

Investigating the Origins of Cancer in the Intestinal Crypt with a Gene Network Agent Based Hybrid Model

Arturo Araujo^{1,4}, Hanxiao Zhang², Albert Rübben³ and Peter J Bentley^{1,4}

¹ Braintree Ltd, London, United Kingdom

² CoMPLEX, UCL, London, United Kingdom

³ Euregio Skin Cancer Center, Department of Dermatology, University Hospital RWTH, Aachen, Germany

⁴ Department of Computer Science, UCL, London, United Kingdom
a.araujo@braintree.com

Abstract

Colorectal cancer (CRC) is the second most common tumour in the world (Bray, 2018). It has been proposed that morbidity and mortality could be mitigated by screening methods that identify key genetic mutations in the DNA of a patient's biosample (Traverso, 2002). However, for this to work, a theoretical understanding of the most likely mutations that initiate malignant transformation, and how they affect subsequent microevolution, is needed. Specifically, we hypothesise that *there is a CRC-proliferative mutation that is more likely to be initially fixated in the crypt*. To investigate this, we developed an agent-based model of cells in the colon crypt that shows emergent biological homeostasis at the tissue level from the cellular and molecular interactions. We equipped each of the cells with a molecular gene network which, in their wildtype state, regulates homeostasis in the crypt and recapitulates known behaviour. We identified and modelled key genes implicated in CRC which, when mutated, alter the rate of death and division of cells. We used this model to study the biological first principles of the fixation of mutations, offering key spatial and temporal understanding of this process. We discuss the impact and clinical relevance of proliferative genetic mutations in isolation, pointing to the KRAS gene as a likely mutation to be initially fixed in the crypt.

Introduction

The development of colorectal cancer (CRC) is thought to occur in marked stages happening throughout an entire decade (Williams, 2016); however, the precise mechanisms of oncogenic initiation are still debated. Because of this slow development, CRC is a highly suitable system for investigating the emergence and accumulation genetic alternations that many cancers have in common. However, *in vivo* and *in vitro* models usually take place in a matter of days or weeks, not being able to fully recapitulate the microevolutionary oncogenic process. Current research points to CRC having an origin at the base of colonic crypts: flask shaped invaginations in the inner lining of the intestine, which produce new cells to support the tissue (Figure 1). Cells at the top of the crypt are continuously worn away by the process of nutrient absorption (through the villi) and passing food, and thus are being continually renovated by stem cells at the bottom. These stem cells divide to replace worn cells and may even displace other stem cells so that at a given time the whole crypt becomes monoclonal- a descendant of one single stem cell. Because of the high rate of division, it is here that key oncogenic mutations are thought to arise

(especially during cell division, which is the most vulnerable state for the DNA of the cell).

It is very difficult to investigate *in vivo* or *in vitro* the impact that key genetic mutations have in isolation. However, this gap in our understanding needs to be addressed to investigate how these initial mutations may shape the subsequent evolution of the disease, and thus offer some therapeutic insights. We propose to overcome the limitations of time, cost and tractability of wet-laboratory experiments with the aid of theoretical approaches, such as agent-based modelling and gene networks. The focus of this work is to determine whether *there is a CRC-proliferative mutation that is more likely to be initially fixated in the crypt*. To answer this question, we have abstracted the natural processes of crypt homeostasis and distilled current genetic knowledge into a network. With this realistic simulation, we can recreate the *in vitro* scenario that isolates the probability of key CRC implicated mutations to be fixated and thus become the steppingstone in oncogenic transformation.

Background

ALife has helped model emergent phenomena in human diseases such as the emergence and evolution of brain tumours (Swanson, 2003). ALife researchers have used the distilled essence of biological behaviour to develop nature-inspired computational techniques that will help us create better tools (Andrews, 2008). It is with this virtuous cycle of the development of ALife methods in mind that we seek to formalize the simulation of complex biosystems and apply these to the understanding of cancer (Araujo, 2010) (Rübben, 2013), with the ultimate goal of discovering novel therapeutic targets such as key switches in cell-cell communication (Bentley, 2014). The ALife methods that us and others have used, such as agent-based modelling, have allowed us to investigate the genetic mutations that occur at a molecular level, but which have repercussions at the tissue level; as well as the cross-talk between cell types that make up the tumour microenvironment (Araujo, 2014). Importantly, these modelling approaches can connect the different time and space scales are needed (Rejniak, 2010), and have shown the feasibility of modelling cells with internal genomes to study cancer initiation (Fontana, 2010).

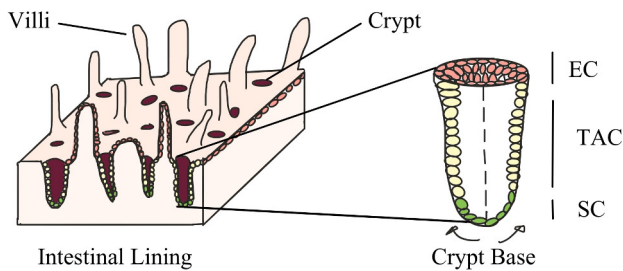


Figure 1. An intestinal crypt is an invagination in the inner lining of the intestine. Stem cells at the base constantly replenish the tissue.

Modelling Colorectal Cancer

There have been advances in the computational study of CRC, such as the effects of variation of cell cycle rate in that may disrupt homeostasis in colon crypts (Smallbone, 2013). Bravo and Axelrod measured the variation in stem cells, proliferating cells, and differentiated cells in multiple crypts in normal human biopsy specimens, offering a metric of the robustness with which crypts recover from chemotherapy and radiation scheduling protocols (Bravo, 2013). Meineke used a lattice-free cylindrical surface to model experimental data showing that cell movement is a consequence of mitotic activity (Meineke, 2001). Venturing into molecular modelling, Leeuwen developed a hybrid model that incorporates the WNT ligand gradient along the crypt axis that is shown to regulate the cell cycle and cell division (Van Leeuwen, 2009), being one of the first truly multiple-scale abstractions that link the subcellular, cellular and tissue level processes.

Agent-Based Modelling (ABM) theory has been further developed to study subcellular behaviour. Using such advances, Mirams was able to show theoretically that the probability of mutations fixation is only weakly associated with the destruction of WNT-dependent cell proliferation (Mirams, 2012). Introducing space into the already complex dynamics, Buske developed one of the first 3D crypt models to study combined effect of WNT and Notch signalling on cell proliferative behaviour (Buske, 2011). There are other ABMs, such as those proposed by (Dunn, 2013) and more recently by (Ingham-Dempster, 2017), that abstract the concept of anoikis (programmed cell death) for a systemic investigation of emergence behaviours and migration dynamics that are difficult to study in vitro. The researchers were able to localize cell death to a small region at the top of the crypt and identifying it as an emergent property in response to changes in cell proliferation rates.

Mutations in the Intestine Crypt Evolution

Recent work in genetics and molecular biology has helped identify key genetic alternations implicated in the tumorigenesis of CRC. These genes have been generally classified as either tumour-suppressor genes (TSGs), due to their anti-growth properties; or proto-oncogenes (POG), which drive proliferation and will be the focus of our investigation. CRC mutations inactivate TSG function, or increase POG function (the mutated version termed oncogene). Typically, mutations are recessive in TSGs but dominant in proto-oncogenes; this means both alleles need to be altered in TSGs

for loss-of-function to occur, whereas mutation in one of the copies of proto-oncogenes is enough to promote proliferative behaviour, and thus have a more measurable effect (Evan 2006). It has been calculated that in normal human cells, the average mutation rate is $\sim 2.5 \times 10^{-8}$ mutations/nucleotides (Nachman, 2000). Mutations can manifest as point mutations altering one specific gene or through small structural aberrations such as short gene duplications, deletions or inversions. Furthermore, cancers may demonstrate chromosomal instability (CIN), where defects during cell division leads to daughter cells with large chromosomal amplifications, deletions or whole chromosome rearrangements (aneuploidy). CIN that lead to changes in chromosome number or structure accounts for 85% of sporadic colorectal cancers (Tsang, 2014). These mechanisms may lead to the loss of a TSG or to a gain in POGs such as KRAS.

Although much work has been done to elucidate the genetic signalling pathways, the connections between all the genes involved in CRC oncogenesis is still not fully understood. The current evidence points to the existence of some key mutations that are consistent in CRC. Amplifications and mutations of POGs KRAS and BRAF have been consistently observed in CRC; especially in CIN tumours (Pino and Chung 2010). While some mutations have been studied both biologically and clinically, the knowledge on the role they have when combined with mutations in other OGs or TSGs is still limited. It is therefore of importance to determine the chances that different mutations have to get fixated in the crypt, as these would then synergise with subsequent mutations in the path to malignant transformation.

We hypothesise that *there is a CRC-proliferative mutation that is more likely to be initially fixated in the crypt*. To provide evidence for this, we will simulate key mutations in an agent-based model that can help us understand 1. how long does it take for a random mutation to be fixed in the crypt? and 2. how fast would a key mutation fixate in the crypt? We will perform multiple simulations on CIN conditions (such as copy number decreased to 1 or 0 or copy number increased to 3) and point mutation in each POG abstraction.

The System

In order to investigate the role that the individual mutations have in CRC initiation, we have designed a computational model that exhibits the same homeostatic behaviour of a healthy crypt. We have abstracted the behaviour at a cellular level and modelled each cell as a circular agent. As described in Figure 1, the colon crypt is an invagination that is in constant renewal. For representation, we have adapted the invagination into a two-dimensional plane made up of cells with a continuous boundary to the left and right of the cells, thus preserving the original three-dimensional geometry (Figure 2). In our computational model, cells are represented by agents that react to the morphogens in the microenvironment, enabling cells to divide, quiesce or die (Figure 3). During normal homeostasis, three populations of morphogen-regulated cells coexist: Stem Cells (SCs) at the bottom, Transit Amplifying Cells (TACs) in the middle and fully differentiated Epithelial Cells (ECs) at the top (Figure 2). Cells are physically able to push other cells in all directions, with higher probability of

pushing cells up or sideways up, a low probability of dividing sideways and an even lower probability of going downwards to the sides or downwards (Table 1). SCs proliferate at the bottom compartment, pushing cells up and supplying a fresh batch of TACs that eventually differentiate at the top of the crypt and are shed away.

We previously studied the intestinal glands in the colon, or colon crypts: invaginations in intestinal tissue that help absorb nutrients as food passes through them (Figure 1). We found that ALife techniques are ideal to address biological complex systems at the molecular, cellular and tissue level; and capable of shedding light on in vivo experiments that report seemingly different findings. Specifically, we were able to bridge different values reported for these contributors, and thus reconcile theories on which one is the biggest contributor the time it takes for a crypt to become the descendant of a single basal stem cell, also known as monoclonality (Araujo, 2018). In our work we focused on two key morphogens the process of cell renewal in the crypt: WNT (promoting the stem-cell phenotype) and EGF (promoting cell division and regulating cell differentiation). We modelled both morphogens as being maximum at the base of the crypt, as they are thought to be provided by Paneth cells which reside there (Sato, 2011). It is currently thought that the morphogens concentrations decrease in a gradient throughout the length of the crypt (Bach, 2000). In our homeostatic model, the WNT ligand concentration, keeping all the cells in contact with it in a stem cell phenotype, is completely depleted 10 to 30 μm (approximately one to three cell diameters) above the base. When cells are out of the WNT concentration, but still within the EGF gradient, they lose stem cell properties, start aging, and are able to divide proportionally to the bio-availability of EGF. Once these transit-amplifying cells are pushed outside of the EGF gradient (approximately 31 μm above the base of the crypt) they become fully differentiated epithelial cells, stop dividing and their likelihood of being shed away or dying is 100%. Besides EGF, other morphogens have been implicated in the regulation of TAC cells (Carulli, 2014) and EGF in the model is only representative of a putative morphogen acting by a gradient.

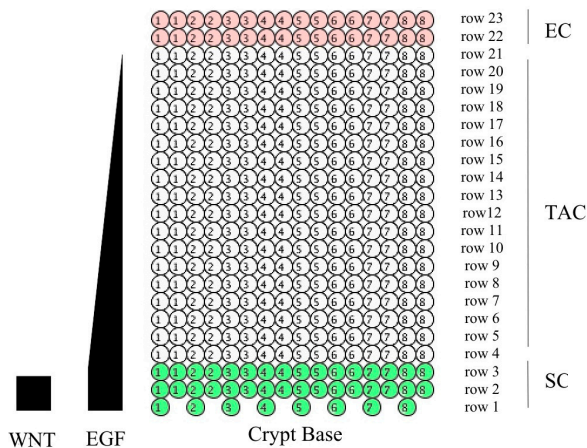


Figure 2. The compartments that make up the colon crypt model. The WNT morphogen (right) maintains cells in SC state, while EGF regulates cell division.

Our previous results (Araujo, 2018) show that the geometry of the crypt such as the total number of stem cells (Fletcher, 2012), the proportion of side cell displacement (Ritsma, 2015) and the number of basal stem cells (Kozar, 2013) all have a profound impact on the time to monoclonality. We showed that a niche of stem cells dividing from the bottom provide a continuous influx of new cells, and that eventually the dynamics (such as cell displacement), generate observable global effects such as a change in the frequency of monoclonality. This is important because it offers a metric for the possible fixation of mutations; but it doesn't consider the phenotype changes that an altered genotype might confer. To obtain a true metric by which genetic alterations will spread through the crypt it is important to include realistic abstractions of oncogenic mutated cells in the intestine.

The Algorithm

We had previously developed an ABM that recapitulates the known mechanics of the healthy crypt (Araujo, 2018). Novel in this work, we have completely redesigned our agents, giving them an internal genetic circuit for decision making. As in the original model, every cell is queried at every time step in an asynchronous update. A random cell that has not been previously updated during the time step is picked and follows the update algorithm by which it is given the chance to decide whether to die, divide or do nothing per cell cycle (Figure 3).

The cell identity is decided by its position with respect to the WNT concentration (SC) and the EGF gradient (TAC) though a probabilistic calculation. SC cells have a fixed rate of division, $division_{SC}$, while under the influence of WNT, which might be altered by a mutation. Their probability of division 100% throughout the wildtype SC compartment is then calculated as:

$$P(SCdiv) = 100 * (division_{SC})$$

Wildtype TACs have 100% chance of dividing as they leave the WNT concentration but are still in contact with the EGF gradient. The wildtype TACs chance of division, $TACdiv$, decreases linearly to 0% as it travels to the top of the EGF gradient. To achieve this, we normalize the difference between the EGF and WNT, $\Delta grad$, and calculate the vertical distance, $yPos$, between the TAC cell and the end of the WNT ligand concentration, which modulates $division_{TAC}$:

$$P(TACdiv) = 100 * (division_{TAC}) * (yPos - \Delta grad) / \Delta grad$$

Where $division_{SC}$ and $division_{TAC}$ are biological parameters for wildtype cells, as shown in (Table 1). When a cell divides it pushes one of its neighbours, selected with the baseline probabilities shown in Table 1. The probability of death is 100% when leaving the EGF gradient. We model aging as a decrease in telomere length, therefore reducing the number of times a cell can divide (initially 5 divisions) and persist within the TAC-compartment. Each round, the algorithm queries the system for a user defined condition. For this work we will use one of three user-specific queries: 1. a set time (in days), 2. if the basal SCs have become monoclonal, and 3. if a mutation has disappeared from the basal SC compartment. If the user specific query is not met, it continues updating cells, and advancing to a new time step when every cell has been updated or stopping when the query has been met.

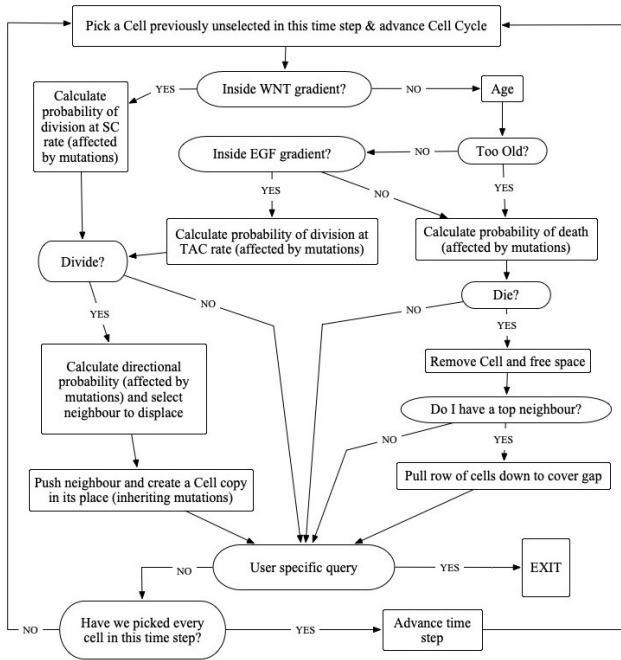


Figure 3. Algorithm for cell dynamics. The program ends when the user specific query (E.g. reaching monoclonality) is met.

Agent design

We designed the agents to offer different information that could be displayed visually with regards to key metrics of interest. In the wildtype visualization, SCs (cells inside the WNT ligand concentration) are tagged green, TACs (cells outside the WNT and in the EGF gradient) are marked white, and ECs (cells outside of both gradients) are coloured peach (Figure 2). To track the lineage of cells, we give a tag number to each of the initial basal SC which is inherited by their progeny throughout the simulation. In the initial baseline condition, as shown in Figure 2, we assume that cells that are in the column immediately on top of the SC are progeny of it and therefore inherit this number, which is displayed at the centre of each cell in the visualization. This gives a representation and a clear pathway of how cell mixing and eventually monoclonality occur. Other metrics such as the number of divisions and age are also stored in each agent and can be shown as the number displayed on each cell.

Biological Parameters

The structure of the seven gene network used in this paper was extracted from and confirmed against literature (De Roock, 2011) (Sartore-Bianchi, 2016) (Strubberg, 2017) (Pan, 2017) (Berg, 2012). Inputs (WNT and EGF levels at crypt positions), outputs (calculated cell properties) and trigger values (threshold levels for different outputs) of the gene network algorithm in the baseline case (no mutations), were selected based on the parameters from Table 1, while mutation impacts are shown in Figure 3. The gene output for copy number changes, gene inactivation/activation mutations, signal thresholds for output changes were selected to reproduce the expected mutation impact outputs (Figure 4 and Table 2).

The gene signalling pathway diagram was converted into a simple algorithm (Figure 4) which returns cell parameters for proliferation probability, death probability, cell cycle length and cell fate (stem cell, TAC or differentiated cell) based on the position of the cell in the crypt. Also included in the algorithm are abstracted impacts for activating or deactivating point mutations and copy number changes e.g. if KRAS was activated proliferation probability was increased. Each stage of the algorithm represents a gene receiving a signal of a certain strength either from the previous gene in the network or from the initial signalling protein (EGF or WNT).

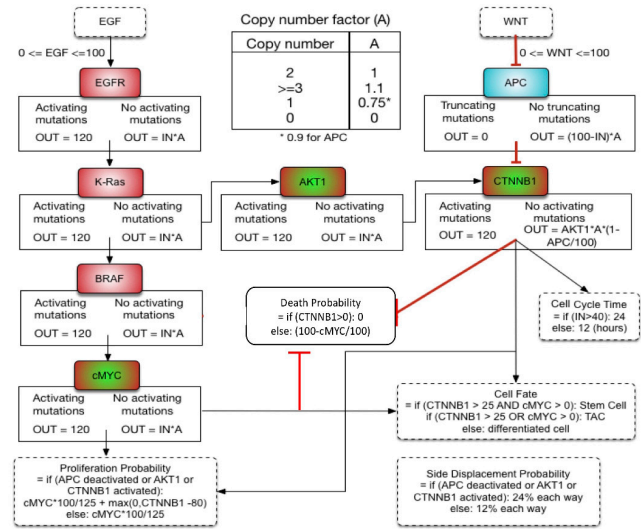


Figure 4. Simplified gene network. The network was constructed based on the mutation frequency of key CRC related genes documented in the COSMIC database ((Forbes, 2017); cancer.sanger.ac.uk). Black arrows with solid line: gene activation; Black arrows with dash line: effects; Red arrows: inhibition; red boxes: proto-oncogenes; blue box: tumour suppressors. The input and output variables as well as the calculation for signal integration are a first approximation to reconcile in-vivo (Ritsma, 2015) and in-vitro (Snippert, 2013) data with yet-unquantified but known genetics (Forbes, 2017).

Parameters	Value	Reference
Cell Diameter	10 μm	(Bach, 2000)
Width and length of crypt	16 x 25 cells	(Bach, 2000; Totafurno, 1987)
SC division rate (<i>division_{SC}</i>)	Once every 24 hours max	(Fischer, 2016)
TAC division rate (<i>division_{TAC}</i>)	Once every 12-24 hours max	(Kozar, 2013)
Probability of Side Displacement per cell division (<i>SDisp</i>)	0.24 (0.12 each side) to 0.74	(Ritsma, 2015) (Snippert, 2013)
Probability of Downwards Displacement per cell division	0.08 for TACs, 0.00 for SCs	(Ritsma, 2015)

Table 1: Baseline parameters used to simulate a murine small intestine crypt.

Genetic Alteration	Change in Probability
Gene copy number increase to 3 or more	output gene signal > 100%
Gene copy number decrease to 1	output gene signal < 100% > 0%
Gene copy number decrease to 0	output gene signal = 0%
Activating mutation	output gene signal > 100% regardless of input
Minimum cell cycle length	12 hours (normally applies to TAC)
Maximum cell cycle length	24 hours (normally applies to SC)

Table 2- Altered probabilities from mutations in the abstracted gene network. The actual values of the variables as well as the calculations describing signal transduction are still being elucidated, and thus we have abstracted these values in a first approximation to recapitulate published data.

Model Validation 1: Positional Mitotic Index

We compare our model with a metric used by us and others in the past: mitotic activity, which causes a pressure-driven passive movement from the bottom to the top of the crypt, as described in (Meineke, 2001). We ran 100 simulations and defined our end user-query as a 30-day period. When tested, the average division per row (Figure 5) recapitulates the data for mitotic index distribution in the crypt presented in Fig 4.iv of (Sunter, 1979), where there is a maximum of mitotic activity a few rows after leaving the base of the crypt and decreases throughout the rest of the rows.

Model Validation 2: Neutral Drift Dynamics

We test for evidence of a neutral drift in the cellular dynamics of intestinal stem cells, and compare to results found by (Snippert, 2010). In their experimental system, a 14 basal stem cell population was calculated to have a probability of side displacement of 0.74, demonstrating neutral competition amongst SCs. In 100 simulations of such experiment (Figure 6) we find found that the dynamics exhibited by the updated model agree with that reported in Figure 7D of (Snippert, 2010).

Experiments

We hypothesise that *there is a CRC-proliferative mutation that is more likely to be initially fixated in the crypt.* To provide evidence for this, we focus on answering: 1. how long does it take for a random mutation to be fixed in the crypt? and 2. how fast would a key POG mutation fixate in the crypt? To investigate the time to mutation fixation and quantify the role that the individual mutations have in CRC initiation, we first establish a baseline to fixation. Subsequently we simulate experiments that explore mutational scenarios. Our experiments are:

Experiment 1- Baseline time to monoclonality. In their cell-tracking experiments Ritsma et al. measured the time it takes for one basal SC sub-clone to divide sideways enough to make the entirety of the crypt a descendant of this cell, or monoclonal

(Ritsma, 2015). The researchers performed this experiment with a number of crypts and describe in their results the percentage of crypts that have become monoclonal in a 140-day period. The researchers use 8 basal SC and 16 suprabasal SCs, each with a probability of side displacement of 0.24% (0.12 each side) per division. We perform 100 simulations with the baseline model, as shown in Figure 7, based on the accepted parameters shown in Table 1 and Table 2, and measure our results using this same methodology.

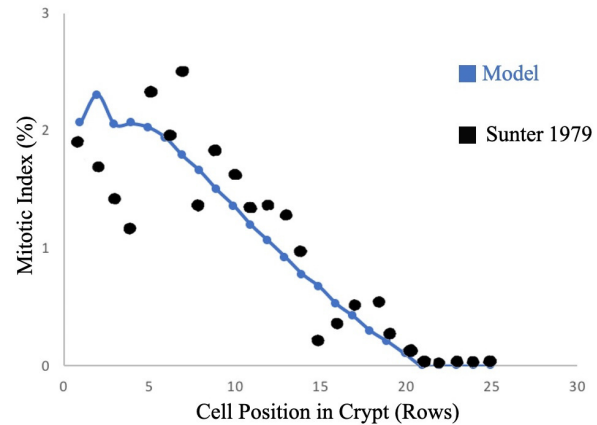


Figure 5. Mitotic index distribution in the crypt (upwards from the crypt as described in in Figure 2) of 100 simulated crypts with 8 basal and 16 suprabasal stem cells over a 30-day period. We compare it to data presented in Fig 4.iv of (Sunter, 1979).

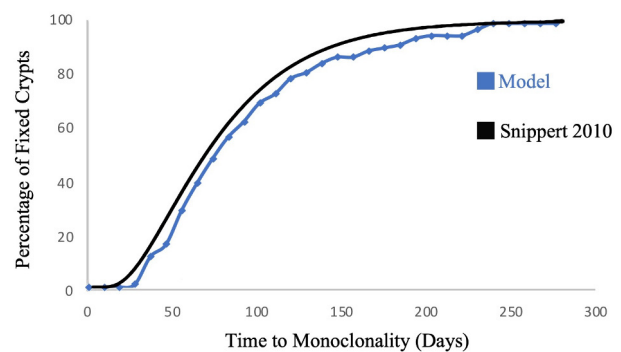


Figure 6. Frequency of monoclonal crypts over time as a percentage of surviving SC clones (out of initial 14) as predicted by neutral drift dynamics (Snippert, 2010).

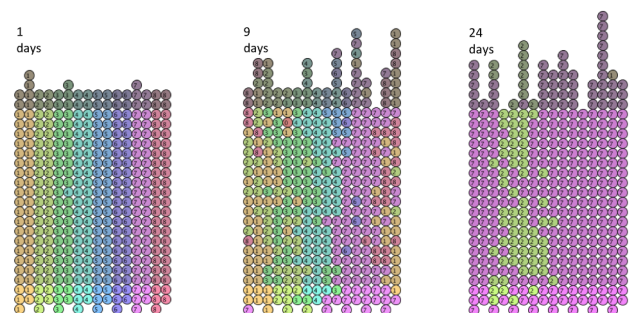


Figure 7. Progression from 8 stem cell lineages towards monoclonality of stem cell 7 (marked purple) over 24 days.

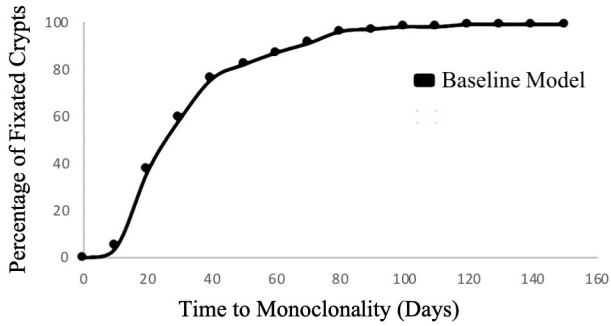


Figure 8: Simulation of biological data for monoclonality based on experiments reported on (Ritsma, 2015). From 100 simulations (as shown in Figure 7) we measure every 10 days the percentage of simulations that have become monoclonal. By day 140, 100% have reached this state.

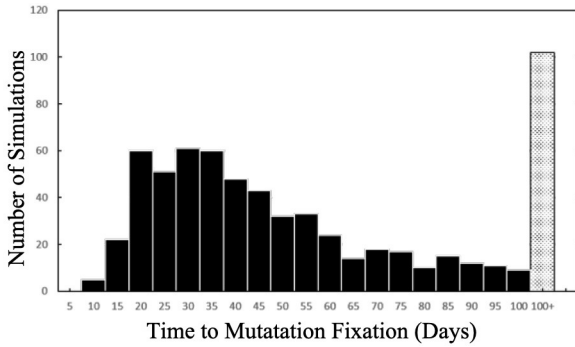


Figure 9. Time to a mutation fixation within 100 days (black boxes) from 4000 simulations. The grey box shows the number simulations that still contained a mutant at the basal SC compartment that could spread throughout the crypt given more simulated time.

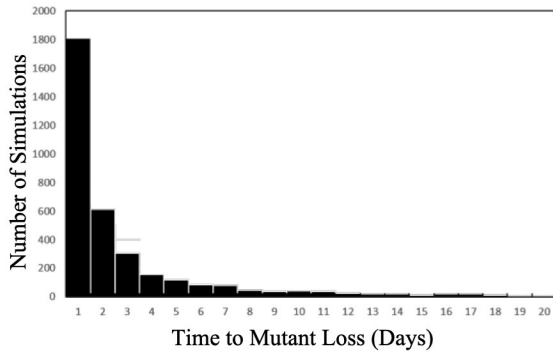


Figure 10. Time to a mutant loss from the crypt from 4000 simulations.

Experiment 2- Time to Mutation Fixation. In Experiment 2, a genetic mutation is randomly introduced into one of the basal stem cells (BSC). Mutations that lead to changes in the output signal of the gene in Figure 4 will ultimately alter proliferation probability, cell death probability and cell cycle length.

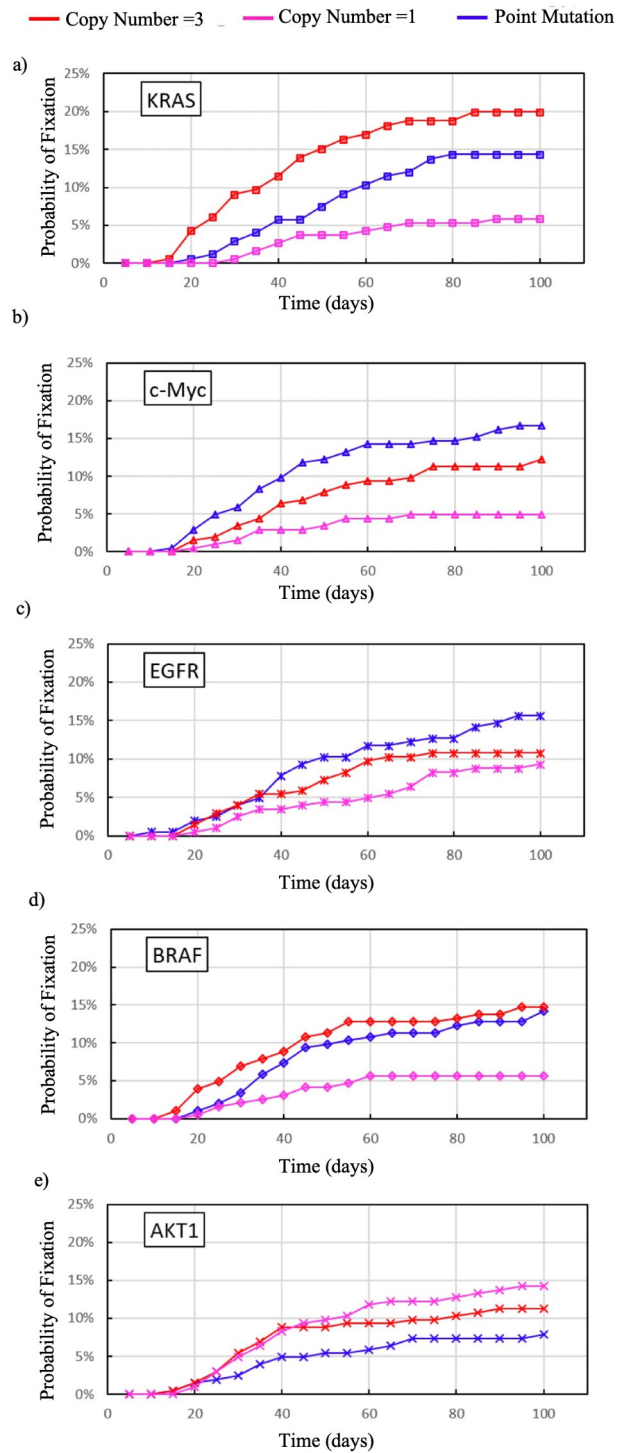


Figure 11. Probability of fixation for each genetic mutation as a cumulative distribution function for the oncogenes KRAS, c-Myc, EGFR, BRAF and AKT1.

The crypt is allowed to evolve for 100 days and the simulation is stopped when one of the three conditions is satisfied: 1. The time limit is reached. 2. Mutation is fixed. 3. The mutation population is swept from the crypt. The

mutation is assumed to be fixed when the mutated lineage takes over all BSCs, as eventually, all cells in the crypt will have the same ancestor and inherit the same type of mutation. Similarly, the mutation is assumed dead when mutated ancestors is lost from the first row.

Experiment 3- Fixation Probability of Oncogenes. Finally, we simulate a point mutation and compare it to the gain of one gene and the loss of one alleles of the proto-oncogenes *EGFR*, *KRAS*, *BRAF*, *c-Myc*, and *AKT1*, with the altered probabilities shown in Figure 4. 200 simulations were run for each genetic mutation with basal stem cell number set to 8.

Results

Results 1: Establishing a baseline time to monoclonality. In our baseline simulations for monoclonality (Figure 8), 100% of the crypts become monoclonal, or fixed, by day 140 following the same trend as that reported by Ritsma (Ritsma, 2015). These dynamics are important to our understanding of the genetic evolution of the crypt, since they would give us an estimate of how fast we can expect a neutral mutation to spread through a healthy crypt. Using this as our baseline, we will proceed to analyse the results of Experiments 2, 3 and 4.

Results 2: Time to Mutation Fixation. Results from Experiment 2 show a distribution of the time it took for a mutation to become fixed in the crypt. The data collected from 4000 simulations ran for 100 days of simulated time shows that, if a mutation arises, it is mostly likely to be fixed within 20 to 40 days (Figure 9). Interestingly, 102 simulations contained at least one mutant cell which could dominate the crypt given more simulated time.

Figure 10 summaries the times at which the mutated ancestor was washed out from the crypt, showing that the survival rate for a mutation is low. Most mutant cells were swept from the crypt within the first 2 days. The results suggest that mutant domination is a slow and inefficient process, as most mutant cells are replaced by WT clone in a stochastic manner (Figure 10). This highlights the importance of monoclonal conversion in cancer initiation as it prevents the accumulation of mutations.

Results 3: Fixation Probability of Oncogenes. Figure 11 shows the probability of fixation for known proto-oncogenes under the scenario of a point mutation (blue), an increase in copy number (red) and loss of a gene (pink). The more likely scenario for true oncogenes activation is an activating mutation or an increase in copy number via CIN. Broadly, the cumulative distribution function for the oncogenes *KRAS*, *c-Myc*, *EGFR*, *BRAF* and *AKT1* under our theoretical conditions (Figure 4) is under 20% by day 100. Specifically, *AKT1* showed an increase in fixation when losing copy numbers (Figure 11.e). Some oncogenes have also been implicated in key cell maintenance processes, so having more than one role makes their precise contribution difficult to predict. Models such as this have helped us detangle the isolated effect of single mutations and suggest a metric for their contribution in oncogenesis. In the case of *c-Myc* (Figure 11.b) and *EGFR* (Figure 11.c), the model suggests a higher frequency of fixation via point mutations,

while gene amplification leads to the highest fixation of *BRAF* (Figure 11.d) and *KRAS* (Figure 11.a). The model points to *KRAS* having the highest likelihood of being fixated. These distributions, although in general agreement with reported behaviour in literature, do not by themselves explain a mutation fixation reported to be more than 25% for *APC* and possibly higher for *KRAS* (Vermeulen, 2013). Further, recent studies show that more than 97% of *KRAS* mutations in CRCs could be due to point mutations, conferring perhaps more plasticity (Bronte, 2015). Our results suggest that the precise mechanisms of oncogenic alterations for the fixation of mutations are not yet fully understood. Collectively, these results suggest that the theoretical impact of individual initiating mutations must be refined to fully capture the complexity joint mutations, which will be the next step in our research.

Conclusions

In this work, we investigated whether *there is a kind of CRC-proliferative mutation that is more likely to be initially fixated in the crypt*. We focused on modelling the *EGFR* and *WNT* genes and their connectivity, with special interest on *CRC* implicated proto-oncogenes. The evidence provided by the abstracted model suggests that *KRAS* could be heavily responsible for cancer initiation. In addition, in agreement with biological observations, oncogenes generally have higher fixation probability for activating mutations and lower probability for non-activating mutations. These results align with clinical evidence. Clinical studies for the rest of genetic mutations in *CRC* initiation are limited. In particular, the mechanism underlying aberrant activation of *c-Myc* is unknown. A major reason for this is that majority of these genetic mutations may occur in later stages of cancer progression or presented in an alternative pathway to *CRC*, thus further evaluation on the genetic network is required. The simplified model was able to provide a theoretical insight into the nature and significance of each genetic mutation in early tumour formation, and we aim for it to help both biologist and computational cancer researchers to interesting areas of exploration.

For our next step, we aim to further investigate the genetic and epigenetic interdependencies in *CRC* initiation, focusing on *in silico* experiments that cannot be done *in vitro* or *in vivo*, but which may have a significant impact on colorectal adenoma (Lao, 2011). The current model already allows shared information across multiple scales, so it can be readily extended to consider a multiple sequence of mutations. Also, epigenetic events (i.e. DNA methylations) that affect the regulation of key genes could be added to the molecular level. Quantifying the impact from both genetic and epigenetic abnormalities can help us shed light on non-intuitive mechanisms underlying cancer initiation. Finally, if information such as patient characteristics and carcinogen influence is available, the model could yield clinically relevant results.

The crypt model aims to investigate individual contribution of key genetic mutation in early stage of colorectal cancer. Our simulated results are in general agreement with evidence from the literature, and we believe that the next step is to further extend this bridge to clinically-relevant human data for therapeutic discoveries. Much work remains to be done to fully understand the *CRC* pathogenesis.

Acknowledgments

We would like to thank Braintree Ltd and UCL CoMPLEX for funding, fruitful discussions, and facilitating this collaborative and interdisciplinary research.

References

- Andrews, Paul S. . 2008. "Investigating Patterns for the Process-Oriented Modelling and Simulation of Space in Complex Systems.." ALIFE.
- Araujo, A, L M. Cook, C C. Lynch, and D Basanta. 2014. "An Integrated Computational Model of the Bone Microenvironment in Bone-Metastatic Prostate Cancer." *Cancer Research* 74(9):2391–2401.
- Araujo, Arturo, A Rubben, David Basanta, and Peter J. Bentley. 2018. "Testing Three Hypotheses of the Contribution of Geometry and Migration Dynamics to Intestine Crypt Evolution." Pp. 1–8 in.
- Araujo, Arturo, Peter J. Bentley, and Buzz Baum. 2010. "Modelling the Role of Aneuploidy in Tumour Evolution.." ALIFE.
- Bach, S P., A G. Renehan, and C S. Potten. 2000. "Stem Cells: the Intestinal Stem Cell as a Paradigm.." *Carcinogenesis* 21(3):469–76.
- Bentley, Katie, Andrew Philippides, and Erzsébet R. Regan. 2014. "Do Endothelial Cells Dream of Eclectic Shape?." *Developmental Cell* 29(2):146–58.
- Berg, Marianne and Kjetil Soreide. 2012. "EGFR and Downstream Genetic Alterations in KRAS/BRAF and PI3K/AKT Pathways in Colorectal Cancer: Implications for Targeted Therapy.." *Discovery medicine* 14(76):207–14.
- Bravo, Rafael and David E. Axelrod. 2013. "A Calibrated Agent-Based Computer Model of Stochastic Cell Dynamics in Normal Human Colon Crypts Useful for in Silico Experiments.." *Theoretical Biology and Medical Modelling* 10(1):66.
- Bray, Freddie . 2018. "Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries." *CA A Cancer Journal for Clinicians* 68(6):394–424.
- Bronte, Giuseppe . 2015. "New Findings on Primary and Acquired Resistance to Anti-EGFR Therapy in Metastatic Colorectal Cancer: Do All Roads Lead to RAS?." *Oncotarget* 6(28):24780–96.
- Buske, Peter . 2011. "A Comprehensive Model of the Spatio-Temporal Stem Cell and Tissue Organisation in the Intestinal Crypt" edited by D. Lauffenburger. *PLoS Computational Biology* 7(1):e1001045–14.
- Carulli, Alexis J., Linda C. Samuelson, and Santiago Schnell. 2014. "Unraveling Intestinal Stem Cell Behavior with Models of Crypt Dynamics." *Integrative Biology* 6(3):243–16.
- De Rook MD, Wendy, Veerle De Vriendt MSc, Nicola N. MD, Prof F. C. MD, and Prof S. T. MD. 2011. "KRAS, BRAF, PIK3CA, and PTEN Mutations: Implications for Targeted Therapies in Metastatic Colorectal Cancer." *Lancet Oncology* 12(6):594–603.
- Dunn, Sara-Jane, Inke S. Näthke, and James M. Osborne. 2013. "Computational Models Reveal a Passive Mechanism for Cell Migration in the Crypt" edited by J. Vera. *PLoS ONE* 8(11):e80516–18.
- Evan, Gerard I. 2006. "Can't Kick That Oncogene Habit." *10(5):345–47.*
- Fischer, Jared M. . 2016. "Single Cell Lineage Tracing Reveals a Role for TgßR2 in Intestinal Stem Cell Dynamics and Differentiation." *Proceedings of the National Academy of Sciences of the United States of America* 113(43):12192–97.
- Fletcher, Alexander G., Christopher J. W. Breward, and S J. Chapman. 2012. "Mathematical Modeling of Monoclonal Conversion in the Colonic Crypt." *Journal Of Theoretical Biology* 300(c):118–33.
- Fontana, Alessandro. 2010. "An Artificial Life Model for Carcinogenesis.." ALIFE.
- Forbes, Simon A. . 2017. "COSMIC: Somatic Cancer Genetics at High-Resolution." *Nucleic Acids Research* 45(D1):D777–83.
- Ingham-Dempster, Tim, Dawn C. Walker, and Bernard M. Corfe. 2017. "An Agent-Based Model of Anoikis in the Colon Crypt Displays Novel Emergent Behaviour Consistent with Biological Observations." *Royal Society Open Science* 4(4):160858–13.
- Kozar, Sarah . 2013. "Continuous Clonal Labeling Reveals Small Numbers of Functional Stem Cells in Intestinal Crypts and Adenomas." *Stem Cell* 13(5):626–33.
- Lao, Victoria V. and William M. Grady. 2011. "Epigenetics and Colorectal Cancer." *Nature Reviews Gastroenterology & Hepatology* 8(12):686–700.
- Meineke, F A., C S. Potten, and M Loeffler. 2001. "Cell Migration and Organization in the Intestinal Crypt Using a Lattice-Free Model." *Cell Proliferation* 34(4):253–66.
- Mirams, Gary R., Alexander G. Fletcher, Philip K. Maini, and Helen M. Byrne. 2012. "A Theoretical Investigation of the Effect of Proliferation and Adhesion on Monoclonal Conversion in the Colonic Crypt." *Journal Of Theoretical Biology* 312(C):143–56.
- Nachman, M W. and S L. Crowell. 2000. "Estimate of the Mutation Rate Per Nucleotide in Humans." *Genetics* 156(1):297–304.
- Pan, Tianhui, Jinghong Xu, and Yongliang Zhu. 2017. "Self-Renewal Molecular Mechanisms of Colorectal Cancer Stem Cells.." *International Journal of Molecular Medicine* 39(1):9–20.
- Pino, M S. and D C. Chung. 2010. "The Chromosomal Instability Pathway in Colon Cancer." *Gastroenterology* 138(6):2059–72.
- Rejniak, Katarzyna A. and Alexander R. A. Anderson. 2010. "Hybrid Models of Tumor Growth." *Wiley Interdisciplinary Reviews: Systems Biology and Medicine* 3(1):115–25.
- Ritsma, Laila . 2015. "Intestinal Crypt Homeostasis Revealed at Single-Stem-Cell Level by in Vivo Live Imaging." *Nature* 507(7492):362–65.
- Rubben, A and O Nordhoff. 2013. "A Systems Approach Defining Constraints of the Genome Architecture on Lineage Selection and Evolvability During Somatic Cancer Evolution." *Biology Open* 2(1):49–62.
- Sartore-Bianchi, A, F Loupakis, G Argilés, and G W. Prager. 2016. "Challenging Chemoresistant Metastatic Colorectal Cancer: Therapeutic Strategies From the Clinic and From the Laboratory." *Annals of Oncology* 27(8):1456–66.
- Sato, Toshiro . 2011. "Paneth Cells Constitute the Niche for Lgr5 Stem Cells in Intestinal Crypts." *Nature Publishing Group* 469(7330):415–18.
- Smallbone, Kieran and Bernard M. Corfe. 2013. "A Mathematical Model of the Colon Crypt Capturing Compositional Dynamic Interactions Between Cell Types." *International Journal of Experimental Pathology* 95(1):1–7.
- Snippert, Hugo J. . 2010. "Intestinal Crypt Homeostasis Results From Neutral Competition Between Symmetrically Dividing Lgr5 Stem Cells." *Cell* 143(1):134–44.
- Snippert, Hugo J., Arnout G. Schepers, Johan H. van Es, Benjamin D. Simons, and Hans Clevers. 2013. "Biased Competition Between Lgr5 Intestinal Stem Cells Driven by Oncogenic Mutation Induces Clonal Expansion." *EMBO reports* 15(1):62–69.
- Strubberg, Ashlee M. and Blair B. Madison. 2017. "MicroRNAs in the Etiology of Colorectal Cancer: Pathways and Clinical Implications." *Disease Models & Mechanisms* 10(3):197–214.
- Sunter, J P., D R. Appleton, M S. Dé Rodriguez, N A. Wright, and A J. Watson. 1979. "A Comparison of Cell Proliferation at Different Sites Within the Large Bowel of the Mouse.." *Journal of anatomy* 129(Pt 4):833–42.
- Swanson, Kristin R., Carly Bridge, J D. Murray, and Ellsworth C. Alvord Jr. 2003. "Virtual and Real Brain Tumors: Using Mathematical Modeling to Quantify Glioma Growth and Invasion." *Journal of the Neurological Sciences* 216(1):1–10.
- Totalfurno, J, M Bjercknes, and H Cheng. 1987. "The Crypt Cycle. Crypt and Villus Production in the Adult Intestinal Epithelium." *Biophysical Journal* 52(2):279–94.
- Traverso, Giovanni . 2002. "Detection of APC Mutations in Fecal DNA From Patients with Colorectal Tumors.." *The New England journal of medicine* 346(5):311–20.
- Tsang, Andy H.-F. 2014. "Current and Future Molecular Diagnostics in Colorectal Cancer and Colorectal Adenoma." *World Journal of Gastroenterology* 20(14):3847–12.
- Vermeulen L, Morrissey E, van der Heijden M, Nicholson AM, Sottoriva A, Buczacki S, et al. Defining stem cell dynamics in models of intestinal tumor initiation. *Science*. 2013;342: 995–998.
- Williams, Marc J., Benjamin Werner, Chris P. Barnes, Trevor A. Graham, and Andrea Sottoriva. 2016. "Identification of Neutral Tumor Evolution Across Cancer Types." *Nature Genetics* 48(3):238–44.