From genotype to phenotype: leveraging functional genomics in oncology

A challenge in drug discovery is that, at best, only around one in five drug candidates will lead to the launch of a medicine. In 2011, AstraZeneca initiated a strategic ‘5R Framework’ focusing on identifying drug projects in consideration of the right target, right tissue, right safety profile, right patient and right commercial potential.

Although we have significantly enhanced the quality of our pipeline since bringing in the 5R model and our success rate is markedly higher than the industry average, the reality is that around 80% of our projects still fail before candidate selection, often resulting from a poor understanding of the relationship between target biology and disease. Therefore, understanding exactly what is going on at a fundamental biological level, so-called target validation, is essential if we are to boost the quality of our drug discovery pipeline and the chances of commercial success.

The advent of next-generation DNA sequencing has led to the creation of vast quantities of genomic data from individuals, tissues and tumours. Through AstraZeneca’s Genomics Initiative and the work of many other researchers around the world, we now have access to an enormous number of DNA sequences that we can use to inform the development of molecularly-targeted therapies, support patient selection for clinical trials and match patients to the therapies most likely to benefit them.

Drug discovery is increasingly informed by genomics, which can significantly boost the chances of a novel therapy successfully navigating the lengthy journey from bench to bedside. But despite the rapid proliferation of genomic information, we still have a limited understanding of how specific genes and genetic variations drive tumour growth and therapeutic resistance. In turn, this limits our ability to identify and validate novel targets that are truly informed by biology and make an impact on disease progression – arguably the most crucial step in the whole drug discovery pathway.

In search of answers, we have established a major collaboration with Cancer Research UK to launch the Functional Genomics Centre in Cambridge, UK. Using the power of gene editing technology, we are investigating the connections between the information encoded within the genes (genotype) and its functional output (phenotype), systematically altering every single gene in the

By Dr Ultan McDermott, Dr Steve Rees, Dr Susan Galbraith, Dr Mike Snowden and Dr Mene Pangalos
genome across a wide range of cell types to see how it affects growth and function or contributes to therapeutic resistance.

**From gene to function**

Our initiative builds on the approach that has been used by geneticists for more than a century as they seek to uncover the link between genotype and phenotype: create a change in the genome then observe the effect. Initially this was done by mutagenesis with chemicals or X-rays, becoming more sophisticated through the use of genetic engineering techniques such as transgenic (‘knock-out’) cell and animal models. Both these approaches have significant drawbacks, including the challenge of tracing a resulting phenotype back to the underlying gene fault in the case of mutagenesis, or the high cost and low efficiency of generating knockouts.

Things began to change around a decade ago with the development of TALENs and zinc finger genome editing tools, which could be used to create highly-targeted DNA alterations at specific sequences. The field made a quantum leap in 2012 with the demonstration that the bacterial DNA cleavage system CRISPR/Cas9 (usually known as CRISPR) could be targeted to any location in the genome of human cells with much higher specificity, allowing researchers to manipulate sequences relatively quickly, cost-effectively and with high efficiency. Since its discovery, CRISPR has proved to be a powerful tool for unpicking the complex connections between genotype and phenotype, particularly in cancer research.

At AstraZeneca, we were quick to adopt gene editing technologies into our drug discovery programmes, altering specific genes to generate more physiological cellular and animal models of disease. Moreover, it is now possible to use these tools at scale, using genome-wide libraries of CRISPR guide RNAs to alter every gene in the human genome in normal or cancer cells, screening 20,000 genes in a single experiment. Running these assays in combination with high-throughput robotics and phenotypic analysis techniques allows for massively parallel screening, dramatically accelerating the identification of novel genes involved in biological processes.

**Up, down and out**

We will initially use genome-wide CRISPRn libraries generated by our collaborators at the Wellcome Sanger Institute in Cambridge, UK, to completely knock out all 20,000 genes across a wide range of cell and tumour types. Future libraries will allow us to upregulate (CRISPR activation or CRISPRa) or downregulate (CRISPR interference, CRISPRi) these genes. We can then quantify a wide range of phenotypic responses, such as cell proliferation or death, RNA or protein expression, metabolism, morphology and resistance to therapy, using this information to develop valuable new biological insights for cancer drug discovery.

It is notable that we are looking at the impact of up- or down-regulating each gene as well as knocking out its function. CRISPRn usually relies on a specific guide RNA to direct a DNA nuclease
(Cas9) to the desired location in the genome. Once the nuclease has made a double-strand cut, the cell’s own repair mechanisms will fix the break. This either leads to a loss of function of the gene or replacement with an alternative sequence, depending on the particular reagents present.

By contrast, CRISPrα and CRISPrβ work by guiding a nuclease-dead Cas9 (‘dCas9’) coupled to either a transcriptional activator or repressor to a specific promoter sequence, where it activates or inhibits gene transcription. Not only does this begin to provide insights into the potential impact of alterations in non-coding regulatory elements that control gene expression levels, it also more realistically mimics the physiological effect of a drug in the body, which might reduce but not completely remove the target. Initially we plan to focus on the exome, altering the function of protein-coding genes, expanding to encompass more esoteric targets such as microRNAs and key regions of the non-coding genome in the future.

Our initial screens are being carried out using cancer cell lines and primary tumour cells, for example from CRISPR-derived cellular or animal models of disease, and from patient samples. However, we are aiming to move from simple cell culture to more realistic situations that incorporate elements of the tumour microenvironment, such as hypoxia or immune cell infiltration. Manipulating culture conditions may also reveal phenotypes that are only expressed under certain physiological conditions, such as oxidative stress or increased acidity. Carrying out screening or target validation in more physiologically-relevant systems such as organoids, co-cultures, patient samples, animal models or organ-on-a-chip technology is more likely to provide results that will bear up in clinical testing, greatly increasing the chances of success.

CRISPR offers a completely new approach for the discovery and validation of novel targets. It is efficient – assaying 20,000 genes simultaneously – and we can skip straight from a CRISPR ‘hit’ to the underlying biological target. This approach also has wider applications outside of oncology, and we are applying the concept of large-scale, high-throughput CRISPR screening to identify and validate novel targets across a wide range of therapeutic areas including respiratory, cardiovascular, metabolic and renal diseases.

**AI approaches for rational discovery**

While we anticipate that manipulating individual genes across the genome will reveal novel targets, cancer cells are complex systems, with multiple genes working together in pathways that can be genetically or epigenetically ‘rewired’ in response to treatment. It is therefore likely that we will need to investigate the impact of altering combinations of multiple genes at the same time, in multiple cell types and with multiple drug combinations.
One approach is to use CRISPR to generate polygenic cellular or animal models of disease, then carry out genome-wide screens on cells or organoids derived from them. Even against that background, manipulating a single additional gene may not be sufficient to generate a detectable phenotype, so it might be necessary to alter two or more genes at the same time.

However, the numbers quickly become unmanageable – even something as simple as pairwise combinations of less than a hundred genes in a handful of cell types quickly adds up to more than 140,000 interactions\(^6\). It is impossible to test all possible permutations, so it is vital that we develop ways to identify combinations that are most likely to have an effect.

Our Functional Genomics team will employ cutting-edge AI and machine learning algorithms to analyse large-scale genomic information and other biological datasets from model organisms, experimental results and patients. By drilling down into the biology of cancer cells, the algorithm can reveal rational combinations of genes or drugs that are likely to act within the same network or synergistically (for example, to induce synthetic lethality).

Applications in oncology
Fundamentally, we see this integration of genomics, functional genomics and AI technology becoming a routine part of drug discovery, putting science front and centre in strategic decision-making about target identification and validation to increase the chances of a novel drug candidate successfully making it to patients. The most important application of this research programme is therefore the identification and validation of biologically relevant targets that impact disease progression.

Many potential therapeutics have biological plausibility – they have been designed to hit a particular molecular component in a pathway that has been implicated in disease and ‘should’ work. Yet numerous potential medicines have failed because they turn out not to act upon their expected target, with some failing as late as Phase III clinical trials. Others have unacceptable toxicity or pharmacokinetic properties, which are also related to selecting an inappropriate target.

Our functional genomics approach provides a new way of quickly and efficiently generating biologically-relevant targets for further investigation, hand in hand with more conventional methods of target hunting such as genomic analysis. It also provides a means of rapid phenotypic validation, enabling us to screen out misleading targets at an early stage that do not validate in ‘real life’, despite having strong biological plausibility.

For example, if altering the activity of a particular gene in a cancer cell leads to cell death, then that reveals a potential novel drug target for further investigation and validation. Alternatively, if a CRISPR screen reveals a gene that appears to be essential for a cancer cell to become resistant to a particular therapy, we can then carry out further experiments in cells or animal models – also generated using CRISPR – to investigate that mechanism of resistance and develop ways of targeting it. As proof of principle, this approach has already been used by the Wellcome Sanger Institute team to reveal genetic vulnerabilities and potential therapeutic targets in leukaemia\(^7\). We already have
targets entering our discovery portfolio identified through this approach.

Furthermore, there is an urgent need to explore novel therapeutic spaces. Drug development research is focused on targets that are believed to be ‘druggable’, such as kinases or cell surface receptors that can be targeted with small molecule inhibitors. By targeting every single gene in the genome and understanding the networks in which they function, hand in hand with new modality approaches such as antisense oligonucleotides, we can expand the therapeutic world that is available to us.

One area of particular interest is the DNA damage response. Alterations in genes involved in various DNA damage response and repair mechanisms can have a profound impact on the risk of cancer, tumour progression and response to various types of therapy. For example, women carrying a germline mutation in BRCA1 or BRCA2 have a significantly increased risk of developing breast or ovarian cancer due to errors in the homologous recombination DNA pathway. However, inhibiting an alternative repair pathway mediated by PARP results in chromosomal instability and cell death, as the cancer cells are left without any kind of functional DNA repair.

The discovery of this ‘one-two punch’, known as synthetic lethality, led to the approval of the first-in-class PARP inhibitor, olaparib. We are now using a functional genomics approach to map out the landscape of synthetic lethality in cancer cells, searching for similar combinations of co-dependent pathways that can be manipulated using drugs or other therapies.

Another key area of interest is identifying and targeting mechanisms of resistance to cancer therapy in metastatic disease. It is becoming increasingly clear that the underlying genetic heterogeneity within a tumour will inevitably generate clones of resistant cells that emerge as a result of the selective pressure applied by treatment, meaning that a patient whose disease initially responds to therapy will ultimately relapse. Using our functional genomics platform, we can investigate the impact of manipulating thousands of genes in cancer cells that have become resistant to a particular therapy to uncover the molecular mechanisms of resistance.

While it might seem like there are infinite ways that cancer cells can evade the effects of treatment, it is more likely that there is a limited number of escape routes depending on the cancer type and the therapies employed. Some types of cancers will prefer to ‘rewire’ in certain ways – which may depend on their developmental history and tissue of origin – but this is potentially predictable and can be uncovered using a functional genomics approach.

For example, the presence of a T790M mutation in the epidermal growth factor receptor (EGFR) gene is the most frequent mechanism of acquired resistance in non-small cell lung cancer patients, appearing in around 60% of tumours that initially responded to first-generation EGFR inhibition.

Identifying the landscape of genetic alterations that might make a cancer resistant to a given drug could reveal potential targets for novel therapies that overcome or prevent resistance. We can also use AI to tease out potential combinations of therapies that would create a ‘double bind’ of mutually exclusive resistance mechanisms in tumour subclones.

As well as understanding the general mechanisms of drug resistance in various cancer types, we can also use our functional genomics platform to investigate the phenomenon of drug tolerant persister cells that neither proliferate nor die, creating a ‘sleeper pool’ from which highly-resistant clones can eventually emerge. CRISPR screening could reveal key genes responsible for dormancy and reactivation, identifying targets for more effective therapies that prevent long-term relapse and significantly extend survival from metastatic disease.

In addition to using CRISPR to leverage the genome to improve drug discovery, there are further opportunities for AstraZeneca to better understand and improve the tools themselves; for example, by increasing the accuracy of targeting and reducing off-target effects or by improving the delivery of CRISPR tools into cells.

A collaborative effort

The Functional Genomics Centre is based at the new Milner Therapeutics Institute in Cambridge, right next door to our global headquarters and strategic R&D centre within the biggest life sciences research hub in Europe. The Milner Therapeutics Institute provides a unique collaborative space and environment on the Cambridge Biomedical Campus.

The initiative consolidates expertise across the academic, clinical, charitable and industrial sectors to create standardised, bench-marked platforms for discovery science. It is being established under the guidance of world-leaders in functional genomics, including Professor Greg Hannon, Director of the Cancer Research UK Cambridge Institute. Our investment is pushing the UK to the forefront of CRISPR applications and bringing vital new insights into cancer, leveraging the secrets
hidden within the human genome to create potential life-changing therapies.

There are also wider benefits for the scientific community. The scale, costs and expertise required to run large-scale functional screens are prohibitive for most individual labs, but by working with Cancer Research UK’s Therapeutic Discovery Laboratories, we are democratising access to this technology for the many thousands of scientists and clinicians funded by the charity. In particular, Cancer Research UK’s strategic focus is on hard-to-treat cancers including lung, pancreatic, oesophageal and brain tumours, which have seen small gains in survival over the years and where there is much to gain from the discovery of novel therapeutic approaches.

By combining the power of genomics, CRISPR screening and AI, the Functional Genomics Centre aims to bring powerful insights into some of the thorniest challenges in medicine and drug discovery today, particularly the rapid evolution of resistance to targeted therapies in oncology.

Ultimately, we are confident that the Functional Genomics Centre will help to advance science by bringing together academia and industry to reveal insights into the biology underpinning the mechanisms of disease, target validation and therapeutic resistance.

**DDW**

---

**Dr Ultan McDermott** is Chief Scientist – Oncology at AstraZeneca where he focuses on drug resistance in cancer. He is a clinician scientist with a lifelong interest in understanding how cancer genomes affect response to therapy in the clinic and, more recently, the use of genetic screens to identify mechanisms of drug resistance in cancer. Ultan also holds an Honorary Consultant post in the Oncology department at Addenbrooke’s Hospital where he treats colorectal cancer patients.

**Dr Steve Rees** is Vice-President of Discovery Biology at AstraZeneca with global responsibility for functional genomics, reagent generation and assay development. He is Chair of the European Laboratory Research and Innovation group (ELRIG) and is a member of the Scientific Advisory Board for LifeArc and the Centre for Membrane Protein and Receptor research at the Universities of Nottingham and Birmingham.

**Dr Susan Galbraith** is Senior Vice-President of Early Oncology at AstraZeneca. Susan is responsible for the discovery and early development of molecules in the core biology areas of tumour drivers and resistance mechanisms, DNA damage response and epigenetics. Susan is regularly invited as a speaker/session chair at the American Association of Cancer Research (AACR), National Cancer Research Institute and European Society of Medical Oncology. She has also served on the AACR Scientific, Finance and Clinical Trial Committees and is a Fellow of the Academy of Medical Sciences.

**Dr Mike Snowden** is Senior Vice-President and Head of Discovery Sciences, R&D BioPharmaceuticals at AstraZeneca. His group focuses on internal and collaborative research to provide a comprehensive range of technical and translational activities to preclinical project teams including, HTS and SAR screening, X-ray crystallography, medicinal chemistry, translation, technology for the many thousands of scientists and clinicians funded by the charity. In particular, Cancer Research UK’s strategic focus is on hard-to-treat cancers including lung, pancreatic, oesophageal and brain tumours, which have seen small gains in survival over the years and where there is much to gain from the discovery of novel therapeutic approaches.

By combining the power of genomics, CRISPR screening and AI, the Functional Genomics Centre aims to bring powerful insights into some of the thorniest challenges in medicine and drug discovery today, particularly the rapid evolution of resistance to targeted therapies in oncology.

Ultimately, we are confident that the Functional Genomics Centre will help to advance science by bringing together academia and industry to reveal insights into the biology underpinning the mechanisms of disease, target validation and therapeutic resistance.

**DDW**

---

**Dr Ultan McDermott** is Chief Scientist – Oncology at AstraZeneca where he focuses on drug resistance in cancer. He is a clinician scientist with a lifelong interest in understanding how cancer genomes affect response to therapy in the clinic and, more recently, the use of genetic screens to identify mechanisms of drug resistance in cancer. Ultan also holds an Honorary Consultant post in the Oncology department at Addenbrooke’s Hospital where he treats colorectal cancer patients.

**Dr Steve Rees** is Vice-President of Discovery Biology at AstraZeneca with global responsibility for functional genomics, reagent generation and assay development. He is Chair of the European Laboratory Research and Innovation group (ELRIG) and is a member of the Scientific Advisory Board for LifeArc and the Centre for Membrane Protein and Receptor research at the Universities of Nottingham and Birmingham.

**Dr Susan Galbraith** is Senior Vice-President of Early Oncology at AstraZeneca. Susan is responsible for the discovery and early development of molecules in the core biology areas of tumour

---

**References**

3. https://www.nature.com/articles/ng.3314.