage, high-throughput screening of compounds, and have proven to be highly reproducible and cost- and time-effective. However, due to the limited histological and phenotypic similarities to primary cancer, lack of native tumor microenvironment, and limited correlation of response to patients, PDXs are better suited for later stages of compound development to confirm therapy responses. Investigating how concordant the drug responses are between 3D *in vitro* and *in vivo* models will be necessary to credential this approach but it holds great promise for improving the how well preclinical results translate into the clinic thereby advancing better cancer treatment outcomes for patients.

REFERENCES

1. Kola I, Landis J. Can the pharmaceutical industry reduce attrition rates? *Nature reviews Drug discovery.* 2004;3(8):711-715.
2. DiMasi JA, Reichert JM, Feldman L, Malins A. Clinical approval success rates for investigational cancer drugs. *Clinical pharmacology and therapeutics.* 2013;94(3):329-335.
3. Johnson JI, Decker S, Zaharevitz D, et al. Relationships between drug activity in NCI preclinical in vitro and in vivo models and early clinical trials. *British journal of cancer.* 2001;84(10):1424-1431.
4. Wilding JL, Bodmer WF. Cancer cell lines for drug discovery and development. *Cancer research.* 2014;74(9):2377-2384.
5. Gillet JP, Varma S, Gottesman MM. The clinical relevance of cancer cell lines. *Journal of the National Cancer Institute.* 2013;105(7):452-458.
6. Liu H, Patel MR, Prescher JA, et al. Cancer stem cells from human breast tumors are involved in spontaneous metastases in orthotopic mouse models. *Proceedings of the National Academy of Sciences of the United States of America.* 2010;107(42):18115-18120.
7. Whittle JR, Lewis MT, Lindeman GJ, Visvader JE. Patient-derived xenograft models of breast cancer and their predictive power. *Breast cancer research : BCR.* 2015;17:17.
8. Hidalgo M, Amant F, Biankin AV, et al. Patient-derived xenograft models: an emerging platform for translational cancer research. *Cancer discovery.* 2014;4(9):998-1013.
9. Tentler JJ, Tan AC, Weekes CD, et al. Patient-derived tumour xenografts as models for oncology drug development. *Nature reviews Clinical oncology.* 2012;9(6):338-350.
10. Rosfjord E, Lucas J, Li G, Gerber HP. Advances in patient-derived tumor xenografts: from target identification to predicting clinical response rates in oncology. *Biochemical pharmacology.* 2014;91(2):135-143.
11. Lodhia KA, Hadley AM, Haluska P, Scott CL. Prioritizing therapeutic targets using patient-derived xenograft models. *Biochimica et biophysica acta.* 2015;1855(2):223-234.
12. Lai Y, Wei X, Lin S, Qin L, Cheng L, Li P. Current status and perspectives of patient-derived xenograft models in cancer research. *J Hematol Oncol.* 2017;10(1):106.
13. Hidalgo M, Bruckheimer E, Rajeshkumar NV, et al. A pilot clinical study of treatment guided by personalized tumorgrafts in patients with advanced cancer. *Molecular cancer therapeutics.* 2011;10(8):1311-1316.
14. Garralda E, Paz K, Lopez-Casas PP, et al. Integrated next-generation sequencing and avatar mouse models for personalized cancer treatment. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2014;20(9):2476-2484.
15. Izumchenko E, Paz K, Ciznadija D, et al. Patient-derived xenografts effectively capture responses to oncology therapy in a heterogeneous cohort of patients with solid tumors. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO.* 2017;28(10):2595-2605.
16. Gao H, Korn JM, Ferretti S, et al. High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nature medicine.* 2015;21(11):1318-1325.

17. Bruna A, Rueda OM, Greenwood W, et al. A Biobank of Breast Cancer Explants with Preserved Intra-tumor Heterogeneity to Screen Anticancer Compounds. *Cell.* 2016;167(1):260-274 e222.

18. Siolas D, Hannon GJ. Patient-derived tumor xenografts: transforming clinical samples into mouse models. *Cancer research.* 2013;73(17):5315-5319.

19. Eirew P, Steif A, Khattra J, et al. Dynamics of genomic clones in breast cancer patient xenografts at single-cell resolution. *Nature.* 2015;518(7539):422-426.

20. American Cancer Society, Lifetime Risk of Developing or Dying From Cancer. 2018; https://www.cancer.gov/about-cancer/understanding/statistics. Accessed August 20, 2018, 2018.

21. National Cancer Institute, Cancer Statistics. https://www.cancer.gov/about-cancer/understanding/statistics. Accessed August 20, 2018, 2018.

22. Albrecht B, Andersen, S., Chauhan, K., Graybosch, D., and Menu, P. Pursuing breakthroughs in cancer-drug development. In. *McKinsey.* Vol 2018. www.mckinsey.com2018.

23. Buffery D. Innovation Tops Current Trends in the 2016 Oncology Drug Pipeline. *Am Health Drug Benefits.* 2016;9(4):233-238.

24. Sharpless NE, Depinho RA. The mighty mouse: genetically engineered mouse models in cancer drug development. *Nature reviews Drug discovery.* 2006;5(9):741-754.

25. DiMasi JA, Grabowski HG, Hansen RW. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of health economics.* 2016;47:20-33.

26. Croce CM. Oncogenes and cancer. *The New England journal of medicine.* 2008;358(5):502-511.

27. Society AC. History of Cancer. 2016; http://www.cancer.org/acs/groups/cid/documents/webcontent/002048-pdf.pdf.

28. Friedman AA, Letai A, Fisher DE, Flaherty KT. Precision medicine for cancer with next-generation functional diagnostics. *Nature reviews Cancer.* 2015;15(12):747-756.

29. Aronson SJ, Rehm HL. Building the foundation for genomics in precision medicine. *Nature.* 2015;526(7573):336-342.

30. Vanneman M, Dranoff G. Combining immunotherapy and targeted therapies in cancer treatment. *Nature reviews Cancer.* 2012;12(4):237-251.

31. Klein ME, Dabbs DJ, Shuai Y, et al. Prediction of the Oncotype DX recurrence score: use of pathology-generated equations derived by linear regression analysis. *Modern pathology : an official journal of the United States and Canadian Academy of Pathology, Inc.* 2013;26(5):658-664.

32. Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *The New England journal of medicine.* 2004;351(27):2817-2826.

33. *NCCN Clinical Practice Guidelines in Oncology (NCCN Guidelines).* www.nccn.org: National Comprehensive Cancer Network;2018.

34. Whiteside TL. The tumor microenvironment and its role in promoting tumor growth. *Oncogene.* 2008;27(45):5904-5912.

35. Alizadeh AA, Aranda V, Bardelli A, et al. Toward understanding and exploiting tumor heterogeneity. *Nature medicine.* 2015;21(8):846-853.

36. Binnewies M, Roberts EW, Kersten K, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nature medicine.* 2018;24(5):541-550.

37. Reina-Campos M, Moscat J, Diaz-Meco M. Metabolism shapes the tumor microenvironment. *Curr Opin Cell Biol.* 2017;48:47-53.

38. Islam F, Gopalan V, Smith RA, Lam AK. Translational potential of cancer stem cells: A review of the detection of cancer stem cells and their roles in cancer recurrence and cancer treatment. *Experimental cell research.* 2015;335(1):135-147.

39. Ledford H. US cancer institute to overhaul tumour cell lines. *Nature.* 2016;530(7591):391.

40. Fatehullah A, Tan SH, Barker N. Organoids as an in vitro model of human development and disease. *Nat Cell Biol.* 2016;18(3):246-254.

41. Evrard YA, Ahalt MMG, Alcoser SY, et al. Abstract 986: The National Cancer Institute's patient-derived models repository (PDMR). *Cancer research.* 2018;78(13 Supplement):986-986.

42. Day CP, Merlino G, Van Dyke T. Preclinical mouse cancer models: a maze of opportunities and challenges. *Cell.* 2015;163(1):39-53.

43. Wong AH, Li H, Jia Y, et al. Drug screening of cancer cell lines and human primary tumors using droplet microfluidics. *Scientific reports.* 2017;7(1):9109.

44. Frese KK, Tuveson DA. Maximizing mouse cancer models. *Nature reviews Cancer.* 2007;7(9):645-658.

45. Lee H. Genetically engineered mouse models for drug development and preclinical trials. *Biomol Ther (Seoul).* 2014;22(4):267-274.

46. Hollingshead MG. Antitumor efficacy testing in rodents. *Journal of the National Cancer Institute.* 2008;100(21):1500-1510.

47. Aparicio S, Hidalgo M, Kung AL. Examining the utility of patient-derived xenograft mouse models. *Nature reviews Cancer.* 2015;15(5):311-316.

48. Richmond A, Su Y. Mouse xenograft models vs GEM models for human cancer therapeutics. *Dis Model Mech.* 2008;1(2-3):78-82.

49. Rubio-Viqueira B, Jimeno A, Cusatis G, et al. An in vivo platform for translational drug development in pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2006;12(15):4652-4661.

50. George E, Kim H, Krepler C, et al. A patient-derived-xenograft platform to study BRCA-deficient ovarian cancers. *JCI Insight.* 2017;2(1):e89760.

51. Byrne AT, Alferez DG, Amant F, et al. Interrogating open issues in cancer precision medicine with patient-derived xenografts. *Nature reviews Cancer.* 2017;17(4):254-268.

52. Kilkenny C, Browne WJ, Cuthill IC, Emerson M, Altman DG. Improving bioscience research reporting: The ARRIVE guidelines for reporting animal research. *J Pharmacol Pharmacother.* 2010;1(2):94-99.

53. Laajala TD, Corander J, Saarinen NM, et al. Improved statistical modeling of tumor growth and treatment effect in preclinical animal studies with highly heterogeneous responses in vivo. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2012;18(16):4385-4396.

54. Liang H. Comparison of antitumor activities in tumor xenograft treatment. *Contemp Clin Trials.* 2007;28(2):115-119.

55. Wu J, Houghton PJ. Interval approach to assessing antitumor activity for tumor xenograft studies. *Pharm Stat.* 2010;9(1):46-54.

56. Hather G, Liu R, Bandi S, et al. Growth rate analysis and efficient experimental design for tumor xenograft studies. *Cancer Inform.* 2014;13(Suppl 4):65-72.

57. Migliardi G, Sassi F, Torti D, et al. Inhibition of MEK and PI3K/mTOR suppresses tumor growth but does not cause tumor regression in patient-derived xenografts of RAS-mutant colorectal carcinomas. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2012;18(9):2515-2525.

58. Shultz LD, Goodwin N, Ishikawa F, Hosur V, Lyons BL, Greiner DL. Human cancer growth and therapy in immunodeficient mouse models. *Cold Spring Harbor protocols.* 2014;2014(7):694-708.

59. Krupke DM, Begley DA, Sundberg JP, Richardson JE, Neuhauser SB, Bult CJ. The Mouse Tumor Biology Database: A Comprehensive Resource for Mouse Models of Human Cancer. *Cancer research.* 2017;77(21):e67-e70.

60. Marangoni E, Vincent-Salomon A, Auger N, et al. A new model of patient tumor-derived breast cancer xenografts for preclinical assays. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2007;13(13):3989-3998.

61. Xu S, Li S, Guo Z, Luo J, Ellis MJ, Ma CX. Combined targeting of mTOR and AKT is an effective strategy for basal-like breast cancer in patient-derived xenograft models. *Molecular cancer therapeutics.* 2013;12(8):1665-1675.

62. Nemati F, Sastre-Garau X, Laurent C, et al. Establishment and characterization of a panel of human uveal melanoma xenografts derived from primary and/or metastatic tumors. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2010;16(8):2352-2362.

63. Azmi AS, Mohammad M, Kaseb AO, Sarkar FH, Mohammed RM. Utility of Animal Models in Pancreatic Cancer Research In: Lowy AM, ed. *Pancreatic Cancer.* New York: Springer; 2008:577-599.

64. Geran R, Greenberg N, MacDonald M, Schumacher A, Abbott B. *Protocols for screening chemical agents and natural product against animal tumors and other biological systems.* Vol 3. 3rd ed1972.

65. Huynh H, Ngo VC, Fargnoli J, et al. Brivanib alaninate, a dual inhibitor of vascular endothelial growth factor receptor and fibroblast growth factor receptor tyrosine kinases, induces growth inhibition in mouse models of human hepatocellular carcinoma. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2008;14(19):6146-6153.

66. Mooney CZ, Duval RD. *Bootstrapping : a nonparametric approach to statistical inference.* Newbury Park, Calif.: Sage Publications; 1993.

67. Teicher BA. Tumor models for efficacy determination. *Molecular cancer therapeutics.* 2006;5(10):2435-2443.

68. Carter CA, Chen C, Brink C, et al. Sorafenib is efficacious and tolerated in combination with cytotoxic or cytostatic agents in preclinical models of human non-small cell lung carcinoma. *Cancer Chemother Pharmacol.* 2007;59(2):183-195.

69. Fiebig HH, Schuler J, Bausch N, Hofmann M, Metz T, Korrat A. Gene signatures developed from patient tumor explants grown in nude mice to predict tumor response to 11 cytotoxic drugs. *Cancer Genomics Proteomics.* 2007;4(3):197-209.

70. Zhang X, Lewis MT. Establishment of Patient-Derived Xenograft (PDX) Models of Human Breast Cancer. *Curr Protoc Mouse Biol.* 2013;3(1):21-29.

71. Yu J, Qin B, Moyer AM, et al. Establishing and characterizing patient-derived xenografts using pre-chemotherapy percutaneous biopsy and post-chemotherapy surgical samples from a prospective neoadjuvant breast cancer study. *Breast cancer research : BCR.* 2017;19(1):130.

72. Krepler C, Sproesser K, Brafford P, et al. A Comprehensive Patient-Derived Xenograft Collection Representing the Heterogeneity of Melanoma. *Cell reports.* 2017;21(7):1953-1967.

73. Jimeno A, Tan AC, Coffa J, et al. Coordinated epidermal growth factor receptor pathway gene overexpression predicts epidermal growth factor receptor inhibitor sensitivity in pancreatic cancer. *Cancer research.* 2008;68(8):2841-2849.

74. Garrido-Laguna I, Uson M, Rajeshkumar NV, et al. Tumor engraftment in nude mice and enrichment in stroma- related gene pathways predict poor survival and resistance to gemcitabine in patients with pancreatic cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research.* 2011;17(17):5793-5800.

75. Wang X, Fu X, Hoffman RM. A new patient-like metastatic model of human lung cancer constructed orthotopically with intact tissue via thoracotomy in immunodeficient mice. *International journal of cancer Journal international du cancer.* 1992;51(6):992-995.

76. Loukopoulos P, Kanetaka K, Takamura M, Shibata T, Sakamoto M, Hirohashi S. Orthotopic transplantation models of pancreatic adenocarcinoma derived from cell lines and primary tumors and displaying varying metastatic activity. *Pancreas.* 2004;29(3):193-203.

77. Zhao X, Liu Z, Yu L, et al. Global gene expression profiling confirms the molecular fidelity of primary tumor-based orthotopic xenograft mouse models of medulloblastoma. *Neuro-oncology.* 2012;14(5):574-583.

78. Zhang X, Claerhout S, Prat A, et al. A renewable tissue resource of phenotypically stable, biologically and ethnically diverse, patient-derived human breast cancer xenograft models. *Cancer research.* 2013;73(15):4885-4897.

79. DeRose YS, Wang G, Lin YC, et al. Tumor grafts derived from women with breast cancer authentically reflect tumor pathology, growth, metastasis and disease outcomes. *Nature medicine.* 2011;17(11):1514-1520.

80. Heo EJ, Cho YJ, Cho WC, et al. Patient-Derived Xenograft Models of Epithelial Ovarian Cancer for Preclinical Studies. *Cancer Res Treat.* 2017;49(4):915-926.

81. Kabos P, Finlay-Schultz J, Li C, et al. Patient-derived luminal breast cancer xenografts retain hormone receptor heterogeneity and help define unique estrogen-dependent gene signatures. *Breast cancer research and treatment.* 2012;135(2):415-432.

82. Ben-David U, Ha G, Tseng YY, et al. Patient-derived xenografts undergo mouse-specific tumor evolution. *Nature genetics.* 2017;49(11):1567-1575.

83. Li H, Zhu Y, Tang X, et al. Integrated analysis of transcriptome in cancer patient-derived xenografts. *PloS one.* 2015;10(5):e0124780.

84. Ma XJ, Dahiya S, Richardson E, Erlander M, Sgroi DC. Gene expression profiling of the tumor microenvironment during breast cancer progression. *Breast cancer research : BCR.* 2009;11(1):R7.

85. Ding L, Ellis MJ, Li S, et al. Genome remodelling in a basal-like breast cancer metastasis and xenograft. *Nature.* 2010;464(7291):999-1005.

86. Reyal F, Guyader C, Decraene C, et al. Molecular profiling of patient-derived breast cancer xenografts. *Breast cancer research : BCR.* 2012;14(1):R11.

87. Ananda G, Mockus S, Lundquist M, et al. Development and validation of the JAX Cancer Treatment Profile for detection of clinically actionable mutations in solid tumors. *Exp Mol Pathol.* 2015;98(1):106-112.

88. Woo XY, Srivastava A, Graber JH, et al. Bioinformatics workflows for genomic analysis of tumors from Patient Derived Xenografts (PDX): challenges and guidelines. *bioRxiv.* 2018.

89. Adzhubei IA, Schmidt S, Peshkin L, et al. A method and server for predicting damaging missense mutations. *Nature methods.* 2010;7(4):248-249.

90. Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC. SIFT missense predictions for genomes. *Nature protocols.* 2016;11(1):1-9.

91. Zerbino DR, Achuthan P, Akanni W, et al. Ensembl 2018. *Nucleic acids research.* 2018;46(D1):D754-D761.

92. Jones P, Binns D, Chang HY, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics.* 2014;30(9):1236-1240.

93. Stenson PD, Ball EV, Mort M, Phillips AD, Shaw K, Cooper DN. The Human Gene Mutation Database (HGMD) and its exploitation in the fields of personalized genomics and molecular evolution. *Curr Protoc Bioinformatics.* 2012;Chapter 1:Unit1 13.

94. Landrum MJ, Lee JM, Benson M, et al. ClinVar: improving access to variant interpretations and supporting evidence. *Nucleic acids research.* 2018;46(D1):D1062-D1067.

95. Patterson SE, Liu R, Statz CM, Durkin D, Lakshminarayana A, Mockus SM. The clinical trial landscape in oncology and connectivity of somatic mutational profiles to targeted therapies. *Hum Genomics.* 2016;10:4.

96. Petrillo LA, Wolf DM, Kapoun AM, et al. Xenografts faithfully recapitulate breast cancer-specific gene expression patterns of parent primary breast tumors. *Breast cancer research and treatment.* 2012;135(3):913-922.

97. Romanelli A, Clark A, Assayag F, et al. Inhibiting aurora kinases reduces tumor growth and suppresses tumor recurrence after chemotherapy in patient-derived triple-negative breast cancer xenografts. *Molecular cancer therapeutics.* 2012;11(12):2693-2703.

98. Kuperwasser C, Chavarria T, Wu M, et al. Reconstruction of functionally normal and malignant human breast tissues in mice. *Proceedings of the National Academy of Sciences of the United States of America.* 2004;101(14):4966-4971.

99. Boudreau JE, Liu XR, Zhao Z, et al. Cell-Extrinsic MHC Class I Molecule Engagement Augments Human NK Cell Education Programmed by Cell-Intrinsic MHC Class I. *Immunity.* 2016;45(2):280-291.

100. Walsh NC, Kenney LL, Jangalwe S, et al. Humanized Mouse Models of Clinical Disease. *Annu Rev Pathol.* 2017;12:187-215.

101. Pauli C, Hopkins BD, Prandi D, et al. Personalized In Vitro and In Vivo Cancer Models to Guide Precision Medicine. *Cancer discovery.* 2017;7(5):462-477.

APPENDIX

Table A-1 Study group data of cisplatin treated PDX models

Model	Primary Site	Specimen Site	Clinical Diagnosis	Primary/Met	Passage	Study Duration	Control				Treatment			
							N	Average Start Size (mm3)	Min Start Size (mm3)	Max Start Size (mm3)	N	Average Start Size (mm3)	Min Start Size (mm3)	Max Start Size (mm3)
J000079689	Retroperitoneum	Retroperitoneum	neuroblastoma	Metastatic	2	28	5	220.6	103.2	326.9	6	180.3	100.8	272.5
J000080739	Breast	Breast	invasive ductal carcinoma	Primary	2	21	10	143.1	72.3	286.4	10	141.3	73.0	264.1
J000096652	Lung	Lung	lung adenocarcinoma	Primary	6	21	4	147.0	80.7	207.6	6	150.5	89.6	221.3
J000099327	Breast	Breast	invasive ductal carcinoma	Recurppendixrent	3	14	10	138.3	70.3	262.9	10	137.8	78.7	293.5
J000100646	Bladder	Bladder	urinary bladder cancer	Primary	3	28	10	130.9	95.7	223.4	10	131.3	96.9	221.4
J000100672	Lung	Lung	lung adenocarcinoma	Metastatic	4	28	8	137.2	81.8	226.6	7	137.9	76.9	212.5
J000100674	Breast	Breast	invasive ductal carcinoma	Primary	3	21	11	144.9	77.1	272.4	11	142.4	81.6	215.1
J000100675	Breast	Breast	invasive ductal carcinoma	Primary	4	28	9	138.9	73.0	228.8	11	139.0	74.8	243.1
J000101173	Breast	Breast	invasive ductal carcinoma	Primary	3	28	9	124.0	68.2	208.1	9	121.5	82.7	222.0
J000102184	Breast	Breast	Invasive ductal carcinoma	Primary	4	28	8	118.7	86.2	204.4	8	109.1	71.7	138.8
J000102630	Colon	Colon	colon adenocarcinoma	Metastatic	4	28	8	129.5	93.0	194.6	7	130.7	94.5	159.3
J000102680	Colon	Colon	colon adenocarcinoma	Metastatic	3	28	8	133.2	104.5	166.5	8	133.3	99.9	195.9
J000103634	Breast	Breast	invasive ductal carcinoma	Primary	3	28	9	117.9	70.3	218.6	11	118.8	78.6	189.1
J000103917	Breast	Breast	invasive ductal carcinoma	Primary	4	28	4	134.4	92.1	209.3	6	145.1	91.6	244.4
J000104256	Bladder	Bladder	urinary bladder cancer	Primary	4	21	10	169.6	99.5	295.0	10	169.6	86.1	293.0
J000104518	Bladder	Bladder	urinary bladder urothelial carcinoma	Primary	5	28	10	126.6	93.2	194.9	9	126.8	90.1	203.6
J000105006	Bladder	Bladder	squamous cell carcinoma	Primary	3	28	9	102.9	80.7	165.1	9	101.7	82.4	147.5
J000106560	Skin	Skin	Melanoma	Primary	8	28	9	135.2	96.5	197.4	9	136.9	108.9	193.4
TM00090	Breast	Breast	breast metaplastic carcinoma	Primary	3	21	2	143.4	76.4	237.9	4	130.0	94.3	169.3
TM00091	Breast	Breast	invasive ductal carcinoma	Primary	3	28	8	149.7	73.9	258.5	8	142.7	78.8	218.4
TM00096	Breast	Breast	invasive ductal carcinoma	Metastatic	3	28	8	144.2	63.4	229.3	9	145.2	74.6	256.3

TM00097	Breast	Breast	invasive ductal carcinoma	Primary	3	28	9	139.0	72.8	218.5	2	143.0	81.3	227.1
TM00098	Breast	Breast	invasive ductal carcinoma	Primary	3	28	9	121.5	73.9	167.4	9	121.4	75.1	161.2
TM00099	Breast	Breast	invasive ductal carcinoma	Primary	3	21	9	151.6	71.5	224.0	9	150.4	75.2	227.4
TM00103	Breast	Breast	invasive ductal carcinoma	Primary	4	28	9	121.3	91.0	191.4	10	118.0	95.5	148.4
TM00107	Breast	Breast	invasive ductal carcinoma	Recurrent	4	21	8	152.6	82.3	228.3	8	153.6	97.4	205.3
TM00185	Lung	Lung	lung squamous cell carcinoma	Primary	3	28	7	148.3	71.3	252.3	7	155.9	80.8	296.8
TM00186	Lung	Lung	lung adenocarcinoma	Primary	3	21	7	157.7	74.7	281.6	8	157.7	77.8	223.3
TM00188	Lung	Lung	lung squamous cell carcinoma	Primary	3	28	8	172.0	71.8	288.3	8	175.4	83.5	286.0
TM00192	Lung	Lung	lung adenocarcinoma	Primary	5	21	10	132.9	75.1	236.1	10	131.4	77.5	218.7
TM00193	Lung	Lung	lung adenocarcinoma	Primary	5	28	8	135.4	71.1	239.3	8	141.6	88.4	210.5
TM00194	Lung	Lung	lung small cell carcinoma	Metastatic	4	28	8	110.2	78.4	167.8	8	107.2	79.1	156.0
TM00196	Lung	Lung	lung adenocarcinoma	Primary	4	28	5	175.0	102.1	278.9	6	163.4	76.6	262.3
TM00199	Lung	Lung	lung adenocarcinoma	Primary	5	28	10	110.9	73.8	219.6	10	110.5	72.7	155.2
TM00202	Lung	Lung	lung adenocarcinoma	Metastatic	4	28	7	111.9	67.5	259.7	7	127.0	73.9	177.2
TM00203	Lung	Lung	lung adenocarcinoma	Metastatic	3	28	8	129.8	88.8	209.5	8	126.0	92.4	177.4
TM00208	Lung	Lung	lung squamous cell carcinoma	Primary	3	28	10	161.8	71.8	295.6	10	160.7	87.2	257.2
TM00212	Lung	Lung	large cell neuroendocrine carcinoma	Primary	3	28	6	124.6	59.2	264.5	8	126.0	77.2	195.1
TM00213	Lung	Lung	lung adenocarcinoma	Primary	3	28	6	147.2	75.4	281.9	6	133.7	88.1	218.3
TM00214	Lung	Lung	lung adenocarcinoma	Primary	3	28	9	184.7	96.6	293.3	10	183.7	96.9	280.4
TM00219	Lung	Lung	lung adenocarcinoma	Metastatic	6	28	9	104.6	86.1	142.3	9	103.7	81.5	134.0
TM00222	Lung	Lung	pleomorphic carcinoma	Primary	3	28	6	139.5	61.1	265.5	4	137.2	71.4	232.4
TM00226	Lung	Lung	lung adenocarcinoma	Primary	3	28	10	186.0	84.3	295.5	10	188.9	96.8	294.2
TM00231	Lung	Lung	lung squamous cell carcinoma	Primary	3	21	10	174.6	92.4	276.0	20	173.9	92.7	274.1
TM00233	Lung	Lung	lung adenocarcinoma	Primary	3	28	7	145.8	92.7	271.7	6	144.6	80.5	234.4
TM00246	Lung	Lung	lung mixed small cell and squamous cell carcinoma	Primary	4	28	7	134.4	72.2	241.4	7	137.6	106.4	229.5
TM00253	Lung	Lung	lung adenocarcinoma	Metastatic	3	28	10	116.3	81.6	184.2	10	115.1	86.6	167.6

TM00256	Lung	Lung	lung squamous cell carcinoma	Primary	6	28	10	99.6	71.6	160.3	10	98.5	77.1	138.6
TM00298	Prostate gland	Prostate gland	adenocarcinoma	Primary	8	28	9	178.8	100.2	287.3	9	170.9	104.4	272.4
TM00302	Lung	Lung	lung adenocarcinoma	Metastatic	5	28	8	105.2	76.4	166.0	9	84.5	50.0	171.5
TM00335	Ovary	Ovary	papillary serous adenocarcinoma	Primary	6	28	10	112.9	77.3	173.2	10	112.7	74.2	154.4
TM00355	Lung	Lung	lung squamous cell carcinoma	Primary	4	28	9	104.4	75.2	227.9	9	103.3	69.7	149.0
TM00356	Lung	Lung	lung adenocarcinoma	Metastatic	3	21	7	121.8	74.7	164.2	8	136.1	80.5	227.1
TM00386	Breast	Breast	invasive ductal carcinoma	Primary	6	28	8	128.6	80.7	205.0	8	124.8	101.9	193.3
TM00387	Kidney	Kidney	renal cell carcinoma	Metastatic	4	28	9	109.1	70.7	171.6	10	108.7	71.3	150.7
TM00702	Skin	Skin	melanoma	Metastatic	7	28	8	136.4	93.4	239.8	8	136.0	99.9	220.0
TM00784	Lung	Lung	lung adenocarcinoma	Primary	5	14	10	191.4	102.2	274.5	10	190.1	109.3	251.7
TM00832	Lung	Lung	lung squamous cell carcinoma	Recurrent	3	28	10	170.0	127.7	213.6	10	168.9	137.3	202.3
TM00877	Lung	Lung	lung adenocarcinoma	Primary	5	28	10	122.4	77.0	191.6	10	124.6	78.6	229.7
TM00999	Breast	Breast	invasive ductal carcinoma	Primary	4	28	8	102.6	62.2	213.5	8	102.3	64.4	162.7
TM01029	Bladder	Bladder	invasive bladder transitional cell carcinoma	Primary	4	14	8	154.5	70.6	258.4	8	152.0	75.9	234.9
TM01039	Kidney	Kidney	renal cell carcinoma	Primary	4	21	3	116.6	73.9	196.7	4	116.8	82.8	188.4
TM01075	Colon	Colon	adenocarcinoma	Primary	3	28	9	144.6	88.8	200.3	8	145.5	90.7	204.3
TM01079	Breast	Breast	invasive ductal carcinoma	Primary	4	28	9	119.1	71.1	189.8	10	117.8	77.5	176.4
TM01087	Adrenal gland	Adrenal gland	neuroblastoma	Metastatic	3	28	8	140.4	70.2	257.8	8	139.5	75.3	244.0
TM01117	Breast	Breast	invasive ductal carcinoma	Primary	4	28	9	127.2	84.7	195.5	9	130.8	93.7	199.2
TM01149	Skin	Skin	melanoma	Metastatic	9	21	9	138.5	83.2	230.0	10	138.9	89.9	229.4
TM01273	Breast	Breast	invasive ductal carcinoma	Primary	4	14	11	154.2	78.1	302.4	11	156.0	71.1	264.5
TM01278	Breast	Breast	invasive ductal carcinoma	Primary	3	28	10	136.8	77.7	231.2	11	139.4	81.8	213.1
TM01563	Lung	Lung	lung squamous cell carcinoma	Primary	4	28	8	126.4	70.0	242.8	8	126.4	70.6	171.9

Table A-2 Study group data of docetaxel treated PDX models

Model	Primary Site	Specimen Site	Treatment Naïve?	Passage	Study Duration	Control **N**	Treatment **N**
J000100675	Breast	Breast	Yes	4	28	9	11
J000101173	Breast	Breast	No	3	28	9	9
J000102184	Breast	Breast	No	4	28	8	8
J000103634	Breast	Breast	No	3	28	9	11
J000100646	Bladder	Bladder	Yes	3	28	10	10
TM00199	Lung	Lung	No	5	28	10	10
TM00208	Lung	Lung	Yes	3	28	10	10
TM00226	Lung	Lung	Yes	3	28	10	10
TM00253	Lung	Lung	No	3	28	10	10
TM00256	Lung	Lung	Yes	6	28	10	10
TM00335	Ovary	Ovary	Yes	6	28	10	10
TM00832	Lung	Lung	No	3	28	10	10
TM00877	Lung	Lung	Yes	5	28	10	10
TM00091	Breast	Breast	No	3	28	8	8
TM00214	Lung	Lung	No	3	28	9	10
TM00387	Kidney	Kidney	No	4	28	9	10
TM00096	Breast	Breast	No	3	28	8	9
J000104518	Bladder	Bladder	Unknown	5	28	10	9
TM00098	Breast	Breast	Yes	3	28	9	9
J000105006	Bladder	Bladder	Unknown	3	28	9	9
J000106560	Skin	Skin	Yes	8	28	9	9
TM00103	Breast	Breast	No	4	28	9	10
TM00219	Lung	Lung	No	6	28	9	9
TM00298	Prostate gland	Prostate gland	No	8	28	9	9
TM00355	Lung	Lung	Yes	4	28	9	9
TM00386	Breast	Breast	No	6	28	8	8
TM00302	Lung	Lung	No	5	28	8	9
TM01075	Colon	Colon	Yes	3	28	9	8
TM00999	Breast	Breast	Yes	4	28	8	8
J000102680	Colon	Colon	No	3	28	8	8
TM01079	Breast	Breast	Yes	4	28	9	10
TM01117	Breast	Breast	Yes	4	28	9	9
TM00188	Lung	Lung	Yes	3	28	8	8
TM00193	Lung	Lung	Yes	5	28	8	8
TM00194	Lung	Lung	No	4	28	8	8
TM00203	Lung	Lung	No	3	28	8	8
TM00702	Skin	Skin	Yes	7	28	8	8
TM01278	Breast	Breast	No	3	28	10	11
TM01087	Adrenal gland	Adrenal gland	Yes	3	28	8	8
TM01563	Lung	Lung	Yes	4	28	8	8
TM00212	Lung	Lung	Yes	3	28	6	8
J000100672	Lung	Lung	Yes	4	28	8	7

J000102630	Colon	Colon	No	4	28	8	7
TM00185	Lung	Lung	Yes	3	28	7	7
TM00202	Lung	Lung	Yes	4	28	7	7
TM00246	Lung	Lung	Yes	4	28	7	7
TM00233	Lung	Lung	No	3	28	7	6
TM00213	Lung	Lung	Yes	3	28	6	6
J000079689	Retroperitoneum	Retroperitoneum	0	2	28	5	6
TM00196	Lung	Lung	Yes	4	28	5	6
J000103917	Breast	Breast	No	4	28	4	6
TM00222	Lung	Lung	Yes	3	28	6	4
TM00097	Breast	Breast	Yes	3	28	9	2

BIOGRAPHY OF THE AUTHOR

Joan Malcolm was born in Bangor, Maine on April 1, 1984. She was raised in Gardiner, Maine and graduated from Gardiner Area High School in 2002 and the University of Maine in 2007 with a Bachelor's degree in Biological Engineering. She began her career at The Jackson Laboratory where she worked in a variety of different positions and at all three U.S. sites. She spent seven years in project management and five years in business management within diverse clinical and biomedical research programs.

Joan met her husband, Joshua, in 2014 while working at the JAX Sacramento facility and currently lives a charmed life with him and their two dogs, Peanut and Buster, in Menlo Park, California. She has recently joined a startup and will continue to focus her career on advancing clinical applications of genomic technologies to improve human health. Joan is a candidate for the Doctor of Philosophy degree in Biomedical Sciences and Engineering from the University of Maine in December 2018.