





# What's in our translational science toolbox?

Preclinical oncology model selection in the genomic era – a biotech's perspective

- Current problems and opportunities; what have we learnt from past experience?
- First-time right? Delivering POC studies to inform the path to the clinic
- From EMT to TME: exploiting data-driven model selection (exemplified by case studies)



Current problems and opportunities - what have we learnt from past experience?

# **Evaluating efficacy** – 10 years on



### **BIOBUSINESS BRIEFS**

TRIAL WATCH

NATURE REVIEWS | DRUG DISCOVERY VOLUME 10 | MAY 2011 | 1

### Phase II failures: 2008–2010

Well-conducted Phase II clinical trials provide the data required to determine whether there is a case to be made, both scientifically and commercially, for progressing a drug candidate into Phase III trials, At present, however. Phase II success rates are lower than at any other phase of development. Analysis by the Centre for Medicines Research (CMR) of projects from a group of 16 companies (representing approximately 60% of global R&D spending) in the CMR International Global R&D database reveals that the Phase Il success rates for new development projects have fallen from 28% (2006-2007) to 18% (2008-2009), although these success rates do vary between therapeutic areas and between small molecules and biologics. As the current likelihood of a drug successfully progressing through Phase III to launch is 50% (Nature Rev. Drug Discov. 10, 87; 2011), the overall attrition of late-stage drug development seems to be unsustainably high.

To help understand these trends, Thomson Reuters Life Science Consulting analysed the 108 reported Phase II failures from 2008 to 2010 for new drugs and major new indications of existing drugs (Drug News Perspect. 22, 39-51; 2009; Drug News Perspect. 23, 48-63; 2010; Drugs Today, 47, 27-51; 2011). Out of these, 87 reported the reasons for failure (FIG. 1a): 51% (44 out of 87) were due to insufficient efficacy, 29% (25 out of 87) were due to strategic reasons and 19% (17 out of 87) were due to clinical or preclinical safety reasons. Out of the 25 failures that were terminated for strategic reasons, 16 involved validated targets such as peroxisome proliferator activated receptor-y (PPARy) and factor Xa, therefore suggesting that some

of these failures were due to inadequate differentiation from more advanced drugs in the same class or from drugs with similar indications in another mechanistic class. Out of the 21 failures for which reasons were not reported, 17 involved validated targets, although not always in an approved indication for drugs affecting that target. Again, it would seem reasonable to conclude that some of these failures were due to insufficient evidence of an efficacy advantage over a more advanced drug; however, it is important not to rule out that failure could be due to the change in the benefit-risk balance of a known target in a new patient population. These data also show that 68% (73 out of 108) of failures fell into four therapeutic areas (FIG. 1b): alimentary/ metabolism, cancer, cardiovascular, and neuroscience. Notably, 61% (14 out of 23) of failures in alimentary/metabolism are for

Although it is difficult to draw conclusions from these data, the finding that a substantial proportion of Phase II failures were due to strategic reasons suggests that one important underlying factor could be overlapping R&D activity between companies with drugs in Phase II trials. This raises the question of whether an increase in collaborative efforts between companies up to the point of proof-of-concept for novel targets or mechanisms might be more cost- and

John Arrowsmith is a Scientific Director at Thomson Reuters, 77 Hatton Garden, London EC1N 8JS, UK.

Efficacy

Pharmacokinetics

/bioavailability 1%

Biostatistics (2019) 20, 2, pp. 273-286 doi:10.1093/biostatistics/kxx069 Advance Access publication on January 31, 2018

### Estimation of clinical trial success rates and related parameters

### CHI HEEM WONG, KIEN WEI SIAH

MIT Computer Science and Artificial Intelligence Laboratory & Department of Electrical Engineering and Computer Science, Cambridge, MA 02139, USA and MIT Sloan School of Management and Laboratory for Financial Engineering, Cambridge, MA 02142, USA

### ANDREW W. LO\*

MIT Computer Science and Artificial Intelligence Laboratory & Department of Electrical Engineering and Computer Science, Cambridge, MA 02139, USA, MIT Sloan School of Management and Laboratory for Financial Engineering, Cambridge, MA 02142, USA, and AlphaSimplex Group, LLC, Cambridge, MA 02142, USA

alo-admin@mit.edu

### SUMMARY

curated by the pharmaceutical industry and are subject to potential selection biases. Using a sample of 406 038 entries of clinical trial data for over 21 143 compounds from January 1, 2000 to October 31, 2015, we estimate aggregate clinical trial success rates and durations. We also compute disaggregated estimates across several trial features including disease type, clinical phase, industry or academic sponsor, biomarker presence, lead indication status, and time. In several cases, our results differ significantly in detail from widely cited statistics. For example, oncology has a 3.4% success rate in our sample vs. 5.1% in prior studies. However, after declining to 1.7% in 2012, this rate has improved to 2.5% and 8.3% in 2014 and 2015, respectively. In addition, trials that use biomarkers in patient-selection have higher overall success probabilities than trials without biomarkers.

Keywords: Clinical phase transition probabilities; Clinical trial statistics; Probabilities of success.

# Previous estimates of drug development success rates rely on relatively small samples from databases

Strategic **b** Other (35) 29% 32% ■ Alimentary/ metabolism (23) ■ Cancer (21) 20% ■ Neuroscience (17) Cardiovascular (12)

Figure 1 | Phase II failures: 2008-2010. The 108 failures are divided according to reason for failure when reported (87 drugs) (a) and therapeutic area (b).

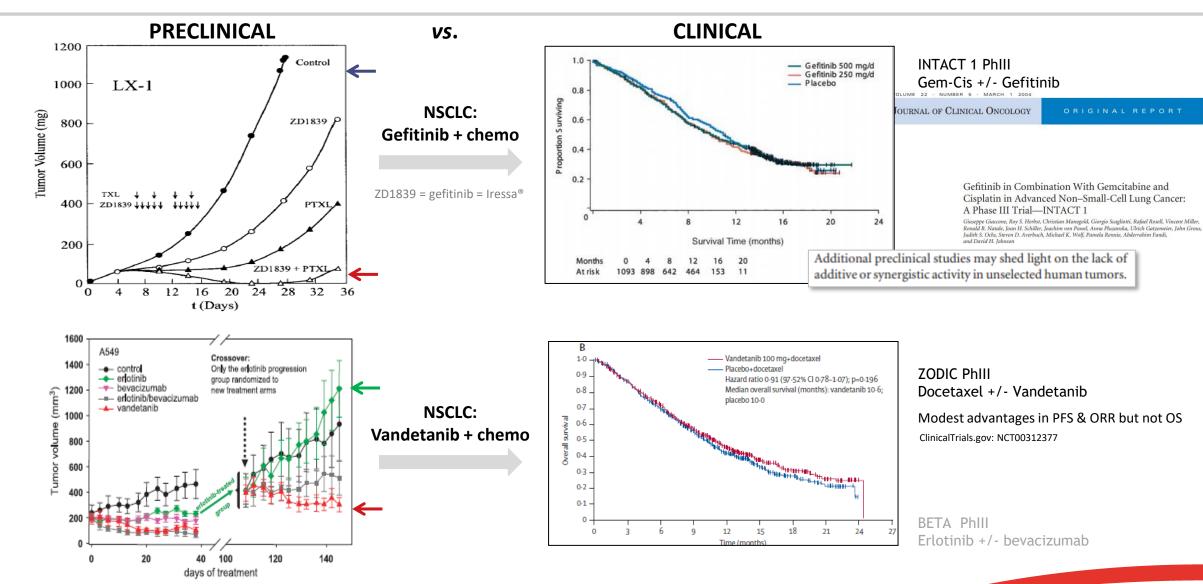
### Probability of Success<sup>2</sup> by Clinical Trial Phase and Therapeutic Area

	P1 to P2	P2 to P3	P3 to Approval	Overall
Oncology	57.6	32.7	35.5	3.4
Metabolic/Endocrinology	76.2	59.7	51.6	19.6
Cardiovascular	73.3	65.7	62.2	25.5
Central Nervous System	73.2	51.9	51.1	15.0
Autoimmune/Inflammation	69.8	45.7	63.7	15.1
Genitourinary	68.7	57.1	66.5	21.6
Infectious Disease	70.1	58.3	75.3	25.2
Ophthalmology	87.1	60.7	74.9	32.6
Vaccines (Infectious Disease)	76.8	58.2	85.4	33.4
Overall	66.4	48.6	59.0	13.8
Overall (Excluding Oncology)	73.0	55.7	63.6	20.9

Source: Chi Heem Wong, Kien Wei Siah, Andrew W Lo. "Estimation of clinical trial success rates and related parameters." Biostatistics 20(2): April 2019, Pages 273-286, Published online: 31 January 2018, DOI: 10.1093/biostatistics/kxx069

# **Problem:** Lack of historical correlation?





# Important to translate back: Lessons from FLEX Trial



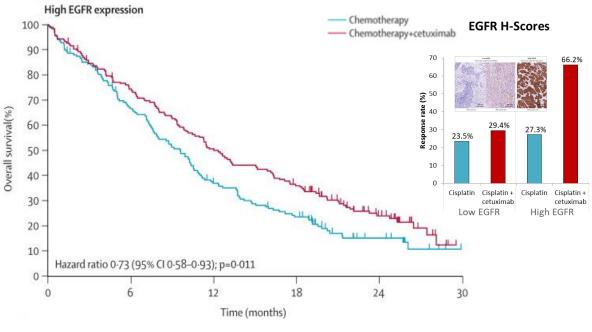
ClinicalTrials.gov number NCT00148798

### FLEX CLINICAL TRIAL

# Chemotherapy — Chemot

### Figure taken from Pirker et al., The Lancet Oncology (2012) 13, 33-42

## FLEX CLINICAL TRIAL (retrospective analysis)



ClinicalTrials.gov number NCT00148798

- No survival benefit to cisplatin + cetuximab treatment observed
- Response to cisplatin + cetuximab treatment significantly associated with high EGFR expression (retrospectively)
- NSCLC PDX data confirmed benefit of combination but only when EGFR H-score was re-assessed (cut-off ≥ 200)... A case of 'informed predictability'?

### Believe it or not: how much can we rely on published data on potential drug targets?

Florian Prinz, Thomas Schlange and Khusru Asadullah

house target identification campaigns, in-licensing and public sourcing, in particular based on reports published in the Bierature and preserred at conference. During the transfer of creticets from an academic to a commany of projects from an academic to a company

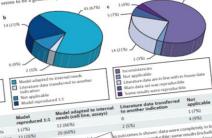
A recent report by Arronomish noted that the success rates for new development projects in practice for the success rates for new development projects in for pursuing a fault blown drug discovery and 1 grant II trials have fallen from 28% to 18% in development programme for a particular tart. The success years, with insufficient efficacy being recent years, with insufficient efficacy being recent years.

nous sources, including in-nification campaigns, in-projects that were started in our company could not be reproduced. Talking to scientists, both in academia and in industry, there

ablic recognition (for example, see REPS 2.5), public recognition (for example, see MEFS 2.5), and the surprisingly few scientific publica-tions dealing with this topic. Indeed, to our knowledge, so far there has been no published

try, with a dedicated budget and scientists who mainly work on target validation to increase recent year, with insufficient efficacy being the most frequent reason for failure ("Plane II allures") on the properties of the properties of the most frequent reason for failure ("Plane II allures") on the properties of the pr in To mitigate some of the risks of such irrests:

times ultimately being warded, most pharmacutical companies run in-house trayer validation programmer. However, validation programmer, However, validation programmer, the programmer of the progra



Editorial: preclinical data reproducibility for R&D -

zantoriai: preciinicai data reprodi the challenge for neuroscience

knowledge, so far there can been no plantament in-depth, systematic analysis that compares reproduced results with published results for

data obtained and their relationship to the pub-

IS THERE A REPRODUCIBILITY

> A Nature survey lifts the lid on how researchers view the 'crisis' rocking science and what they think will help.

> > BY MONYA BAKER

Yes, a significant

38% Yes, a slight

7% Don't know

1,576 RESEARCHERS SURVEYED EDITORIAL

No, there is no crisis

# 4 swan in the making

incentives, culture, and tool sets

"Achieving reproducible

movement to change the

of Riffyn, Inc., in Oakland, CA, and an advisor to the European Union Scientific E-mail: tg@riffy research...will take a collective



OPEN Improving reproducibility improving reproductivity splitting in animal research by splitting the study population into several mini-experiments



First-time right? Delivering POC studies to inform the path to the clinic

# History: Olaparib (Lynparza®)

# From discovery to first-in-class clinical practice and the models in between



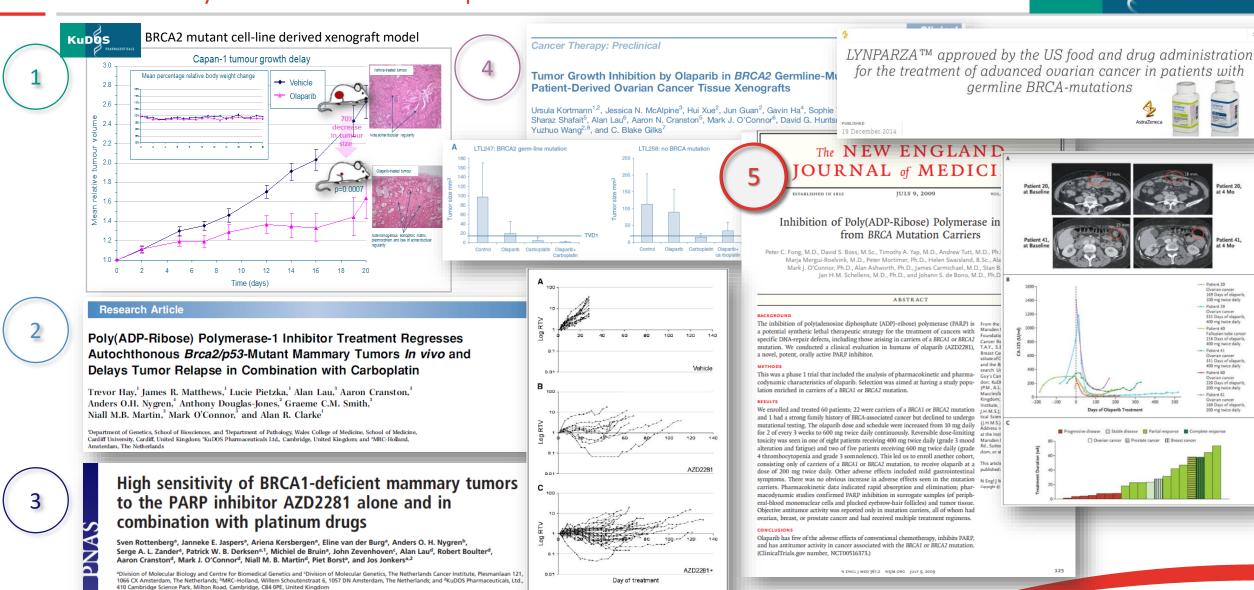


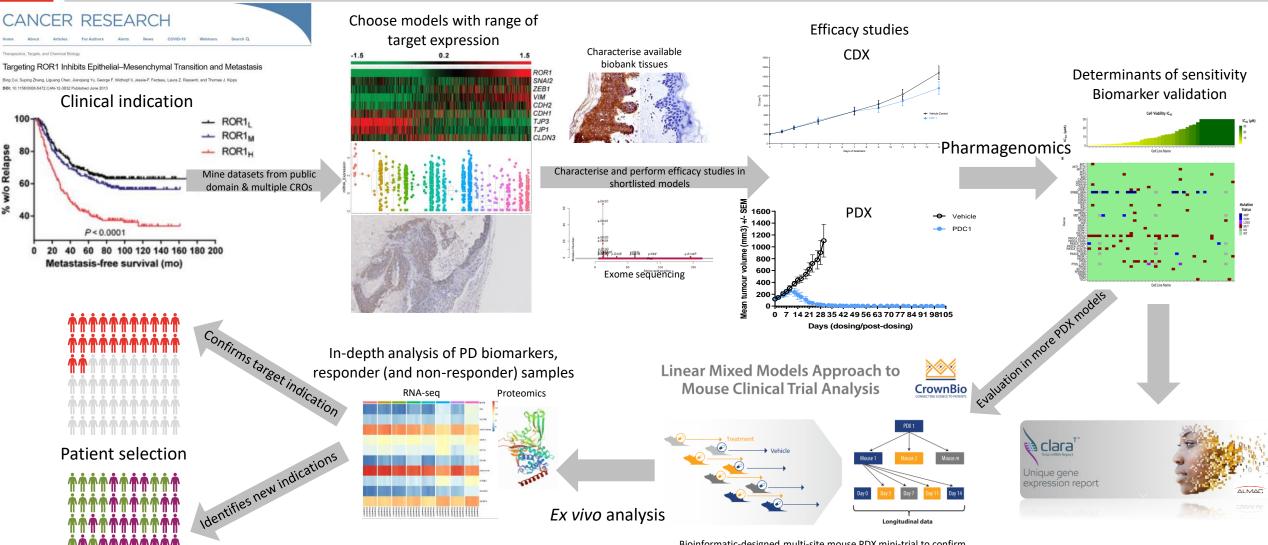
Figure 2. Long-term responses of Brca2/p53-deficient tumors from mice treate daily with 50 mg/kg AZD2281. A, 28 d of daily vehicle treatment. B, 28 d of daily 50 mg/kg AZD2281 treatment. C, continuous daily 50 mg/kg AZD2281 treatment. X axis, day of treatment; Y axis, log RTV; each line represents

the response of an individual tumor

# **POC Studies to inform the Path to Clinic**

# From model short-listing to patient selection



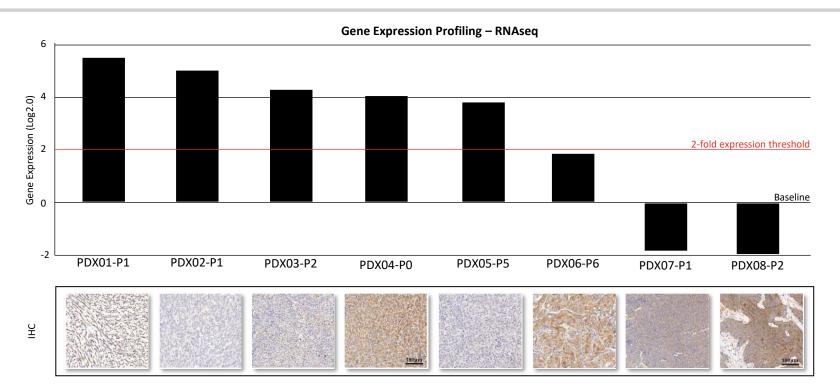


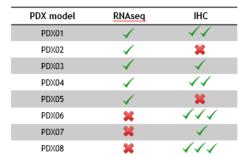
Bioinformatic-designed multi-site mouse PDX mini-trial to confirm additional indications and refine patient selection approach

# RNAseq screening as a gate to select models – a cautionary tale



- Models screened for target expression by RNAseq and ranked according to gene expression score (shown opposite); data kindly supplied gratis by CRO
- Confirmatory IHC screen conducted to validate model choice (shown opposite)
- Concordance of model selection approaches by RNASeq and IHC is less than 40%. Raises questions about validity of model selection on a single method





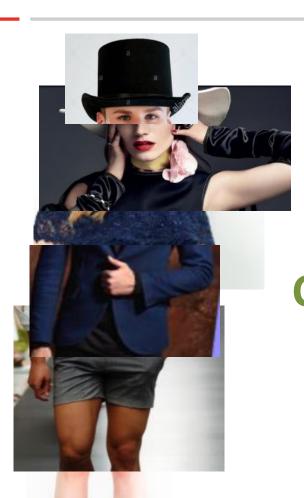
Data shown is real-world data but PDX model IDs are redacted to protect confidential company intellectual property

RNA-IHC concordance = 37.5% (insufficient confidence)

Could be explained by heterogeneity of PDX/sampling, cross-reactivity of IHC Ab for human and mouse target expression, cross-reactivity with another [unknown] protein or simply that RNA levels do not translate to protein expression

# Building a better picture: the right model for the right question





**GEMMs** 

orthotopics

humanised models
co-cultures Syngeneics

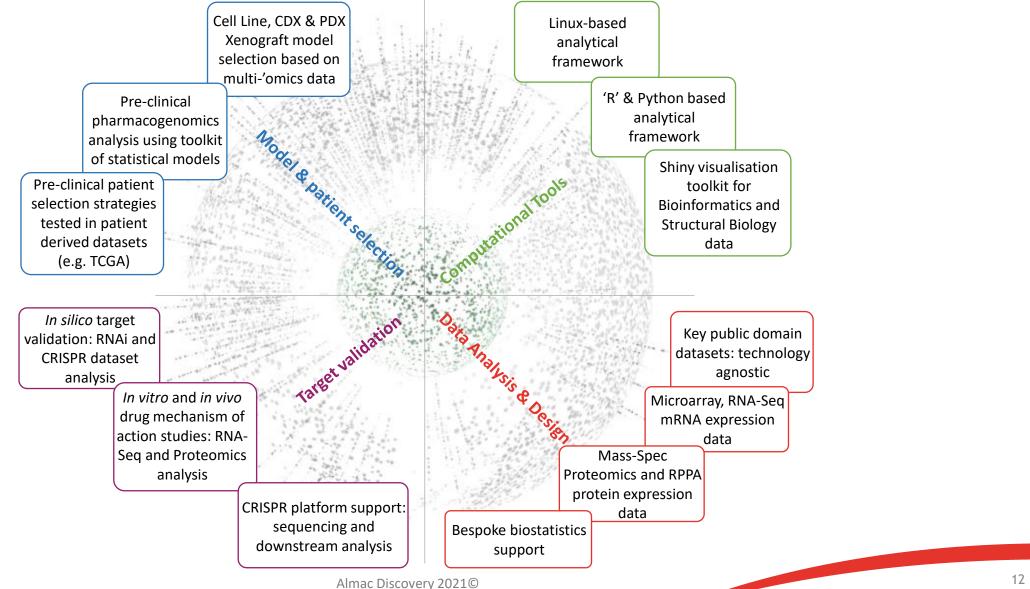
CRISPR

Organoids/PDXOs

COX

# **Systems Biology Toolbox:** connecting the dots to translate discoveries to cures



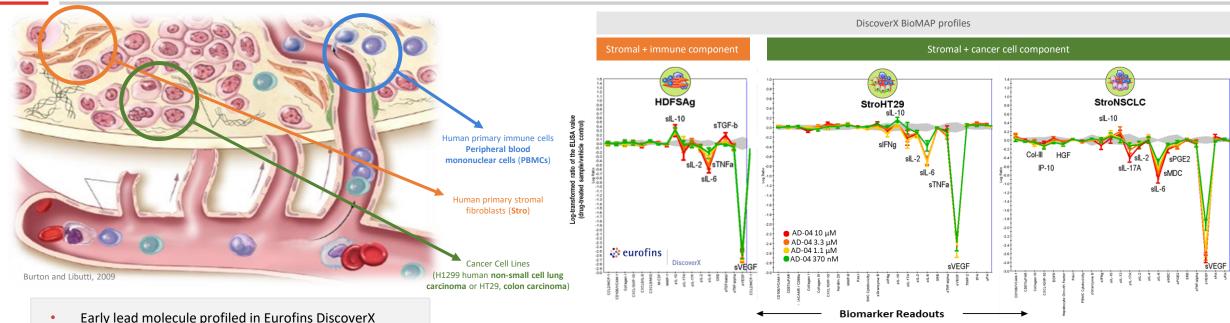




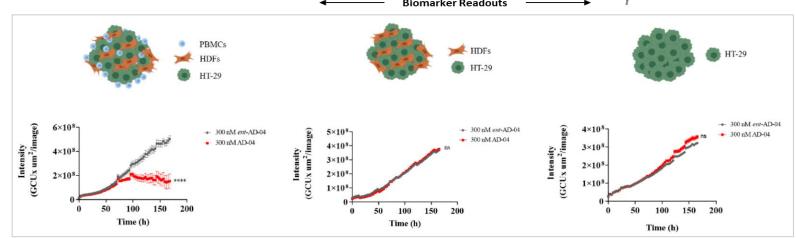
# From EMT to TME: exploiting data-driven model selection

# Deconstructing the Tumour Microenvironment:





- Early lead molecule profiled in Eurofins DiscoverX BioMAP co-culture biomarker discovery system
- Inhibitor generated very striking biological marker profiles when evaluated in complex co-culture cellular systems
- Strong modulation of sVEGF observed across dose range in HDF + PBMC and HDF + cancer cell panels
- Suggests biological mechanism is mediated via the stromal cell component
- Validation of the biological MOA was achieved in-house in spheroid [triple] co-culture systems → → →



AD-04 = USP7 inhibitor: ent-AD-04 = inactive enantiomer

# Reconstructing the Tumour Microenvironment:



### TME models:

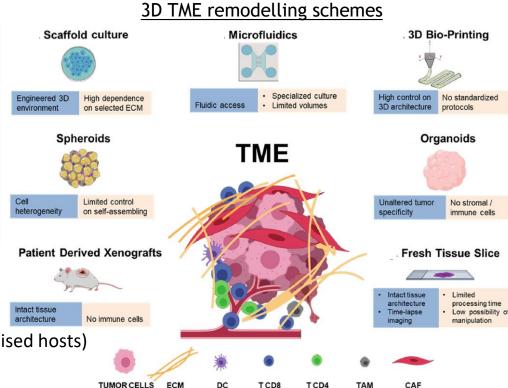
### 2D culture options:

- The use of conditioned medium
- The use of transwell techniques
- Seeding cancer cells on fibroblasts

### 3D culture options:

- Colony assay in soft agar or Matrigel
- Spheroids
- Organoids, PDXO

# DI Modugno et al. Journal of Experimental & Clinical Cancer Research 1,2019;38:117 https://doi.org/10.1186/s13046-019-1086-2 REVIEW Open Access 3D models in the new era of immune oncology: focus on T cells, CAF and ECM Francesca Di Modugno ", Cristina Colosi", Paola Trono", Giuseppe Antonacd", Giancario Ruocco" and Paola Nistico" Abstract Immune checkpoint inhibitor therapy has changed clinical practice for patients with different cancers, since these agents have demonstrated a significant improvement of overall survival and are effective in many patients. However, an intrinsic or acquired resistance frequently occur and biomarkers predictive of responsiveness should help in patient selection and in defining the adequate treatment options. A deep analysis of the complexity of the tumor microenvironment is likely to further advance the field and hopefully identify more effective combined immunotherapeutic strategies. Here we review the current knowledge on tumor microenvironment, focusing on T cells, cancer associated fibroblasts and estracellular matrix. The use of 3D cell culture models to resemble tumor microenvironment factograph and the proprior of the proprior of the complexity of the tumor microenvironment, Immune oncology, 3D culture models, T cells, Cancer associated fibroblasts, Extracellular matrix



### Figure taken from: Di Modugno et al., 2019. J Exp & Clin Can Res 38:117

### In vivo models:

- Syngeneics (limitations: mouse immune system, limited tumours)
- CDX (limitations: homogeneous, drift/selection on plastic, immunocompromised hosts)
- PDX (limitations: immunocompromised hosts, [stromal] tissue drift)
- Humanised PDX (fewer limitations? but expensive)
- GEMMs (limitations: mouse immune system, slow to grow)

### Goals

To develop a 3D high-throughput assay that allows us to investigate the infiltration, activation and function of immune cells, and tumour cell cytotoxicity upon different immune checkpoint inhibitors/compound treatments

In-house development of tumour immune microenvironment (TIME), using tumour cell lines with or without primary dermal fibroblasts (HDFs) or cancer-associated fibroblasts (CAFs) co-cultured with human immune cells in non-adherent growth conditions

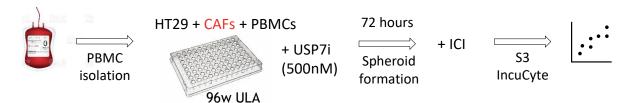
# Reconstructing the Tumour Microenvironment (ex vivo):

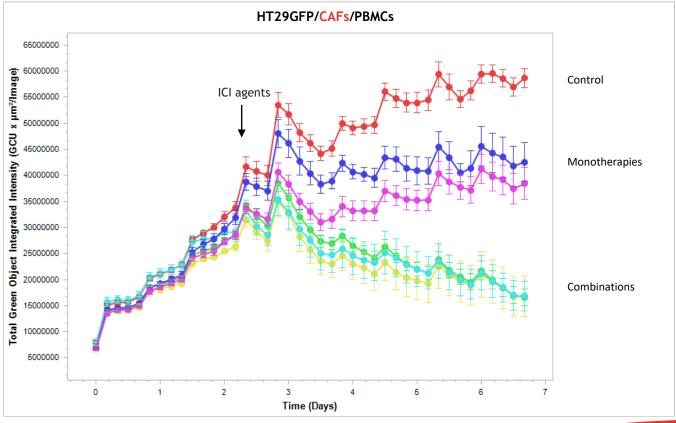


### **3D Spheroid Tri-culture Assays for Compound Profiling:**

- 3D spheroid tri-culture of cancer cells, CAFs and PBMCs
- 96-well plate format; 1 spheroid per well
- Treat with test agents (mono or combo)
- Measure in real time on IncuCyte





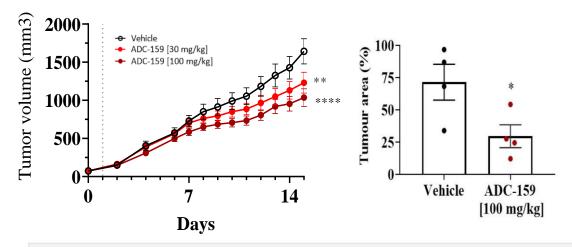


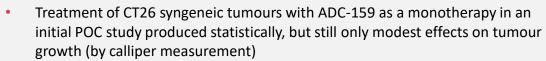
HT29GFP + CAF + PBMCs (freshly isolated PBMCs) + ICI agent(s)

# Syngeneics, [serendipity] and systems biology

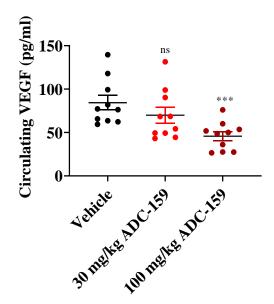


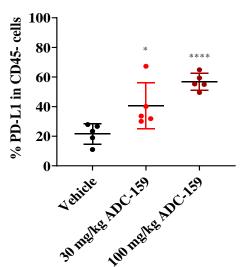
Helping to connect the dots to translate from discoveries to cures

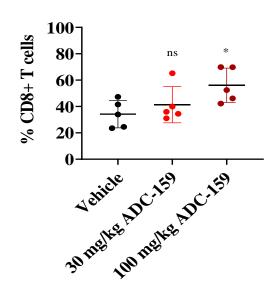




- However, analysis of the tumours revealed remodelling of the TME on multiple levels
- Analysis of circulating VEGF revealed a dose-dependent decrease with ADC-159 treatment (as predicted by the DiscoverX co-culture assay)
- FACS analysis of the treated CT26 tumours demonstrated an increase in CD8<sup>+</sup> T cells suggesting a modulation of the tumour microenvironment
- Also of note, we observed a dose-dependent increase in the level of PD-L1 on CD45<sup>-</sup> cells (endothelial, fibroblasts and cancer cells)









ADC-159 = oral USP7 inhibitor

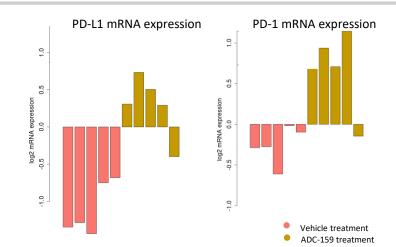
# Syngeneics, [serendipity] and systems biology

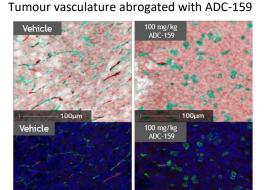


# Helping to connect the dots to translate from discoveries to cures

- Analysis of CT26 syngeneic tumours treated with ADC-159 (an orally bioavailable version of AD-04) at the transcriptomics level confirmed the FACS findings
- Immune checkpoint molecules, PD-1 and PD-L1 are transcriptionally upregulated following treatment with ADC-159
- These finding suggest that a immunologically cold tumour could be sensitised to an agent targeting the PD-(L)1 axis
- CT26 tumours were treated with ADC-159, an anti-PD-L1 agent or the two agents in combination and tumour growth followed over time

As hypothesised, an increase in PDL-1 expression, infiltrating CD8+ T cells, and reduced VEGF levels resulted in improved responses to an anti PD-L1 agent. Provides *in vivo* POC that treatment with ADC-159 can modulate the tumour microenvironment to turn immunologically cold tumours, warm

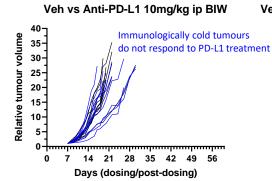


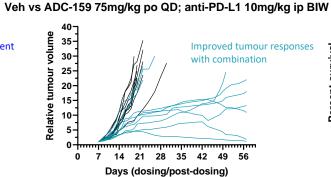


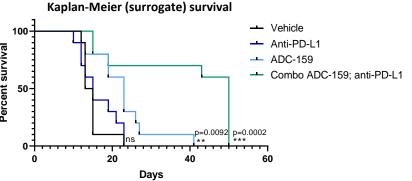
CD31 endothelial vessel marker; NG2 pericyte marker



Confirmation of ↑ in tumour infiltrating cells by *Ultivue* IO multiplex panel (*planned*)



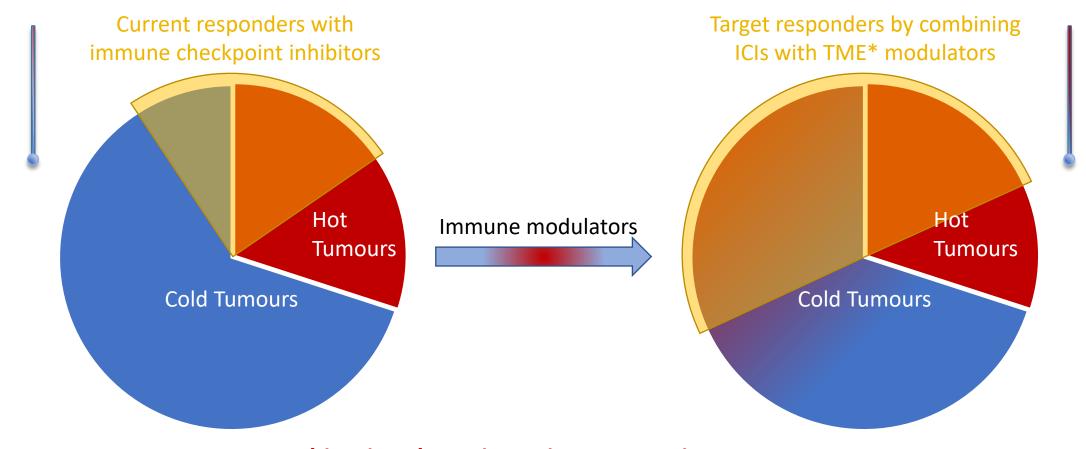




In addition, the combination treatment was very well tolerated with no body weight loss or adverse clinical signs observed over a 7-week dosing period

# Turning up the Temperature in the Tumour Microenvironment



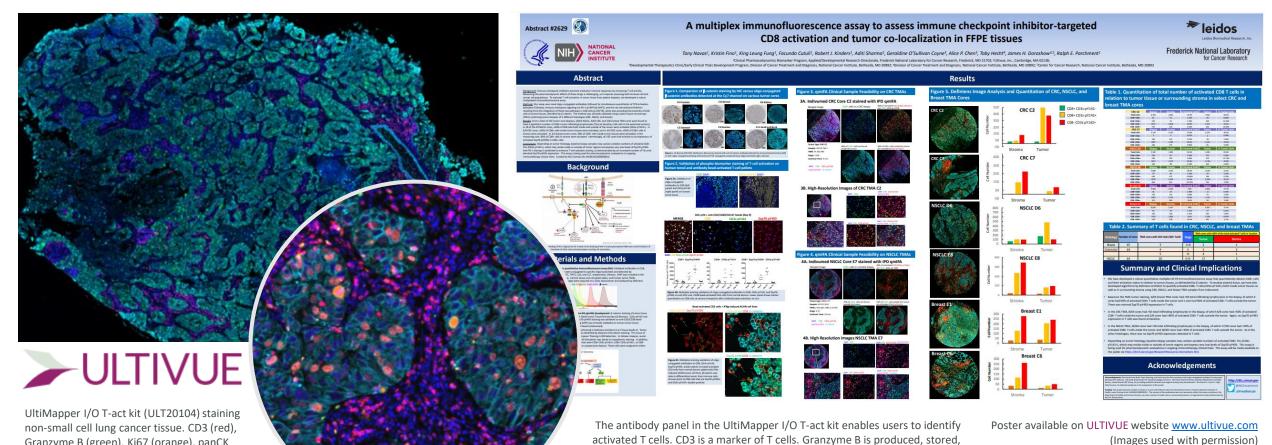


# **Combination therapies to improve patient responses**

• Unmet need in 2<sup>nd</sup> line NSCLC: <20% response rate in a PD-L1 unselected population (Borghaei *et al.*, 2015 NEJM 373:1627-1639; Nishio *et al.*, 2015, 33:15, 8027)

# Turning up the Temperature in the Tumour Immune Microenvironment





Granzyme B (green), Ki67 (orange), panCK (cyan), and nuclear counterstain (blue).

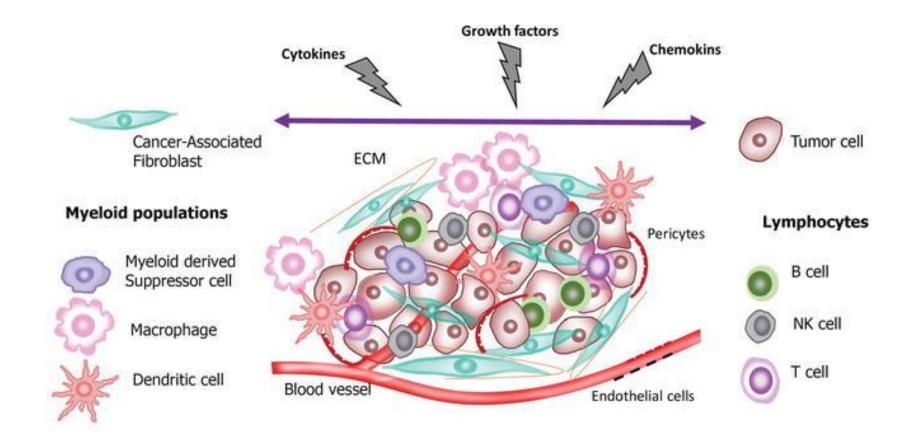
Ultivue's multiplex IF panels enable

T cells. CD3 is a marker of T cells. Granzyme B is produced, stored, and released by cytotoxic T cells to mediate cellular death in combination with perforin. Ki67 is a nuclear protein that is a marker of proliferation. Co-expression of CD3, Granzyme B, and Ki67 indicates proliferating cytotoxic T cells. SOX10 is a tumour marker for melanomas while panCK detects carcinomas (provided in a cocktail).

simultaneous detection of multiple cell types involved in immune and TME biology
Offers an innovative and powerful approach for visualising and quantitating TIME modulation in preclinical (and clinical) IO studies

# **Novel Target Rationale**

Summary: Reprogramming the Tumour Microenvironment (TME)



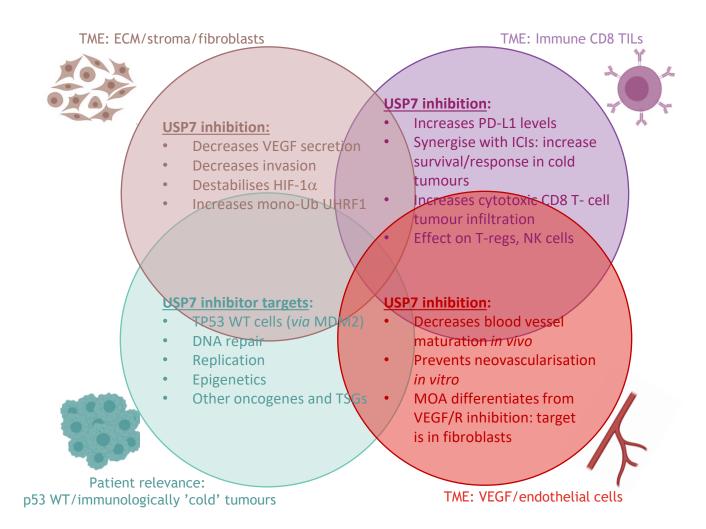
How do we convert the complexity of the tumour microenvironment into actionable/druggable biology?

# **Novel Target Rationale**



22

Summary: Reprogramming the Tumour Microenvironment (TME)



USP7 = ubiquitin specific protease 7

Almac Discovery 2021©

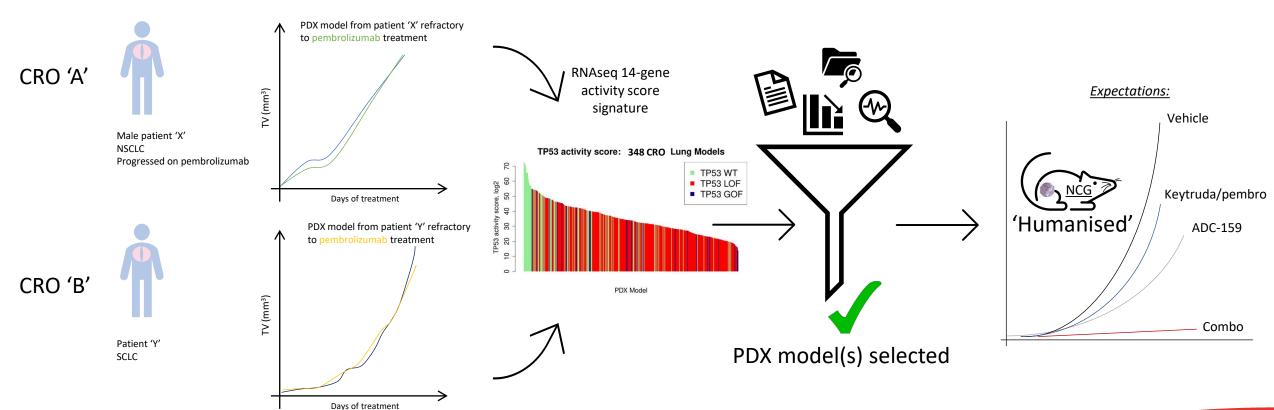
# Looking Ahead...



## Building a Tumour Microenvironment (TME) model for POC studies

### **Models at CROs short-listed based on:**

- Ability of CRO to perform humanisation esp. CD34 human umbilical cord blood
- No response to ICI (potential to potentiate)
- p53 activity score (in-house RNAseq 14-gene signature)
- Cold TME/low immune cell infiltrates (potential to turn cold tumours warm)



<sup>\*</sup>Model information and response data for illustration only; used with permission but redrawn to protect confidentiality

# Looking further ahead...



### New Horizons in Translational Medicine

Volume 2, Issue 1, September 2014, Pages 1-4



Research Articles

# Systems Patientomics: The virtual *in-silico* patient

D.V. Dimitrov 🖾 🕀

Show more >

+ Add to Mendeley 📽 Share 🍠 Cite

https://doi.org/10.1016/j.nhtm.2014.08.002

Get rights and content

- We will undoubtedly continue to see improvements in host strains and humanisation approaches
- Datasets for 'hot' areas e.g. 2<sup>nd</sup>/3<sup>rd</sup> gen ADCs, CAR-T, IO cold/hot will continue to grow and inform predictability/utility of models
- There is a need for reverse-translation from clinical trials to better understand failings (but appetite among biotechs & pharma companies is generally low) - see slide S3 for example of where clinical success is informing preclinical models for IBD
- Likely to see a step-change in computational and systems biology (e.g. AI/ML) approaches and application to in silico patient avatar models



# Acknowledgements

### **Ubi-Plex® DUB Drug Discovery Platform**

Xavier Jacq Ph.D. (VP – Biology) Anamarija Jurisic Ph.D. Peggy Sung Ph.D.

### **PDCs & Protein Therapeutics Group**

Graham Cotton Ph.D. (Head)
Estelle McLean Ph.D. (Technical Leader)
Jen Thom Ph.D. (Technical Leader)

Licensing enquiries to Alan.Lamont@almacgroup.com

# **Bioinformatics & Computational Biology**

Mark Wappett (Group Leader) Wei-Wei Kung Ph.D.

Prof. Tim Harrison (VP – Drug Discovery), Colin O'Dowd Ph.D. (Director - Medicinal Chemistry), Matt Helm Ph.D. (Team Leader), Shane Rountree Ph.D. (Technical Leader)





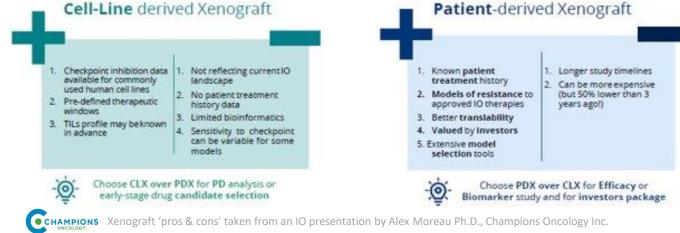


# **Supplemental**

# S1. What have I learnt...?



- Important to cross-validate models and/or approaches take the time to do this! It seems obvious, but presence of target and disease biology is important
- Challenge your assumptions (challenge/stress-test published data)
- Work with the CRO to interrogate and understand the model don't be afraid to ask!
- Perform power analyses upfront; ask CRO for historical data to help with this
- In our biotech world, we are mostly interested in demonstrating in vivo proof-of-concept effects in a couple of models (foothold/gain confidence in molecule/target biology → POC/hypothesis testing)
- Typically, on any given project, these models are complementary (each with their pros & cons):
  - High vs low target expression
  - Simple vs complex (syngeneic vs humanised PDX)
  - Overexpression vs KO
- Models most often used for:
  - Demonstrating efficacy
  - Modulation of target/MOA/downstream biological effects
  - Some toxicology readouts
- When possible, build a package of efficacy data in different models including different indications &/or types of model as this builds confidence (although not necessarily predictability)
- Just as 'one swallow does not make a summer', one good xenograft result does not make a successful clinical trial

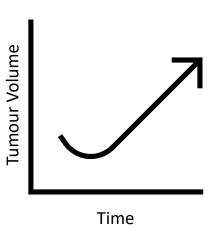


https://www.brighttalk.com/webcast/18129/444843?utm\_source=Champions+Oncology&utm\_medium=brighttalk&utm\_campaign=444843

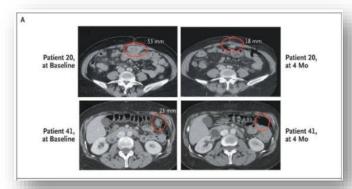
# **S2. Evaluating efficacy** - Interpreting results, building confidence



- Preclinically, anti-tumour effects are often reported as changes in tumour volume over time (relative TV, absolute TV, mean, median, individuals, T/C, %TGI, mRECIST)
- Lack of standardisation/harmonisation between groups (although most published M&M are reasonably clear enabling some comparisons)
- Many published papers cite %TGI as the sole readout and it is often calculated on a day when results look best or on the day when the vehicle group is terminated; need for unbiased analyses/reporting (present best effect, biggest difference, best models)
- Some groups have tried to align tumour volume data from mouse xenograft studies with clinical RECIST\* criteria. Tumour shrinkage and time to disease progression are important readouts to align with clinical parameters, objective response, time to event (progression), survival (note: measurements e.g. 5mm increase over baseline = progression, and inclusion of lymph node assessment is technically limited in mouse studies)
- Justice et al., 2014 and Gao et al., 2015 have presented the case for harmonising the criteria, analyses and reporting of mouse efficacy data (Best Average Response)
- One positive xenograft study ≠ a positive clinical trial result! A bank of positive data in different models provides increasing confidence maybe even some predictability
- Given the complexity and many different cancers/patient disease state, there is unlikely to be a single 'best' model or approach – use the model that best replicates the disease, mechanism or target to build confidence



\*RECIST = Response Evaluation Criteria In Solid Tumours



Ultimately though, species differences and heterogeneity of many patients in a clinical trial with different disease biology, subtypes and stages of disease are most likely to drive the major contribution to the differences observed between clinical and non-clinical results

# **S3. Preclinical Models of IBD** – Clinical success informing preclinical models



29

- In the field of Inflammatory Bowel Disease (IBD), tremendous progress has been achieved with the approval of anti-TNFs, integrin blockers, JAK inhibitors, anti-IL-12/23 agents for treating patients, with S1P antagonists potentially following soon
- However, the success in patients has actually highlighted the flaws in the predictability of preclinical models
- None of the approved treatments are efficacious in all of the rodent models of IBD
- In fact, many of the approved biologic treatments work only in one of the many rodent models
- The choice of the rodent model can avoid false negatives and add predictive value for preclinical drug discovery projects
- It seems obvious, but choice of model should represent the disease biology