www.crukscotlandinstitute.ac.uk/philip_dunne

PHENOTYPIC PLASTICITY IN COLORECTAL CANCER

Group Leader
Philip Dunne

Research Scientists
Natalie Fisher[1]
Sudhir Malla[2]
Raheleh Amirkhah[3]
Shania Corry[4]

Graduate Students
Ryan Byrne
Courtney Bull
Jessica Edwards

[1]CRUK Early Detection & INCISE
[2]MRC National Mouse Genetics Network
[3]CRUK ACRCelerate programme & All Ireland Cancer Research Institute
[4]Left in 2024 for University of Oxford

QUEEN'S UNIVERSITY BELFAST

Our research team are focussed primarily on tumour-based discovery in colorectal cancer (CRC), utilising large collections of both human and preclinical mouse tumour tissue to interrogate molecular and morphological features to gain a better understanding of disease. Our work integrates molecular biology, computational analysis, and translational pathology, particularly in early stage localised disease, to uncover aggressive traits driving metastasis, and therapeutic resistance mechanisms, that in-turn can be used to guid more subtype-specific treatment strategies.

Phenotypic landscape of colorectal cancer.

Our research in 2024 led to the development of the pathway-derived subtyping (PDS) system, (Malla *et al.*, 2024, *Nat Genet*) which has offered previously unseen insights into tumour biology and patient outcomes. Unlike traditional gene-level classifiers, PDS used phenotypic signalling and pathway-level data to identify three subtypes: **PDS1**: Characterised by LGR5+ stem-rich, highly proliferative tumours with good prognosis, **PDS2**: Enriched for immune and stromal signalling, featuring ANXA1+ regenerative traits, and most interestingly **PDS3**: A previously overlooked slow-cycling subtype with reduced stem-like properties, increased differentiated lineages (e.g., enterocytes), and poor clinical outcomes. PDS3 tumours are characterised by transcriptional repression of many features previously defined as being essential hallmarks of colorectal cancer, while conversely exhibiting high levels of canonical epithelial differentiation reminiscent of a normal homeostatic balance and also more subtle traits that align with a neuro or squamous-like cell identities, all unique signalling profiles distinct from existing CRC subtyping systems (Figure 1).

The study also highlights the clinical relevance of transcriptional pathways, such as MYC targets (enriched in PDS1 and PDS2) and polycomb repressive complex (PRC) targets (elevated in PDS3). This axis reflects a stem-to-differentiation spectrum, with PDS3 tumours exhibiting diminished proliferation and stemness. Using disease-positioning to align these human tumour traits with a series of genetically engineered mouse models (GEMMs) from the Sansom lab, while PDS1 and PDS2 align well with these GEMMs, our study identifies PDS3 as being the least represented in these preclinical models and a range of *ex vivo* organoid models, which limits therapeutic exploration and underscores its clinical challenges (Figure 2). Prognostic analysis across large CRC cohorts confirmed the subtype's association with poor relapse-free survival, even when controlling for KRAS mutation status or microsatellite stability. This work emphasises the utility of integrating pathway-level subtyping with existing frameworks to capture overlooked tumour heterogeneity across stem and differentiated lineage states (Figure 3).

Disease positioning of human and mouse tumours.

Driven partly by the absence of robust disease-positioned PDS3 models described above, Owen Sansom's lab have generated a more comprehensive "gold standard" GEMM transcriptional cohort using n=1,200 samples from across a range of developmental stages (normal, precancer/short-term, cancer, metastasis) in both tissue and organoid and across a wide variety of genotypes (20+ genotypes, including combinations of conventional v serrated drivers). This unique resource has allowed our teams to begin to assess, within a heterogeneous collection of lesions, how mouse tumours recapitulate the phenotypic features observed during the development and progression of human CRC. Remarkably, in these preliminary analyses, using an unsupervised clustering method across all tissue samples has revealed a transcriptional pattern that appears to align with the well-understood disease trajectory of both the adeno-carcinoma sequence and, to a certain extent, the progression from non-invasive primary lesions to metastatic lesions.

In 2024, we performed a series of preliminary analyses, to directly address some of the outstanding criticisms of using GEMMs, regarding how well mouse models represent the range of important molecular and clinical subtypes, alongside how they recapitulate the phenotypic heterogeneity seen in human tumours. Using these data, in combination with a range of dual-species classifiers developed by our team, we have successfully disease-positioned mouse tumours according to human CMS, iCMS and the PDS approach. In addition to finding genotypes that consistently returned the same human tumour subtype, we also identified several fixed genotypes that were classified across a range of molecular subtypes (CMS, iCMS and PDS) and stem cell states. This is similar to the intrinsic and microenvironmental phenotypic heterogeneity observed in human tumours.

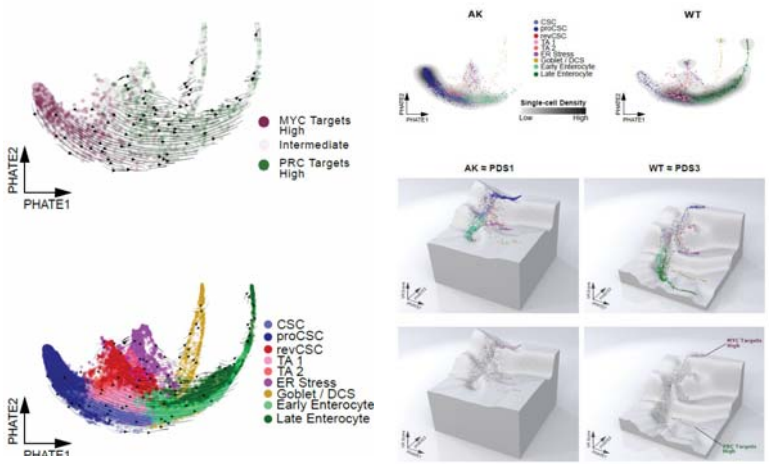

Figure 1. PDS signalling is associated with differentiation along an axis of colon epithelial lineage identities.

Figure 2: When projected onto a Waddington landscape, PDS signalling is associated with stem-like pluripotency and cell entropy, where AK genotypes (Apc and Kras mutant) drive stem cell features reminiscent of PDS1, whereas PDS3 maintains a more normal/WT balance of lineage identities.

Data analytic applications and analysis portals

Generation of transcriptional data has dramatically increased in the past decade, driving the development of analytical algorithms that enable interrogation of the biology underpinning the profiled samples. However, these resources require users to have expertise in data wrangling and analytics, reducing opportunities for biological discovery by 'wet-lab' users with a limited programming skillset. Although commercial solutions exist, costs for software access can be prohibitive for academic research groups. To address these challenges, we have developed a series of open source and user-friendly data analysis platforms for on-the-fly bioinformatic interrogation of transcriptional data by any user, regardless of bioinformatics skillsets, we have been developing a series of data applications to support our commitments to the FAIR principles.

Subtype Explorer (SubtypeExploreR, https://subtypeexplorer.qub.ac.uk). To complement the non-exhaustive PDS characterisations we presented in the *Nature Genetics* study, we developed the 'SubtypeExploreR' platform. This enables any user to interrogate transcriptional genes and/or signatures, including existing signatures from numerous databases or an unlimited combination of *de novo* unpublished classifiers, according to three different molecular CRC subtyping, including, PDS, CMS and iCMS across the bulk cohorts used in that study.

Molecular Subtyping Resource (MouSR, https://mousr.qub.ac.uk/). This internet-accessible analytical tool enables users to easily interrogate their data using an intuitive 'point-and-click' interface, which includes a suite of molecular characterisation options including quality control, differential gene expression, gene set enrichment and microenvironmental cell population analyses from RNA sequencing. The MouSR online tool provides a unique freely available option for users to perform rapid transcriptomic analyses and comprehensive interrogation of the signalling underpinning transcriptional datasets, which alleviates a major bottleneck for biological discovery.

Publications listed on page 120

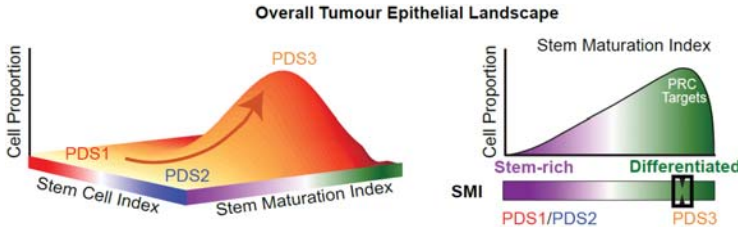

Figure 3: The colorectal phenotypic landscape can be defined using PDS from both a bulk and single cell viewpoint using a combination of signalling related to stem cell polarisation and epithelial cell differentiation status.