University of Southampton

Faculty of Engineering and Physical Sciences

Development of 3D X-ray phase-contrast Imaging and Analysis Tools for Tubular and Branching Structures with Applications in Colorectal Cancer Research

DOI: 10.5258/SOTON/T0026

by

Charalambos Rossides

ORCID 0000-0002-7547-0256

Thesis submitted in partial fulfilment for the degree of Doctor of Philosophy

March 2021

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences

School of Engineering

Thesis for the degree of Doctor of Philosophy

Development of 3D X-ray phase-contrast Imaging and Analysis Tools for Tubular and Branching Structures with Applications in Colorectal Cancer Research

By

Charalambos Rossides

X-ray micro-computed tomography (μCT) is a non-destructive imaging technique used to capture the internal three-dimensional (3D) structure of small samples at high spatial resolutions. As it requires a certain level of knowledge in engineering, material science, mathematics and physics, it and can quickly become challenging to follow for readers with diverse academic backgrounds. In particular, biological tissues can have irregular internal structures, with convoluted shapes and orientations. In such cases, structural quantification becomes remarkably difficult as it often requires advanced knowledge in mathematics (topology) and engineering (digital signal processing). To facilitate research that employs μCT to study the internal microstructure of soft tissues such as the colon, a suite of tools, methods, and workflows have been developed here. The study was motivated by the problem of morphological quantification of tubular glands in the colon called crypts. These begin to branch during early-stage colorectal cancer (CRC), which makes crypt budding of central interest in the study of the disease.

To study crypt budding during early-stage CRC using μCT, a sample-preparation method was developed here that employs a murine model for CRC and preserves the natural tubular shape of the colon. An X-ray phase-contrast μCT imaging protocol-design flowchart was developed to design X-ray imaging protocols for optimal spatial resolution. The flowchart was then employed to non-destructively image the tubular murine colon samples. These were subsequently unrolled digitally, using a tool developed here called '3D cyclorama'. The unrolled views of the colonic samples were used to identify the budding crypts, and a detailed 3D structural phenotype was compiled, employing a mathematical modelling approach that quantified the morphology of individual branches within the crypt.

The tools, methods and workflows developed here, although motivated by the study of early-stage CRC, are applicable to a wide range of problems. The protocol-design flowchart will help researchers to understand, adapt and apply key elements of phase-contrast X-ray imaging for their own research questions, involving materials and tissues that do not significantly absorb X-rays. Digital unrolling may be applied to structures including, the colon to study crypt budding, ancient scrolls to retrieve inscribed text, or industrial applications to flatten curved parts. Finally, morphological quantification of budding structures may also be applied to blood vessel and airway networks, or plant roots. Taken together, the work presented in this thesis will help researchers from diverse backgrounds to harness the great potential of X-ray μCT imaging. The developed tools, workflows and methods will support various studies by facilitating the extraction of useful information from μCT data of samples with challenging geometries.

I am writing this thesis in lockdown, during a pandemic that has quelled the entire world!

The recent spread of a SARS-CoV-2 virus resulted in a global outbreak of the disease known as CoViD-19, which has claimed more than 2 million lives, worldwide, to date.

In memory of those who perished…

Table of Contents

Table of Tables

Table of Figures

Table of Figures

Table of Figures

List of Accompanying Materials

S1. Standard operating procedure: Induction of colorectal cancer to murine tissue samples – Study 1, May 2017

S2. Dataset supporting the work presented in Chapters 2, 3, 4 & 6, and Appendices A & B
DOI: 10.5258/SOTON/D1565

Research Thesis: Declaration of Authorship

Print name: Charalambos Rossides

Title of thesis: Development of 3D X-ray phase-contrast imaging and analysis tools for tubular and branching structures with applications in colorectal cancer research

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as: In progress:
   C. Rossides, S. Pender, and P. Schneider. “Phase-contrast imaging in practice: Protocol design flowchart for X-ray propagation-based phase-contrast imaging”. Under review (2021).
   C. Rossides, S. Pender, and P. Schneider. “3D cyclorama: Digital unrolling of deformed tubes”. Under review (2021).;

Signature: ..... Date: 1st March 2021

Acknowledgements

This study was supported by the Engineering and Physical Sciences Research Council (EPSRC), United Kingdom (UK), the Institute for Life Sciences (IfLS), University of Southampton (UoS), UK and Nikon X-Tek Systems Ltd., Tring, UK.

I want to thank my supervisors Philipp Schneider and Sylvia Pender. Philipp, you have been a true mentor to me! Your guidance and support helped me improve as an academic and build an invaluable set of skills. For this, I am deeply grateful. The immense patience from your side, the comforting words when needed, and the natural kindness, made it a pleasure to work with you, even when receiving strict and exhaustive feedback. Sylvia, thank you for your much needed guidance through foreign territories to me. Especially, thank you for the patience and hands-on support with respect to laboratory-related work.

I would also like to thank the µ-VIS X-ray Imaging Centre, UoS, the biomedical research facility (BRF), UoS, the histochemistry research unit (HRU), UoS and the biomedical imaging unit (BIU), UoS for granting access to their facilities. Special thanks goes to Jonathan Ward and Jenny Norman for their help in the HRU and the mood boost during long days of repetitive work. Thanks also to the diamond light source (DLS), Didcot, UK, and the Swiss light source (SLS), Villigen, Switzerland for granting beamtime for synchrotron experiments, and the staff Kazimir Wanelik, Andrew Bodey, Shashidhara Marathe, and Elena Borisova for their support and guidance during the experiments.

Substantial aid was received from UoS staff, Orestis Katsamenis and Richard Boardman, who were not formally required to provide any. Orestis, I am grateful for the guidance throughout these years, for providing that extra support when needed, and for the extensive help during the first experiment at DLS. Rich, nothing would have been possible without your backup on the technical side. Thank you for going the extra mile and helping with whatever absurd technical need I asked help with, and for resolving all urgent issues in time. Thanks goes also to David Chatelet, for the fruitful discussions regarding biomedical image processing and for the thorough testing of the 3D Cyclorama plugin. Thanks to Dimitrios G. Fatouros, Georgios K. Eleftheriadis, and Orestis L. Katsamenis for sharing the µCT stack of the pharmaceutical film as a case study for the 3D cyclorama plugin. Special thanks goes to Joseph Bevitt and Carla Raymond for accepting the offer to collaborate on an international level, in an attempt to digitally unroll a rare ancient artefact using 3D cycloramas.

Acknowledgements

Overnight, lengthy shifts over the course of several days were necessary during the three synchrotron experiments performed for this study. For this, I am grateful to Orestis Katsamenis, Katy Williams, Elis Newham, Alisha Sharma and Elaine Ho, who spent time and energy to help conduct these experiments. Thank you for this great sacrifice, for the team spirit and the great time we had together. These trips will never be forgotten!

Moral support and the sense of belonging, emanating from working in the same office together with co-PhD students kept me going during the years. Thanks goes to all colleagues at Highfield and the Southampton general hospital (SGH). Special thanks to Erik Meilak, Tomas Mabey, Joshua Steer, Sara Ferri, Jenny Bramley and Gareth LuTheryn for the fun times at Highfield, as well as Becky Holding, Olivia Lee, Rachel Meadows and Ahmad Alzahrani at SGH. Thanks goes to Matthew Lawson for the fruitful discussions we had over the years, and to Patricia Goggin who was always bringing an uplifting mood. Thanks also to Eileen Ying Li for her help related to data analysis of the synchrotron scans.

Lastly, I would like to thank my family, Chrystalla Rossidou, Michalis Rossides and Timothea Rossidou, for being emotionally and financially by my side, for their love and support throughout my entire life. Whatever I have achieved or will achieve in the future, is a culmination of your efforts and self-sacrifice. For that I am truly grateful and could not thank you enough. Special thanks goes to my partner Stella Manoli, for believing in me, for supporting me during the hard times and being by my side no matter what.

Definitions and Abbreviations

Abbreviation	Full form
2D	two-dimensional
3D	three-dimensional
3R	Replacement Refinement & Reduction
Å	Angstrom unit: $1Å = 10^{-10}m$
a.u.	arbitrary units
ad libitum	Latin: at one's pleasure
ACF	Aberrant Crypt Focus, Aberrant Crypt Foci
AFM	Atomic Force Microscopy
AI	Asymmetry Index
ANSTO	Australian Nuclear Science and Technology Organisation
AOM	Azoxymethane
APC	Adenomatous Polyposis Coli
ARRIVE	Animal Research: Reporting of In Vivo Experiments
ASPA	Animals Scientific Procedures Act
AWERB	Animal Welfare Ethics Research Board
bit	binary digit
BIU	Biomedical Imaging Unit
BMP	Bone Morphogenetic Protein
BRF	Biomedical Research Facility
CAC	Colitis Associated Cancer
CD	Crohn's Disease
CFI	Crypt Fission Index (see also FI)
DAPI	4',6-diamidino-2-phenylindole
DMH	1,2-dimethyldrazine
DNA	Deoxyribonucleic Acid
DSS	Dextran Sodium Sulphate
DTT	Dithiothreitol
DTW	Dynamic Time Warping
e.g.	for example
EDTA	Ethylenediaminetetraacetic acid

Abbreviation	Full form
et al.	et alia (Latin): and others
etc.	et cetera: further, similar items are included
EtOH	Ethanol
ex vivo	takes place outside an organism
FAP	Familial Adenomatous Polyposis
FFPE	Formalin-Fixed Paraffin-Embedded
FI	Fission Index (see also CFI)
Fiji	Fiji is just imageJ
FoV	Field of View
GB	Giga Byte
H&E	Haematoxylin and Eosin
HCL	Hydrogen Chloride
HMF	Highest Mitotic Figure
HRU	Histochemistry Research Unit
i.e.	id est (Latin): that is
in silico	by means of computer modelling or simulation
in situ	Latin: in its original place
in vitro	Latin: within the glass | Work performed outside a living organism
in vivo	Latin: within the living | Work performed in a whole, living organism
I.P.	intra-peritoneal
IMS	Industrial Methylated Spirit
IBD	Inflammatory Bowel Disease
IQR	Inter-Quartile Range
MAC	MACQUARIE University Museum of Ancient Cultures
MBA	Modified Bronnikov Algorithm
MIN	Multiple Intestinal Neoplasia
MT-PBS	Mouse Tonicity Phosphate Buffered Saline
NBF	Neutral Buffered Formaldehyde
OCT	Optical Coherence Tomography
PBS	Phosphate Buffered Saline
pixel	Portmanteau word: pix, i.e. picture
PSF	Point Spread Function
RAM	Random Access Memory

Abbreviation	Full form
RNA	RiboNucleic Acid
ROI	Region Of Interest
SEM	Scanning Electron Microscopy
SNR	Signal to Noise Ratio
SOP	Standard Operating Procedure
TIE	Transport-of-Intensity Equation
UC	Ulcerative Colitis
voxel	Analogous to pixel, with ‘vo’ representing volume
WHO	World Health Organisation
Wnt	Portmanteau word: Wingless & Int-1
wt/vol	weight divided by volume
µCT	micro-Computed Tomography
µm^2	10^-12 m^3
µm^3	10^-18 m^3

Chapter 1 Introduction and background

1.1 Introduction

1.1.1 3D imaging as a tool for interdisciplinary research

Putting together a team of researchers coming from distinct backgrounds is not enough for performing interdisciplinary research. Combining skills from several disciplines is necessary, but key to interdisciplinary research is having unified problem definitions so that the gap between researcher's skills and ideas can be bridged [1]. Often researchers from diverse backgrounds view a given problem from different perspectives, or they use different terminology and scientific methods to approach it [2]. Truth is, that advanced knowledge in more than one scientific fields is typically required from all parties to communicate effectively.

A prime example is X-ray computed tomography (CT), which has been introduced as a three-dimensional (3D) clinical imaging technique in the 1970s and subsequently evolved to image small samples at high spatial resolutions (micro-computed tomography or μCT) [3]. Today, application areas of μCT are broad, including archaeology, art restoration, astrophysics, biology, chemistry, environmental science, forensics, material science, palaeontology and physics [4, 5]. Being an advanced imaging technique, which generally requires a certain level of knowledge in engineering, material science, mathematics and physics, it can quickly become challenging to follow for readers with different academic backgrounds.

Even after imaging, μCT data analysis is not straightforward. The technique produces 3D representations of samples in the form of image sequences (image stacks) that depict its internal structure. Especially for biological tissues, internal features such as blood vessels, villi, etc. can have irregular and convoluted shapes and orientations in 3D. In such cases, structural quantification requires advanced knowledge in mathematics (topology) and engineering (digital signal processing), and as a result, it is often difficult to harness the full potential of μCT.

A specific biological problem is the study of the morphological changes in the colon during the early developmental stages of colorectal cancer (CRC). As discussed in the following sections, understanding the budding process of colonic crypts (tubular glands in the colon) is crucial to understand the early stages of CRC development. A key argument is that certain limitations of traditional tools being used in the study of crypt budding can be overcome by introducing X-ray μCT imaging to complement standard approaches.

1.1.2 The need for new tools in the study of crypt budding

In health, colonic crypts multiply through a bifurcation process called crypt fission (splitting of one crypt into two) [6-10]. Along with the recently observed mechanism of crypt fusion [11, 12] (merging of two crypts into one), a fine balance between the two mechanisms appears to be pivotal for maintaining homeostasis in the colon. Crypt fission is, however, also involved in the pathway to colonic carcinogenesis. Early lesions are believed to originate and replicate through crypt fission [13-19], which eventually evolve to excessive abnormal budding [15, 19-21] (see section '1.2.3 Crypt fission, budding and dysplasia'). Studies focusing on the morphology of budding crypts have subsequently differentiated crypt fission and asymmetric budding [6, 9, 22]. The latter appears to be the aberrant case of crypt formation, as opposed to the normal case of crypt fission, which exhibits symmetric crypt bifurcation. Yet, although crypt fission has been studied for decades, the understanding of the underlying mechanics and the budding crypt morphology remains poor [6, 23, 24].

Methods that enable concurrent functional and morphological studies are crucial for experimental validation of hypotheses that involve morphological changes as a consequence of underlying biological processes. Such methods would enable experimenters to concurrently quantify morphological changes (budding) and biological processes (protein expression), both of which being relevant to the study of CRC development [24]. However, it becomes evident in section '1.3 Standard approaches for crypt budding research' that one of the reasons for the lack of understanding of the crypt's morphological evolution is that, existing methods for the study of crypt budding, including sample preparation and imaging, are not suited for concurrent functional and morphological studies. The currently employed methods for sample preparation include, among others, microdissection, paraffin serial sectioning and mucosal digestion. Imaging methods include among others, light microscopy of serial sections, confocal microscopy and scanning electron microscopy (SEM).

As discussed in section '1.3.4 Computational models of the intestine', computational models have been proven invaluable in furthering the understanding of crypt formation in health and disease. They put forward hypotheses emerging from mathematical modelling of the biological processes, which are subsequently tested through experimental approaches that employ the methods briefly mentioned above. In particular, spatial models are capable of linking observations of structural changes such as crypt budding during the early-stages of CRC (see section '1.2.3 Crypt fission, budding and dysplasia'), with underlying biological processes such as specific gene mutations (see section '1.2.2 Genetic background of colorectal cancer and crypt fission'). However, since the early modelling approaches of Totafurno, Bjerknes *et al.* [10, 25, 26], validation of spatial models

for crypt budding against experimental data has been challenging [6, 27]. As becomes evident in section '1.3 Standard approaches for crypt budding research', methods that enable quantification of large numbers of crypts (e.g. mucosal digestion), typically remove the crypts from their native environment. As a result, they are not suitable for 3D morphological quantification. On the other hand, methods that enable accurate morphological quantification in 3D (e.g. serial sectioning, SEM imaging, 3D confocal imaging) are time-consuming and do not allow concurrent imaging of a large number of crypts. Moreover, methods such as SEM imaging often require destructive sample preparation protocols (staining with metals), which prohibit functional assays of the imaged sample. Therefore, no currently existing method is ideal for *in situ* 3D imaging of a large number of crypts at once, while enabling functional assays at the same time (see Table 1.1).

To this end, a workflow is proposed, methods and software tools are developed in this thesis for the study of crypt budding in the murine colon, which enable non-destructive 3D imaging of samples embedded in paraffin wax (standard procedure). This enables functional assays following standard approaches, while providing structural information for large numbers of crypts in 3D. The key rationale supporting the design of this workflow revolves around the following principles: A well-established sample preparation method is adopted that enables functional assays, a non-destructive imaging method (X-ray µCT) is employed to image large number of crypts in 3D and the natural shape of the colon is retained so that morphological quantification is not biased by the sample preparation method.

In this chapter, an overview of CRC biology revolving around the microstructure of budding colonic crypts is examined in section '1.2 Colorectal cancer and crypt budding', followed by methods currently in use by researchers in section '1.3 Standard approaches for crypt budding research'. Principles of advanced X-ray µCT imaging are discussed in section '1.4 Advanced imaging approaches in research' and the objective and the specific aims of this thesis are outlined in section '1.5 Thesis overview'.

1.2 Colorectal cancer and crypt budding

Colorectal or bowel cancer refers to cancers that emerge in the large intestine, i.e. the part of the gastrointestinal system that includes the colon and rectum. It is currently the third most common cause of cancer-related deaths worldwide and the second in Europe [28]. Risk factors that have been linked to CRC include [29] (i) age: 60 % of UK CRC patients are over 70 years old; (ii) family history: patients with inherited predisposing conditions such as familial adenomatous polyposis (FAP), or long history of inflammatory bowel diseases (IBD) including ulcerative colitis (UC) and

Crohn's disease (CD). Other risk factors include [29] (iii) smoking tobacco, (iv) obesity and (v) diet with too much red and processed meat.

Differentiating cases according to heredity yields two distinct subtypes, familial (hereditary) and sporadic CRC [30]. Further clarification of the characteristics of the disease, yields terms such as inflammation-associated CRC, otherwise called colitis-associated cancer (CAC).

1.2.1 Crypt formation and mucosal homeostasis

Figure 1.1 shows a schematic representation of the human colon and a representative histology section stained with haematoxylin and eosin (H&E). The human colon is composed of four main layers: mucosa, submucosa, muscle layers and serosa. The mucosa is the layer closer to the colonic lumen, and it is lined by a simple columnar epithelium (a single layer of epithelial cells) that forms the inner surface of the lumen. The epithelium or epithelial layer is a continuous sheet of cells comprising a regular pattern of colonic crypts. Colonic crypts, or simply crypts, are mucus, tubular glands with their base (bottom) located at the mucosal/submucosal layer interface, and their top forming an opening into the colonic lumen. The interior of the tubular crypts is called crypt lumen and the imaginary axis passing through its centre is called crypt axis [24, 31]. Each crypt contains two to three thousand cells [32] with pluripotent (i.e. can give rise to different cell lineages) stem cells (SCs) at its base. The epithelial layer is renewed through a process of continuous cell proliferation by SCs at the crypt bases. SCs generate rapidly proliferating cells that differentiate into various cell lineages (absorptive, goblet, etc.) while migrating along the crypt epithelium (epithelial cells) [24]. When they reach the top of the crypt, the differentiated cells die and are shed into the colonic lumen (apoptosis). Maintenance of the balance between cell production and apoptosis is key to mucosal homeostasis and viability [21, 33]. When this balance is lost, colorectal tumorigenesis is initiated in the colonic mucosa with a pattern of abnormal cell replication and enhanced capability of epithelial cells to proliferate. The uncontrollable cell proliferation (excess growth) results in the development of benign polyps (abnormal tissue growths) that eventually differentiate into cancerous tumours [34].

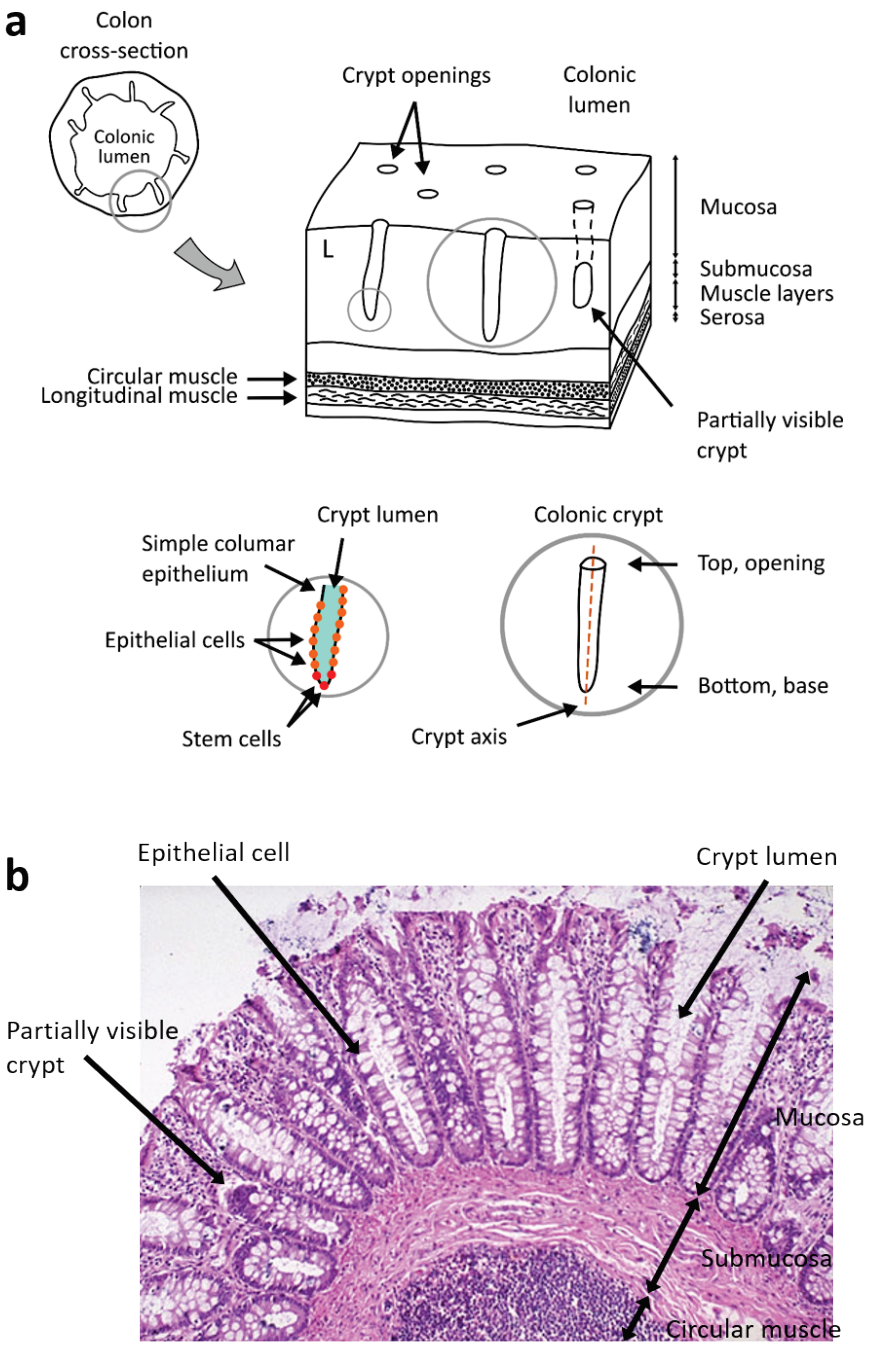


Figure 1.1: Human colonic histology

(a) A 3D schematic of the tissue layers in the healthy human colon. (b) A Haematoxylin and Eosin histology section of a healthy human colon corresponding to plane 'L' in a. Crypts are sliced longitudinally. When the section cuts through a crypt at an angle, the crypt is only partially visible. This image was published in [35], Page 194, Copyright Elsevier (2020). Reproduced with permission from Elsevier.

The murine large intestine is about 14 cm long [36], composed of the caecum, colon and rectum, which ends at the anal orifice (Figure 1.2a) [37, 38]. The colon is further subdivided in a proximal

part closer to the caecum and a distal part closer to the rectum. Histologically, the murine colon is similar to the human colon, both made of the mucosal, submucosal, circular and longitudinal muscle layers and a thin serosal layer (Figure 1.2b). Similarly to the human colon, the normal murine mucosa comprises a uniform pattern of evenly distributed, colonic crypts (Figure 1.2b) [39]. Each crypt in the murine colon contains 235 to 250 cells [40, 41] arranged in a simple columnar epithelium that covers the surface of the colonic crypt lumen. This surface epithelium constitutes a continuous selective barrier that spans the entire mucosal surface.

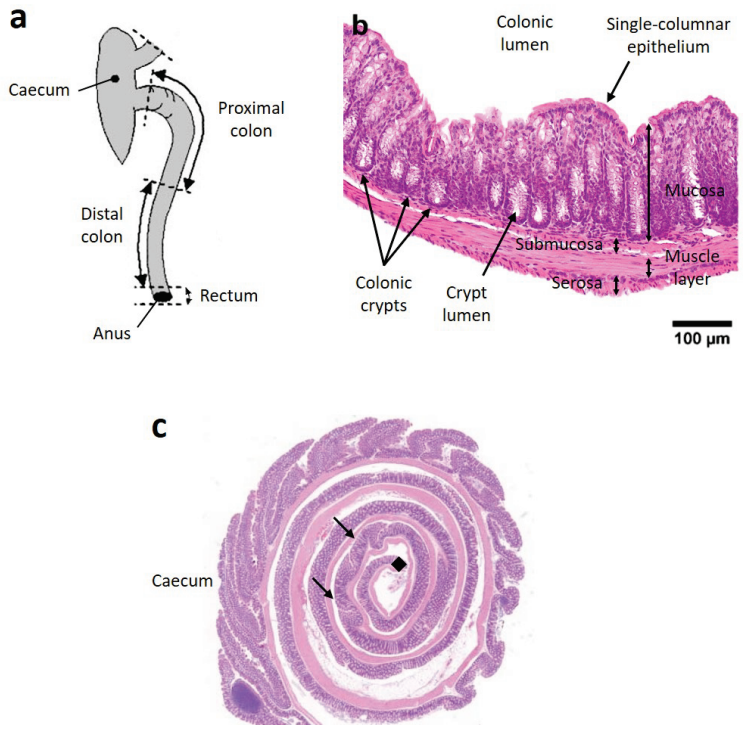


Figure 1.2: Murine colon anatomy and histology layers

(a) Schematic representation of the large intestine anatomy. Image adapted from [42]. (b) Haematoxylin and eosin (H&E) histology section of a healthy murine colon imaged with an Olympus VS110 light microscope. (c) A murine large intestine coiled into a Swiss roll (see section '1.3.2 Standard histological methods for the study of crypt budding'), cut and stained with H&E, featuring the entire large intestine, from the caecum to the rectum (◆). The arrows show longitudinal folds in the distal colon. Image adapted from [37].

As mentioned above, SCs play a fundamental role in the crypt and mucosal homeostasis, in both human and murine colons. The widely popular 'Unitarian hypothesis' suggests that a single SC generates all cell lineages [43-46]. Others support the concept of several SC types [47-50] arranged in a pool of stem and secretory cells (e.g. $Muc2^{+}$) [22, 51]. Regardless of the SC type though, it has long been clear that crypt fission (the mechanism that drives crypt production via a

bifurcation process of existing crypts) is initiated as budding from SCs close to the crypt base [6-10]. Crypt fission begins as an indentation (otherwise referred to as fissure or bud) at the crypt base that progressively moves along the crypt until the parent crypt is eventually split into two separate daughter crypts [10, 52]. Through this process, new crypts are created mostly during a short postnatal period [8, 53, 54] and during recovery from acute trauma (e.g. irradiation [55] or treatment with chemical agents [56]).

Recent studies by Langlands and colleagues [22] suggest that crypt fission is initiated due to biomechanical stiffness that is built up as a consequence of the arrangement (relative position) and numbers of different cell types at the crypt base. The authors suggest that tissue stiffness differentials are created due to an alternating arrangement, and the spatial concentrations, of stem ($Lgr5^{+}$) and secretory (Paneth) cells, typically found in the small intestine. They extrapolated their findings relating the generated tissue stiffness differentials and crypt fission to the colon, by identifying Muc2 secretory cells as the colonic counterparts of Paneth cells. Next, they validated their hypothesis employing organoid cultures (*in vitro* cultures derived from stem cells that maintain many aspects of the normal colonic epithelial organisation).

Such studies that begin to elucidate the link between the colonic crypt structure and stem cell behaviour are discussed later on, in sections '1.2.3 Crypt fission, budding and dysplasia', '1.3.3 Morphological quantification of crypt budding' and '1.3.4 Computational models of the intestine'. As the earliest signs of CRC, are of morphological nature, understanding this link is important to understand the biological mechanisms that govern the early developmental stages of CRC. Notably, it becomes clear later on that, as these first studies begin to attribute crypt budding to certain biological processes (e.g. tissue stiffness differentials forming due to specific cell arrangements), it remains to understand what specific morphological changes may emerge out of these processes. This is achieved by structural computational models of the intestine, discussed in section '1.3.4'.

1.2.2 Genetic background of colorectal cancer and crypt fission

From the genetic perspective, cancer progression is driven by a series of gene mutations in a scheme referred to as the 'multi-hit' cancer model [57]: In brief, the 'classical' genetic modification pathway is initiated with a loss of function of the tumour-suppressing gene adenomatous polyposis coli (APC). Most of the affected cells exhibit one or two mutations of the APC gene before a mutation of the K-Ras gene leads to the development of an adenomatous polyp or adenoma. The polyp turns into a malignant carcinoma after inactivation of the p53 gene. Although the exact order and number of mutations necessary for cancer development is still

debatable, it is clear that cancer incidence is related to age, which provides the time necessary for mutations to accumulate [52, 57]. This makes SCs the primary suspects for the accumulation of mutations, since they are the only cells in the colonic crypt that remain in place long enough for this to happen [52]. This brings attention to mutations of the APC gene in colonic SCs as the initiating event of CRC. The APC protein within SCs controls cell division [58, 59], thus regulating the production of colonic crypts via the fission process [9]. Supporting this statement, increased crypt fission rates have been observed as a result of APC mutation [9, 60-62], suggesting that disturbed stem cell division patterns affect the rates of crypt fission [9, 52, 63].

Several signalling pathways contribute to crypt homeostasis, including the wingless & int-1 (Wnt), the bone morphogenetic protein (BMP), the Notch and the Hippo pathways [27, 64-67]. Wnt is the principal signalling pathway in CRC development as mutations in Wnt components occur first [64]. As detailed by van Leeuwen *et al.* [39], APC is involved in the Wnt signalling pathway via regulation of the levels of β-catenin, which in turn regulates the gene expression of cyclin-D1, survivin, c-myc and ephrin-B. Cyclin-D1 controls the cell cycle [68], survivin prevents cell apoptosis [69], c-myc controls cell differentiation [70], and ephrin-B regulates cell migration and adhesion [71]. In the healthy colon, APC down-regulates Wnt signalling [72, 73]. Therefore deregulation of the levels of β-catenin through the loss of APC function perturbs the Wnt signalling pathway [74-76] in a way that it is continuously activated [77-79]. This leads to excess mitotic rates and thus, colonic crypt formation through increased fissioning [6].

The concentrations of the proteins mentioned above are spatially varying, for instance along the crypt axis [24, 69, 80, 81]. Interestingly, Wnt components including β-catenin and survivin are found in high concentrations at the bottom, and lower concentrations at the top (see Figure 1.3). In contrast, the concentration of APC is inversed, i.e. APC concentrations are higher at the top and lower at the bottom of the crypt. Boman *et al.* [24] suggest that these inverse concentration gradients, in conjunction with the fact that APC down-regulates Wnt signalling, directly affect the microstructure of crypt fission. Specifically, Boman *et al.* argue that since both Wnt components and APC are required for mitosis (i.e. fission or budding) and since their concentrations are inversed along the crypt axis, there should be a 'sweet spot' that is optimal for fissioning.

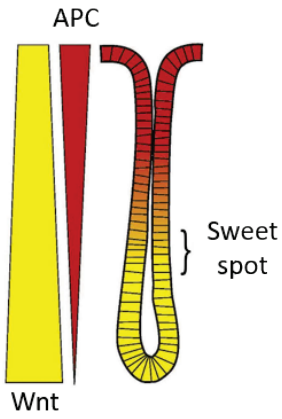


Figure 1.3 APC and Wnt concentration gradients along the crypt axis

Wnt components are found in high concentrations at the bottom of the crypt, whereas the concentration of APC is inversed, i.e. APC concentrations are higher at the top [24, 69, 80, 81]. These inverse concentration gradients, in conjunction with the fact that APC down-regulates Wnt signalling, suggest that there should be a *sweet spot* that is optimal for fissioning [82]. Image adapted from Boman *et al.* [82].

Further, deregulation of the Wnt signalling pathway and APC mutations during the early stage of CRC, or in predisposing conditions (e.g. FAP or IBD patients), would affect the position of the 'sweet spot' for fissioning along the crypt axis, resulting in fissioning higher up the crypt (see Figure 1.3). The relationship between APC mutations and CRC initiation, which remains to be elucidated [24], goes beyond the scope of this thesis. However, the work of Boman *et al.* reveals that studies linking the spatial concentrations of proteins and the crypt structure, require the employed methods to provide both structural and functional information concurrently. However, as discussed in section '1.3.2 Standard histological methods for the study of crypt budding', currently available methods for this purpose are often time consuming, and destructive.

1.2.3 Crypt fission, budding and dysplasia

As discussed above, the underlying biological processes (deregulation of signalling pathways, protein concentrations etc.), potentially result in structural alterations. These alterations are typically studied as changes in the normal histology of the colonic layers. It is believed that carcinogenesis begins with a series of mutations in a stem cell of a normal crypt that develops into a monocryptal lesion, i.e. a lesion consisting of a single mutated crypt [39]. Tissue disorganisation characteristic of dysplasia (structural and cytological alterations, predisposing the tissue to cancer development) is considered to be a result of abnormal and excessive crypt fission [13, 19, 20, 62, 83]. Wong and colleagues [19] underlined the importance of crypt fission in the formation of

adenomas in human CRC, thus bringing the attention to abnormal crypt fission or budding as a putative mechanism behind tumour initiation and growth [9, 52, 84].

Park *et al.* [85] showed that crypts with mutated stem cells develop into monoclonal patches (wherein all cells are derived from a single progenitor) through crypt fission, which was also observed by others [52]. This process of fissioning of mutated monocryptal lesions results in the formation of aberrant crypt foci (ACF) [39, 86] that eventually become adenomas. Adenomas subsequently evolve into cancer (adenocarcinomas) through increasing stages of dysplasia, following histological alterations widely referred to as the 'adenoma - carcinoma sequence' [87-89]. A closer look at the morphology of murine crypts undergoing fission was taken by Tan *et al.* [6], which is discussed in detail in section '1.3.3 Morphological quantification of crypt budding'. In brief, they distinguished between the normal and the abnormal cases of crypt fission and crypt budding, respectively, based on morphological arguments of symmetry of the crypt's shape. In fission the crypt bifurcates symmetrically, resulting in two structurally identical daughter crypts. In contrast, in the case of crypt budding, a daughter crypt grows (asymmetrically) on the side of the parent crypt. Crypt budding is typical in FAP patients and it can be modelled in rodents treated with carcinogenic substances [52]. Tan *et al.* [6] hypothesised that since the normal crypt production sequence is perturbed in the budding case, this might lead to crypts with more than one branches or daughters, thus resulting in dysplastic crypts similar to those found in adenomas by Wong *et al.* [19]. Pin *et al.* [90] showed that crypt budding could be attributed to tissue stiffness differentials due to the arrangement of cells of different types above the crypt base, which results in trifurcating crypts or even crypts with more than three branches in crypt organoids [22].

These recent studies on the structural characterisation of crypt fission and budding, relate the early microstructural changes of the colonic crypts with the well-established adenoma - carcinoma sequence that describes the histological progression of CRC. As crypt budding takes place even before the observation of ACF (early lesions visible on the mucosal surface, currently regarded as the earliest noticeable change in the colonic structure that are visible microscopically [17]), the structural quantification of crypt budding takes a central role in the study of the early developmental stages of CRC.

1.3 Standard approaches for crypt budding research

Taken together, the studies by Pin *et al.* [90], Langlands *et al.* [22] and Tan *et al.* [6] indicate an increasing interest in the study of the morphology of branching crypts (including fission and budding). Therefore, this section discusses existing tissue processing, imaging and computational

methods for crypt budding research, which are relevant to morphological studies. In the next section '1.4 Advanced imaging approaches in research', alternative imaging techniques are presented, namely phase-contrast X-ray CT, which can be used for studying the crypt structure but has not been applied for this purpose yet.

1.3.1 Animal models for colorectal cancer

Animal models have proven invaluable in furthering our understanding of the development of CRC. They provide a controlled and systematic approach to reliably and efficiently reproduce aspects of the disease. De Robertis *et al.* [91] identify three categories of animal models for CRC: xenoplant, genetic and chemically-induced models. In xenoplant models, tumour cells are induced through injection in individuals of immunodeficient mouse strains when tumour progression is of interest. Although straightforward and easy to use, xenoplant models neglect the complexity of tumour initiation and its interaction with its microenvironment. In genetic models, mice carrying mutations of specific genes, including the APC, K-ras and/or p53, are typically used. Such animal models are preferred when the hereditary aspect of CRC is of interest.

Mutations in the APC gene in FAP can be reproduced by the APC^{MIN} genetic mouse model [92, 93]. Multiple intestinal neoplasia (MIN) is a mutation in the murine APC gene [94] that causes $APC^{MIN/+}$ mice to develop more than 50 adenomas in the small and large intestine that develop into tumours leading to early death in C57BL/6 mice (within 120 days) [95]. This model is useful when the function of APC, as a part of the Wnt signalling pathway, is of interest [96]. Tumours emerging in the APC^{MIN} mouse model typically do not develop into cancer and the histological progression from ACF to adenocarcinoma is not observed [97]. However, it is believed that in both FAP patients and $APC^{MIN/+}$ mice, early lesions develop through unrestricted crypt fission [9, 15, 98], which makes the APC^{MIN} mouse model a candidate tool for studies focusing on crypt budding in hereditary CRC [6, 9].

In chemically induced models, cancer emerges from high dose of chemical carcinogens. They are applicable to mice of different genetic backgrounds, highly reproducible, and the pathogenesis recapitulates human CRC in the histological level. These animal models mimic the multistage process of cancer initiation and progression, but rarely metastasis [91, 99]. Chemically induced models are preferred when studies focus on CRC development at the histological level. A wide range of carcinogens has been used including 1,2-dimethyldrazine (DMH) or its metabolite azoxymethane (AOM) [91]. In early studies of crypt budding, weekly injections of DMH have been employed in rats [100, 101], while later studies focusing on CRC histomorphology [102, 103] have switched to AOM due to its lower price and stability in solution. Repeated injections of a

carcinogen, however, requires several (up to 30) weeks of administration before multiple tumours can be observed. Combining AOM treatment with repeated cycles of inflammation, induced by the inflammatory agent dextran sodium sulphate (DSS), was later shown to decrease the required time for tumour development down to 10 weeks [104, 105]. Variations of the resulting azoxymethane and dextran sodium sulphate (AOM/DSS) rodent model quickly gained popularity among researchers, and it is currently the most popular animal model to study inflammation-associated CRC [91, 99]. It achieves carcinogenesis in two steps: firstly, DNA damage (mutation of K-ras and β-catenin genes) is inflicted by the carcinogenic agent AOM, followed by repeated cycles of inflammation and healing, which boost cancer progression. Inflammation is induced by the inflammatory agent DSS and healing comes naturally once the administration of DSS is withdrawn. The AOM/DSS model quickly and reliably recapitulates the phases of initiation and progression of CRC in humans at the histological level while it is affordable and simple to perform [91, 99, 103]. However, it fails to reproduce APC and p53 mutations, which are characteristic genetic abnormalities in human CRC [106, 107].

1.3.2 Standard histological methods for the study of crypt budding

Researchers focusing on the early stages of CRC have used animal models, in conjunction with observations in human biopsies, to elucidate the initiation and progression of the disease. Workflows for the study of crypt budding have not been systematised, with protocols depending on the scientific question at hand. Studying the concentrations of specific cell types along the crypt (functional information) typically requires histological/immunohistochemical assays, while 3D structural quantification (structural information) calls for protocols that preserve or reconstruct each crypt's natural shape. A significant challenge is to combine the two requirements with experimental protocols that provide both functional and structural information in 3D, especially when large numbers of crypts need to be imaged in order to facilitate statistical comparison and testing.

In this section, the prominent of these methods are reviewed, beginning from the well-established, two-dimensional (2D) methods, and proceeding to methods that are better-suited for concurrent structural and functional imaging. Standard histology, which provides 2D cross-sections of the crypts, is presented in section 1.3.2.1. Next, a method that enables the study of the structure of individual crypts (in 2D) is examined in section '1.3.2.2 Microdissection', and a method that enables the study of a large number of crypts is reviewed in section '1.3.2.3 Mucosal digestion'. Proceeding with methods that minimally tamper with the crypt's natural 3D shape, serial paraffin sectioning, scanning electron microscopy and confocal microscopy are reviewed in sections 1.3.2.4, 1.3.2.5 and 1.3.2.6, respectively. Finally, *in vivo* methods are briefly mentioned in

section '1.3.2.7 Endoscopic methods'. Table 1.1 summarises the key competencies of each method with regards to concurrent structural and functional imaging of large numbers of colonic crypts, *in situ*.

1.3.2.1 Standard histology: paraffin embedding, sectioning and light microscopy

Examination of tissues under the light microscope requires the sample to be prepared so that ~4 µm thick tissue sections can be cut. Standard histology workflows are employed to study the colonic histomorphology [108], which may be briefly summarised as follows: Typically, a tissue sample is initially chemically fixed [109] for 24 hours in a solution of 10 % neutral buffered formaldehyde (NBF), commonly known as 10 % formal saline or formalin. After fixation, the tissue can be stored indefinitely in 70 % ethanol (7 parts of ethanol in 3 parts of water) [19, 110]. During a subsequent tissue processing step, the tissue is dehydrated and infiltrated with paraffin wax before it is embedded into paraffin wax blocks. The wax blocks are then mounted onto plastic cassettes that are attached on a microtome to facilitate cutting or sectioning. Tissue sections are then mounted on glass slides, processed with dyes or stains and are subsequently examined under a light microscope.

Common examples of standard histology workflows include immunostaining (employing antibodies that bind specifically to target proteins within the tissue) to differentiate the several tissue and cell types within the sample. A wide range of tissue stains [111] highlights functional or structural characteristics of the tissue providing a great versatility to histology. One of the most commonly used stains is a combination of haematoxylin and eosin (H&E) [112], where haematoxylin stains nuclei blue (binds to negatively charged molecules revealing DNA, RNA, and acidic proteins in a cell [113]), whereas eosin stains cytoplasm and connective tissue pink [111]. H&E can be used to reveal the colonic tissue layers in humans and rodents. It can also be combined with stains specific to mucin (Alcian blue) and collagen (saffron, or Gieson stain) [108, 114] to reveal further functional information down to the cellular level.

Small tissue samples may be isolated from the excised colons, cut-open and embedded in wax blocks for closer examination of the colonic crypts (Figure 1.4). Paraffin sections from the small colon samples may then be taken perpendicularly to the crypts [17] to study the shapes of their cross-sections (see Figure 1.4a,c). Conversely, to capture the entire crypt on a slide, the cross-section can be taken longitudinally to the crypts, cutting through all tissue layers [115, 116] (see Figure 1.4b,d). In the latter case, crypts that are cut at an angle by the sectioning plane are only partially visible, which makes the process sensitive to the sectioning orientation.

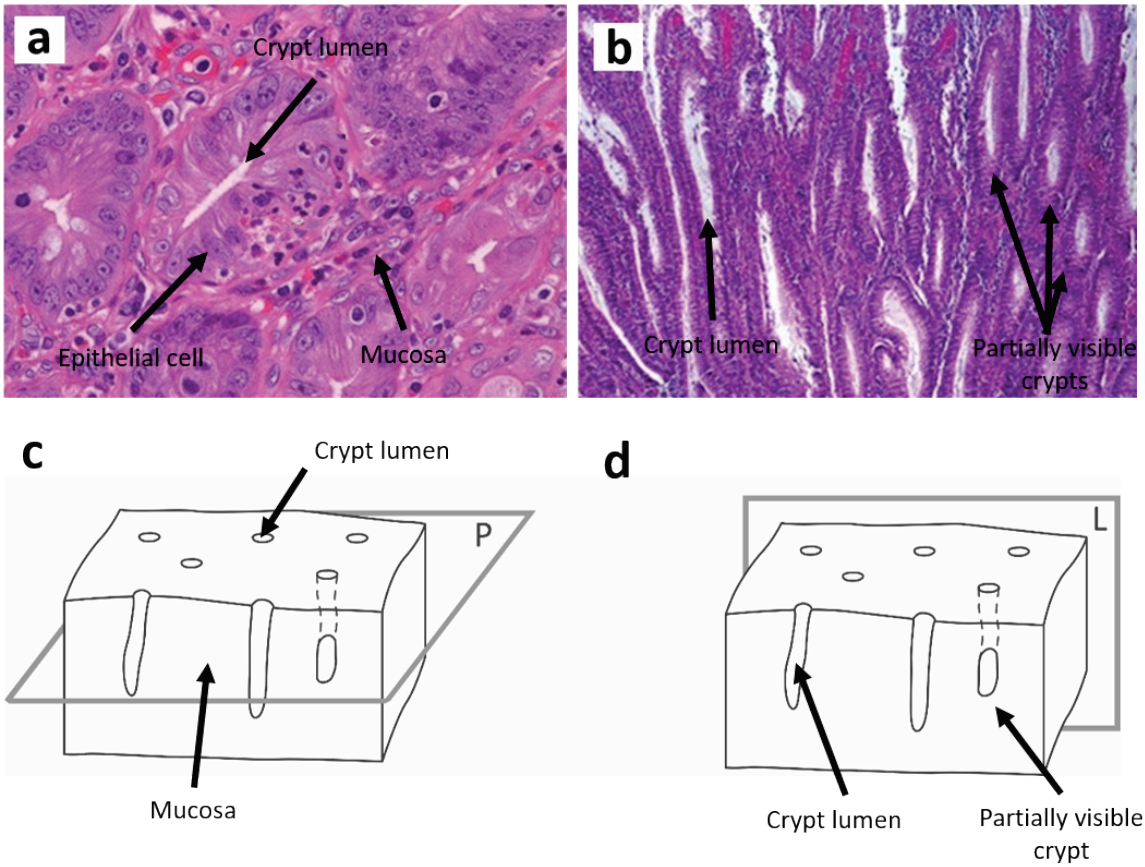


Figure 1.4: Cutting orientations to reveal the colonic crypts

H&E histology slides from the colon of a patient with ulcerative colitis, cut in two different orientations. (a) A representative paraffin section corresponding to plane 'P' in d (perpendicular to the crypts). The cross-sections of the crypt lumens have rounded shapes. (b) A representative paraffin section cut longitudinally to the crypts, corresponding to plane 'L' in d. The cross-sections of the crypt lumens have elongated shapes. Crypts that are cut at an angle by the sectioning plane are only partially visible. Images a and b adapted from [115]. (c & d) Schematic representations of the colonic mucosa, depicting the colonic crypts in 3D.

1.3.2.2 Microdissection

When the morphology of crypt budding is of interest, the challenging shapes of the crypts make it difficult to routinely prepare cross-sections along individual crypts, resulting in partially visible crypts on cross-sections (see Figure 1.1 & Figure 1.4b). Microdissection [117], i.e. using a microscope to assist manual excision of the crypt, can be used in combination with tissue staining to separate individual budding crypts from the mucosa, and study their structure (Figure 1.5). The method [15, 19] requires colon tissue that is excised and stored in 70 % ethanol, to be rehydrated (step-wise immersion in decreasing concentrations of ethanol) and hydrolysed in hydrogen

chloride (HCL), before staining by Feulgen reaction or H&E. The tissue is subsequently moved onto a microscope glass slide with a solution of acetic acid, and teased apart under a dissecting microscope. Eventually, a coverslip is placed on the slide to keep the isolated crypts in place. This method is not only slow but also prone to selection bias, since individual crypts need to be manually identified and excised [6].

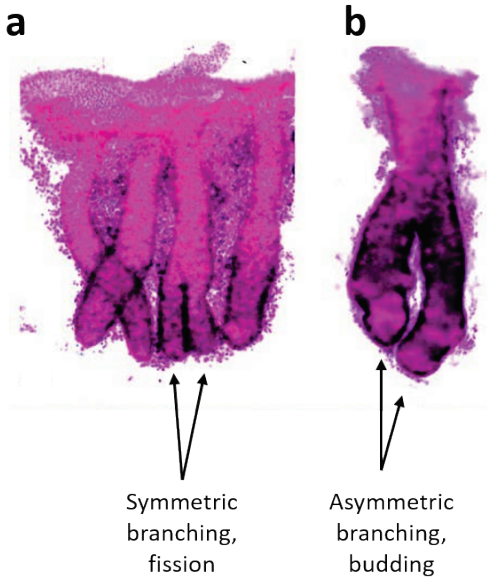


Figure 1.5: Microdissected crypt

Human colonic crypts from excised colons, microdissected and stained with H&E. (a) Healthy crypts, one of which is undergoing the normal process of fission or symmetric branching. (b) Colonic crypt from an adenoma from a patient with familial adenomatous polyposis. The crypt is undergoing asymmetric branching or budding. Images a and b adapted from [15].

1.3.2.3 Mucosal digestion

To collect all crypts within a sample at once, the crypts may be removed from their surrounding tissue following a digestion process [6, 22, 118-120]. The digestion process begins with a freshly harvested colon that is cut open and the mucosal tissue around the crypts is dissolved using a solution of dithiothreitol (DTT) and ethylenediaminetetraacetic acid (EDTA). The suspension is then shaken vigorously to separate the crypts from the sub-mucosa, before it is eventually centrifuged to finalise the separation of the crypts from the surrounding tissue. After isolation, the crypts are fixed in paraformaldehyde (PFA) and stained (e.g. with 4', 6-Diamidino-2-Phenylindole (DAPI) adenine/thymine immunofluorescence), before they are examined under the light microscope (see Figure 1.6) or confocal microscope, as discussed next. The digestion process enables the study of several thousands of crypts at once and it is much quicker than manual

microdissection. It is however highly invasive as it involves repeated shaking of the crypts and centrifuging. It also takes the crypts out of the greater context of their surrounding tissue, by removing them from their native position and orientation in space. Therefore, although protein concentrations along the crypt may be imaged and analysed precisely, 3D morphological information is not available.

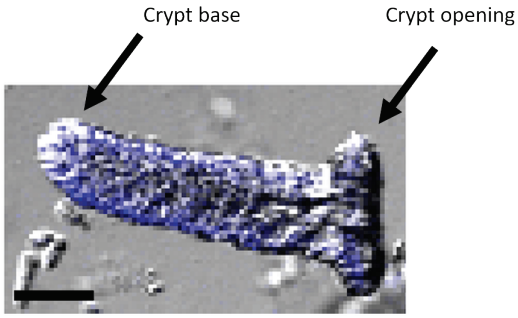


Figure 1.6: Colonic crypt isolated via a digestion process

The mucosal tissue around the crypts in a murine colon is digested to isolate the colonic crypts. The suspension is centrifuged to separate the crypts from the sub-mucosa, and the crypts are fixed and stained with 4', 6-Diamidino-2-Phenylindole (DAPI). The image shows a representative crypt examined under a light microscope, with DAPI staining, yielding a light blue colour. The scale bar is 50 µm. Image adapted from [6].

1.3.2.4 Serial paraffin sectioning

In contrast to the mucosal digestion method, serial paraffin sectioning provides both 3D structural and functional information. Serial paraffin sectioning [9, 121, 122] entails paraffin embedding, sectioning and staining following standard histology procedures described earlier, with the difference that sections are taken along the crypt in a sequence (serial physical sections). During the 3D serial section reconstruction process that follows, the sections are stacked onto each other such that matching features on sequential sections are aligned (image registration). This results in a high-resolution 3D image stack, representing the crypt (Figure 1.7). The 3D reconstruction depicts the crypt down to the cellular level, and allows for histological/immunohistochemical staining of the cross-sections to study the concentrations of particular cell-types and proteins. Serial paraffin sectioning is very labour-intensive, due to the tissue sectioning and the image registration processes, which allows for the study of a few tens of crypts only [9, 121, 122]. Moreover, the natural shape of the colon is altered as it is typically cut open longitudinally, before fixation, embedding and cutting. Finally, paraffin cutting may introduce deformation (cutting

artefacts such as streaks and cracks) [123]. This results in compromised or missing serial sections, reducing the overall quality of the 3D reconstruction of the crypt.

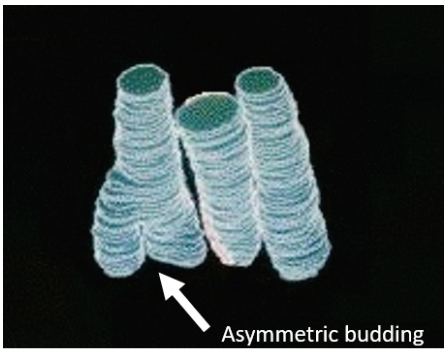


Figure 1.7: Serial sectioning and 3D reconstruction of colonic crypts

Serial paraffin sections of colonic crypts identified within an adenoma of a patient with familial adenomatous polyposis. The paraffin sections were stained, stacked and registered, creating a high-resolution 3D image stack that represents the crypts. The leftmost crypt demonstrates asymmetric budding. Image adapted from [9].

1.3.2.5 Scanning electron microscopy

Accurate reconstruction of isolated colonic crypts can be achieved via SEM [20, 100, 101], such as those shown in Figure 1.8. SEM imaging of colonic crypts involves [101] microdissection and fixation of isolated crypts before they are dehydrated using a critical point dryer to preserve their structure. The samples are subsequently coated with gold particles using an ion coater to enhance their electrical conductivity. Imaging is performed under low vacuum conditions where electrons are propelled towards the sample, and an image is formed by collecting the portion of electrons that bounce off the sample's surface. SEM produces images of superb image quality due to the exceedingly high image contrast. Notably, the technique yields images that appear to be 3D representations. However, SEM provides a single (2D) image, i.e. not a 3D image stack.

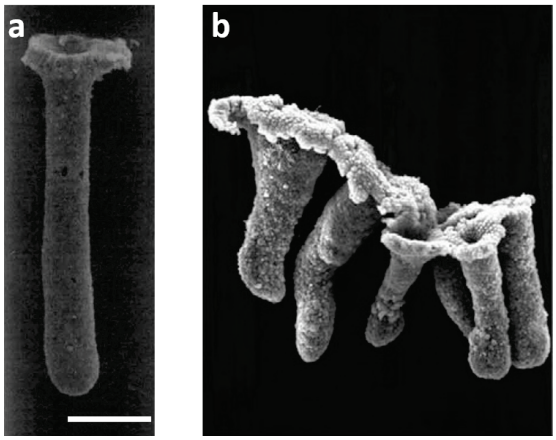


Figure 1.8: Scanning electron microscopy of intestinal crypts

a) A normal human crypt. Scale bar is 100 μm. Image adapted from [20]. (b) Rat colonic crypts within an ACF. The two leftmost crypts are larger (aberrant) compared to the four crypts on the right. Image adapted from [101].

1.3.2.6 Confocal microscopy

A recently developed method that enables tracking of individual cell behaviour and facilitates the understanding of crypt biology is that of *in vitro* crypt organoids [6, 22, 90, 124, 125]. It involves the generation of cell cultures from a single stem cell or isolated crypts that eventually grow into an epithelial layer that surrounds a lumen [27]. These crypt replicas mimic the behaviour of crypts in a cell-to-cell signalling level and bud formation such as that observed during fission. Crypt organoids can be imaged using a standard light microscope [125] or using a confocal microscope, providing 3D information. Confocal microscopy [6, 22, 125] involves staining with antibodies that target specific antigens, to reveal for instance adenine/thymine-rich regions, β-catenin and E-cadherin proteins, eventually providing 3D functional information down to the cellular level. In brief [6], to prepare for confocal imaging, the samples (crypt organoids or microdissected crypts) are placed in a medium permeable to light (e.g. agarose mixture) and mounted on a glass slide before they are placed under the confocal microscope. Laser light at specific wavelengths (e.g. 405 nm, 488 nm or 546 nm) is then used to illuminate the semi-transparent sample. Each wavelength excites a different antibody, yielding a 3D image stack per wavelength, or in fact a single image stack containing 3 colour channels, one for each antibody. The image stack forms a 3D representation of the relevant tissue components (cells containing proteins revealed by each antibody) that can be turned into a photorealistic rendering such as the one shown in Figure 1.9. Confocal imaging yields 3D image stacks without physically cutting the organoids, at exceedingly high spatial resolutions (virtual slice thickness at 0.33 μm) and fully 3D functional mapping (protein expression in each individual cell of the crypt).

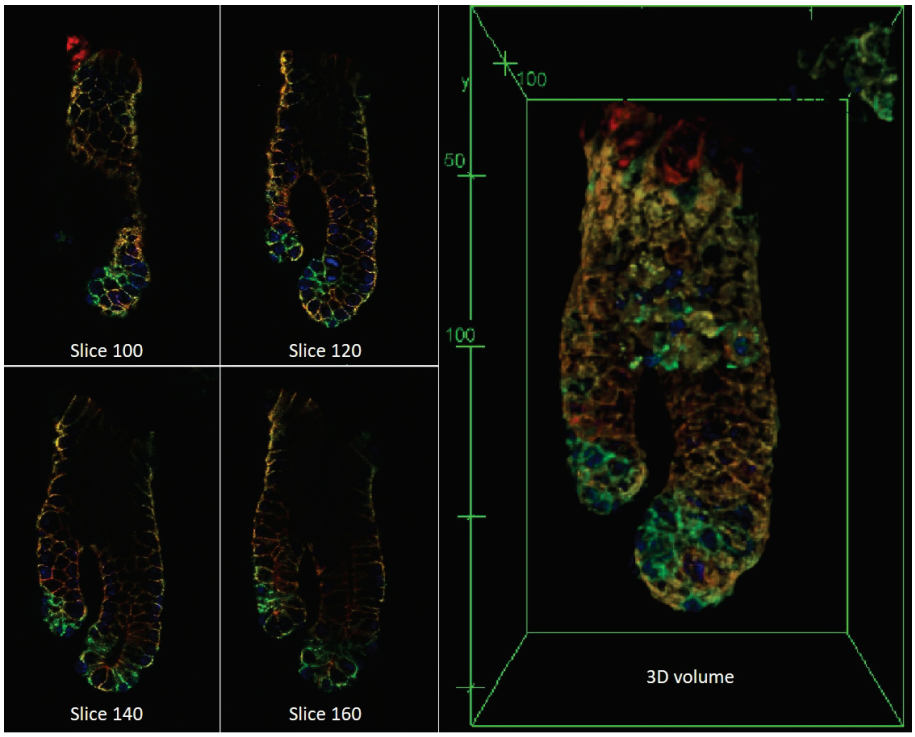


Figure 1.9: Confocal imaging of intestinal crypts

Isolated murine colonic crypts undergoing budding are stained with antibodies specific to 4', 6-Diamidino-2-Phenylindole (DAPI), β-catenin and E-cadherin and placed in a medium that is permeable to light (e.g. agarose mixture). Confocal imaging yields a 3D image stack containing 3 colour channels (red: DAPI, green: β-catenin, blue: E-cadherin) made up of 267 images with a thickness of 0.33 µm. When put together, the images compose a 3D image stack with functional information throughout the crypt. Image adapted from Tan *et al.* [6].

1.3.2.7 Endoscopic methods

Although cancer diagnosis falls outside the scope of this thesis, a few *in vivo* endoscopic methods are worth mentioning as experimentalists have employed them to quantify the colonic crypt structure in 3D. High-magnification chromoscopic colonoscopy (HMCC) is typically used to identify ACF on the colonic surface. HMCC is performed using an endoscope with a high-magnification (× 40) camera attached [126], and involves spraying the lumen with methylene blue (most frequently) or indigo carmine to reveal the ACF out of the surrounding healthy crypt openings [17], as shown in Figure 1.10. Endoscopic confocal microscopy may then be used to visualise deeper tissues (a few hundred µm tissue depth) [127, 128]. To visualise the entire crypt (~ 1 mm) in 3D, *in vivo,* Qi *et al.* [129] used optical coherence tomography (OCT) instead. The method produces high-resolution real-time imaging of the intestinal mucosa down to a tissue depth of about 1.5 mm, which is appropriate to visualise the colonic crypts and quantify their 3D structure.

Notably, OCT [129-134] is emerging as a viable diagnostic tool for sub-surface colorectal cancer [133]. In rodents, intravital microscopy [11, 135] is an alternative approach for *in vivo* imaging. It requires the implantation of an abdominal imaging window [136] into living mice to obtain visual access to the intestine. The mouse is then placed on top of an inverted confocal, or multi-photon microscope, which enables imaging of individual live cells in the small intestinal crypt base. Such methods have been used to identify individual cells of different types, to validate predictions of computational models (see section '1.3.4 Computational models of the intestine').

1.3.2.8 Standard methods summary and outlook

To date, morphological quantification of budding crypts in 3D remains a challenge. Importantly, 3D quantification of the budding crypt structure in combination with the capability of performing functional assays in 3D is of particular interest. Yet, current methods are not suited for this purpose, especially for large number of crypts, as evident in Table 1.1. Therefore, new methods need to be developed for 3D imaging of large numbers of crypts, *in situ*, which should be capable of providing concurrent structural and functional information. The particular requirements for a such a method, especially suited for crypt budding are revealed in the following section, where specific studies focusing on the budding crypt morphology are reviewed.

Table 1.1 Comparison of standard methods for the study of crypt budding

The key the key competencies of prominent methods in literature with regards to crypt budding. The symbols tick (✓) and cross (✗) stand for *yes* and *no*, respectively. The numbers of crypts shown here are only indicative to provide a rough comparison of the capabilities between the methods.

Method	*In situ*	3D	Functional information	Number of crypts
Standard histology	✓	✗	✓	~ 100
Microdissection	✗	✗	✓	10 - 50
Mucosal digestion	✗	✗	✓	> 1000
Serial paraffin sectioning	✓	✓	✓	~ 100
SEM	✗	✗	✗	10 - 50
Confocal microscopy	✗	✓	✓	10 - 50
Endoscopic methods	✓	✓	In specific cases	~ 100

1.3.3 Morphological quantification of crypt budding

The majority of human and animal studies on the development of CRC and early microstructural changes in the colon have not focused on morphological quantification of budding crypts. Instead, objective quantification of crypt budding has been confined in differentiating (normal) crypt fission and asymmetric budding as a consequence of a specific treatment in animal models or underlying disease in human tissue. This section examines the several quantitative measures used in studies involving crypt budding, which are summarised in Table 1.2.

The crypt fission index (FI or CFI) is defined as the percentage of crypts in fission, or the ratio of the number of crypts in fission divided by the total number of crypts [6, 9, 19, 22, 62, 85]. This global measure (it refers to a collection of crypts) does not differentiate between normal and asymmetric budding but it is typically used to identify increased fission rates (fissioning as a function of time) as a consequence of treatment in animals or underlying condition in humans.

To quantify the budding crypt's shape, two classes of measures can be identified: those that characterise the shape of the crypt openings on the mucosal surface and those that characterise the shape of the bulk of the crypt.

As ACF are reportedly the first steps in cancer development [31, 126, 137] and tightly related to crypt budding, the methods for ACF identification are naturally relevant to the morphological quantification of budding crypts. In humans, ACF observed on the colonic lumen's surface are identified *in vivo*, through HMCC (Figure 1.10). Gupta *et al.* [138] summarised the endoscopic criteria used by several studies for the identification of ACF, with most endoscopic studies identifying ACF as darker staining clusters of crypts with one or more of the following characteristics: larger mean diameter (or size) of the crypts in the cluster compared to those of the surrounding mucosa [13, 126, 139], thicker epithelial lining (thickness of the crypt epithelium or epithelial thickness) [126, 137, 140-142], dilated crypt lumen [143, 144] and/or slightly raised openings (opening height ~ 2 mm) from the surrounding mucosa [139, 143-145]. Some studies attempt to better quantify the shape of the crypt lumen by including additional characteristics such as compressed, non-distinct [126], slit-like and elongated [31, 143] or serrated/star-shaped [18, 126, 146, 147] lumen. There is thus a lack of general agreement on a consistent set of criteria used to identify and define ACF on endoscopic images [138]. Specifically, morphological criteria (epithelial thickness, crypt lumen's diameter, elongated or serrated lumen) are often loosely defined and the methods for their quantification are often unclear and/or only qualitative.

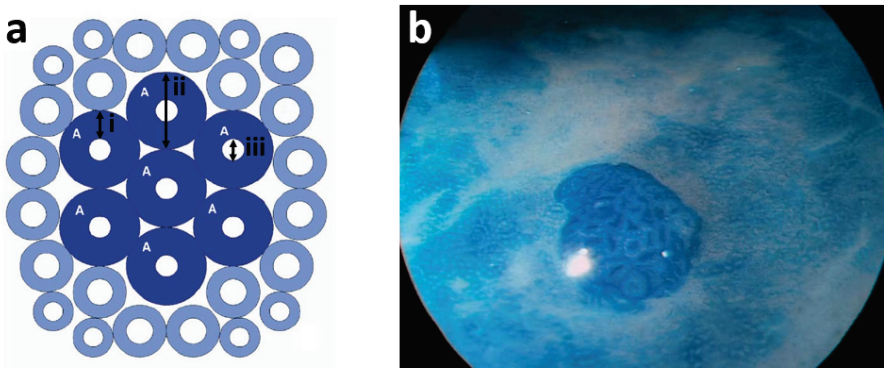


Figure 1.10: Endoscopic identification of aberrant crypt foci

(a) Schematic diagram of the endoscopic view of a cluster of crypts representing an ACF marked with 'A'. The smaller rings represent normal surrounding crypts. The characteristics that have been used by several studies in the past to define an ACF are as follows: darker staining than the surrounding normal crypts, thicker epithelial lining (i), larger crypt size (ii), and compressed lumen (iii). (B) Human ACF as seen via HMCC. Images a and b adapted from [31].

Once identified through HMCC [138], tissue biopsies are taken through the endoscope and processed following standard histology workflows (see section '1.3.2 Standard histological methods for the study of crypt budding') to produce, usually H&E, histology slides. These are used for histologic validation of the existence of intraepithelial neoplasia or pre-invasive neoplastic change of the epithelium (i.e. formation of abnormal tissue growth that has not yet penetrated other tissues)[1]. It is therefore straightforward to attempt quantification of the crypt lumen's shapes on histology slides. An interesting approach in this direction has recently been taken by Lee and colleagues [148] who have defined shape-specific measures to quantify H&E histology sections of crypts in tubular adenomas. Their measures estimate how much a planar polygon (trace around the crypt lumen cross-sections on arbitrarily oriented histology slides) differs from a circle. They employed multivariate discriminant analysis, a statistical method used to classify groups of measurements, employing their three measures to classify the cross-sections in three groups: non-branching, mild branching and moderate branching (branching in this context refers to the cross-sectional shape of the crypt lumen, not the 3D morphology of the entire crypt). Although Lee and colleagues' method does not extrapolate their findings to the 3D shape of the crypts, it shows that shape-based measures can successfully assess and classify

[1] As stated in the World Health Organisation (WHO) guideline that defines the nomenclature for the human case [48], the term intraepithelial neoplasia was introduced in an attempt to resolve confusion surrounding the terms 'dysplasia', 'carcinoma in situ' and 'atypia'.

irregular branching structures (in 2D). A 3D approach would imminently pose the question about the orientation at which the histology sections were taken, since the shape of 2D cross-sections of a 3D structure is dependent on the orientation of the cross-sectional plane [149, 150].

ACF that closely mirror the pattern seen in humans and depict similar histopathological features can be induced in animal models [91]. In mice (and rats) ACF are identified on excised and chemically fixed colons, with the use of methylene blue and a dissecting microscope [38, 99]. The colons are typically cut open and flattened before fixation, resulting in well-oriented crypts that are aligned and directly comparable [62, 100, 137, 151]. Wargovich *et al.* summarised the criteria used across several rodent studies for the identification of ACF [17]: Similarly to the human case, shapes of the crypt openings on the mucosal surface are altered. Quantitative measures include the crypt diameter and the epithelial thickness, while qualitative characteristics include the observation of elongated or serrated lumens.

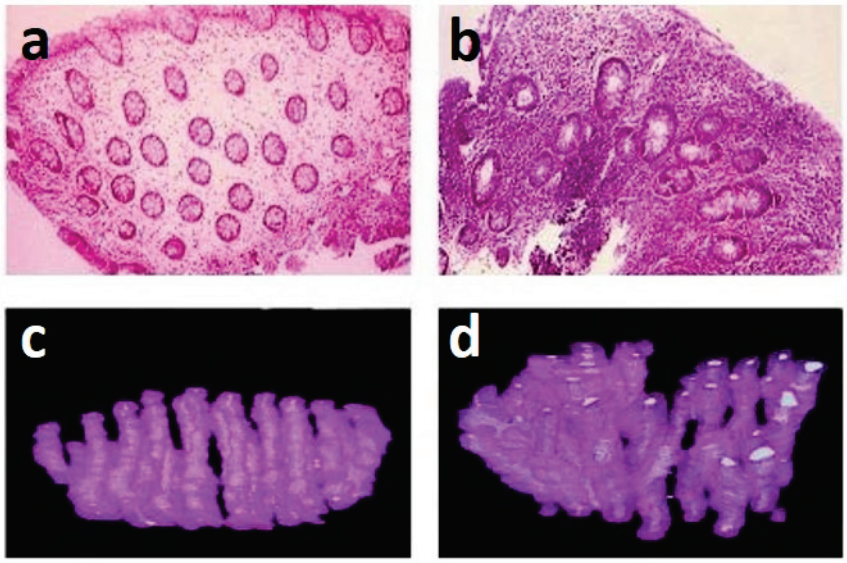


Figure 1.11: 3D reconstruction of human colonic crypts via serial paraffin sectioning

(a & b) H&E histology slides of single sections from biopsies of patients with ulcerative colitis. Original magnification × 100. (c & d) 3D reconstruction of the crypts shown in a and b, respectively. Images a-d adapted from [121].

Notably, crypts are 3D structures, and morphological alterations during the early stages of CRC are not confined to their openings on the mucosal surface (morphological changes observed on ACF), but extend further to the bulk of the crypts. An elegant approach that studied both the crypt openings on the mucosal surface and the shape of the bulk of the crypts in human colonic biopsies was taken by Qi *et al.* [129]. Superficial images were taken with a commercial optical camera from freshly resected tissue biopsies that were stained with methylene blue. 3D image

volumes were acquired using OCT on the same sites and registered with the 2D images. The authors used the area, the major/minor axis lengths and the eccentricity of the best-fitting ellipse, and the solidity (ratio between the area of the crypt lumen and the area of its convex hull, indicating the extent to which a shape is convex or concave) to quantify the 2D appearance of the crypt openings in the superficial 2D images. The authors did not study budding but they focused on the orientation of neighbouring crypts. For 3D quantification of the crypt orientation, they extracted the skeleton of the crypts, i.e. the curve representing the centre line of the crypt lumen. They then used the crypt skeletons to quantify the crypt straightness by fitting a straight line and computing the coefficient of determination, which indicates the extent to which the skeleton resembles a straight line. They subsequently quantified the direction of the fitted lines for each crypt skeleton and obtained a value representing the crypt parallelism, i.e. the extent to which the crypts are parallel to each other.

Furukawa *et al.* [121] on the other hand took a more conventional approach to quantify the 3D shape of abnormal crypts, by applying serial sectioning and H&E staining before reconstructing the crypt's 3D shape (Figure 1.11). The authors studied morphological abnormalities in patients with ulcerative colitis by quantifying the crypt lumen volume, the crypt lumen surface area, the volume of the crypt epithelium and the external surface area of the crypt epithelium.

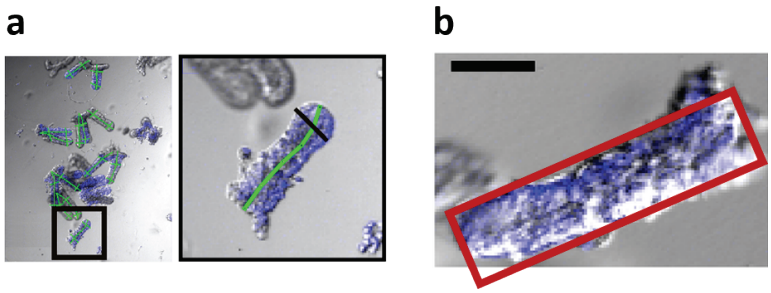


Figure 1.12: Morphological quantification of digested crypts

Crypts isolated via mucosal digestion and suspended in medium. (a) Confocal image of crypts being characterised by Tan *et al.,* with a green segmented line marking the length of the crypt and a black line marking the width of the crypt base (drawn at about 20 % of the crypt length from the crypt base). (b) Tan *et al.* estimated the crypt size as the rectangular area of the crypt (area of the red rectangle) on the confocal image. Scale bar is 50 µm. Images a and b adapted from Tan *et al.* [6].

Mucosal digestion is a popular technique to study crypt budding (see section '1.3.2.3 Mucosal digestion'). Figure 1.12 shows light microscope images of crypts after isolation, yielding a 2D projection of their 3D structure. The crypt width [6], length [6, 22] and area [6, 19, 62] may then be estimated from these 2D images. Note however that the naming convention widely used for

the crypt width and length in this case is misleading, since these morphological characteristics are altered due to the sample preparation method (picking the soft, pliant tissue on a flat surface forces it to change shape). The crypt area is estimated as the area occupied by the crypt on the 2D projection (Figure 1.12), which is also inaccurate and it should not be confused with the actual 3D area of the crypt epithelium.

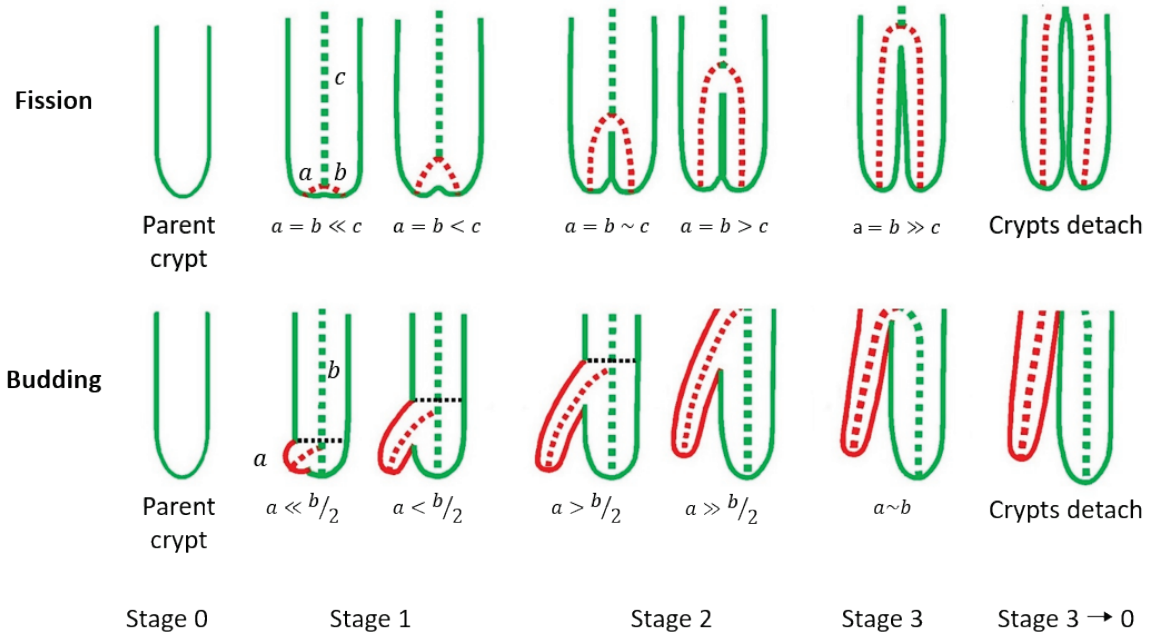


Figure 1.13: Crypt budding stages

The asymmetric crypt production concept is proposed by Tan *et al*. In the schematic representations red shows the new bud, green shows original crypt and the black dotted line shows the developing front. Stages are defined quantitatively by measuring the lengths of the respective crypt sections. Stage 0: non-budding; Stage 1 to 3: onset mid and late budding stage, respectively. Stage 2 is defined when a daughter crypt appears at the lower half of the parent crypt and while it is still smaller than the parent crypt. Stage 3 refers to a finished budding event when the point of division is at the upper half of the parent crypt and the two crypts are about to be separated. Image adapted from Tan *et al*. [6].

Tan *et al.* [6] applied a structured approach for the quantification of crypt budding as follows: They defined the stages that a crypt undergoes during the symmetric (fissioning) and asymmetric budding process (Figure 1.13): A non-budding crypt is at stage 0 until a crypt bud (point of division) near the bottom of the crypt appears (stage 1). Stage 2 is defined when a daughter crypt appears at the lower half of the parent crypt, and while it is still smaller than the parent crypt. Finally, stage 3 refers to a finished budding event when the point of division is at the upper half of the parent crypt, and the two crypts are about to be separated. Notably, this staging system

assumes an individual crypt undergoing a bifurcation process (branching into two crypts), and it is thus not directly applicable to crypts with more than two branches. Based on this model, the authors defined the length of each of the two (parent and daughter) crypts as the lengths A and B, measured from the top of the bifurcating crypt to the base of the parent and daughter crypts, respectively (Figure 1.14). This enabled them to compute the mean crypt length C as the average of the parent and daughter crypt lengths $C = \frac{A+B}{2}$, and define the asymmetry index (AI) as the mean relative deviation from mean crypt length: $AI := \Delta AB = mean(\Delta A, \Delta B)$ where $\Delta A = \frac{|A-C|}{C}$ and $\Delta B = \frac{|A-B|}{B}$. The asymmetry index quantifies the amount of asymmetry in the bifurcating crypt and enables the classification between symmetric (fission) and asymmetric budding by simple thresholding (budding events with $AI > 1\ \%$ were classified as asymmetric).

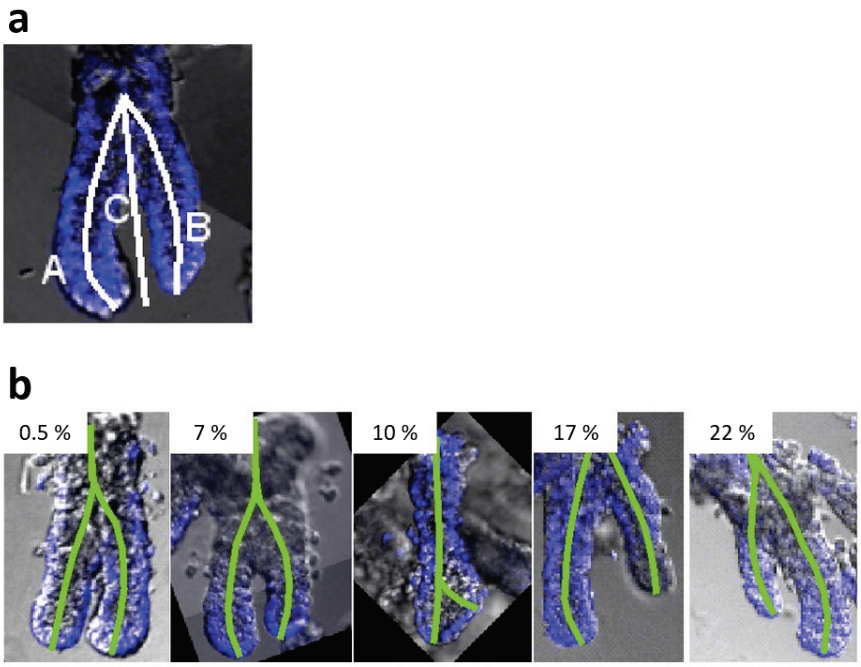


Figure 1.14: Budding crypt length and asymmetry index (AI)

(a) Tan *et al.* defined the length of each of the two (parent and daughter) crypts as the lengths A and B measured from the top of the bifurcating crypt to the base of the parent and daughter crypts, respectively. They defined the mean crypt length C as the average of the parent and daughter crypt lengths $C = \frac{A+B}{2}$, and the asymmetry index (AI) as the mean relative deviation from the mean crypt length: $AI := mean(\Delta A, \Delta B)$, where $\Delta A = \frac{|A-C|}{C}$ and $\Delta B = \frac{|A-B|}{B}$. (b) Examples of budding crypts with increasing per-cent AI labelled above the confocal images. Images a and b adapted from Tan *et al.* [6].

Table 1.2: Summary of crypt morphological measures in the literature

Measure	Studies
Global crypt budding measures	
Crypt Fission Index (FI, CFI)	[6, 9, 19, 22, 62, 85]
Crypt opening shape measures	
Area of ellipse (fitted to the crypt lumen)	[129]
Eccentricity of ellipse (fitted to the crypt lumen)	[129]
Major & minor axis lengths of ellipse (fitted to the crypt lumen)	[129]
Solidity	[129, 148]
Diameter (or size) including epithelial thickness	[13, 126, 139]
Epithelial thickness	[126, 137, 140-142]
Crypt lumen size	[143, 144]
Opening height (raised above the surrounding mucosa)	[139, 143-145]
3D measures	
Straightness	[129]
Parallelism	[129]
Crypt volume	[121]
Crypt epithelial area	[121]
Lumen volume	[121]
Lumen surface area	[121]
Measures on 2D projections of microdissected crypts	
Crypt length	[6, 22]
Crypt width	[6]
Crypt area	[6, 19, 62]
Topological measures of crypt budding	
Mean crypt length (average of parent and daughter crypt lengths)	[6]
Asymmetry index (AI) (deviation from mean crypt length)	[6]
Symmetry ratio (SR) (length of short daughter/length of long daughter)	[22]
Highest mitotic figure (HMF)	[9]
Mitoses per crypt	[9]

Quantification of budding asymmetry is however not consistent across studies, where alternative definitions have been used. For instance, Langlands *et al.* [22] defined/quantified the symmetry ratio (SR) of a crypt in fission by dividing the length of the shorter daughter crypt by the length of the longer daughter crypt. Wasan and colleagues [9] on the other hand, divided each crypt into ten equal compartments (base=0.1 to top=1). They then assessed the highest mitotic figure as the uppermost compartment where mitosis occurred. They defined asymmetric fission events as those commencing on the lateral (upper 80 %) crypt wall, i.e. highest mitotic figure greater or equal to 0.2. The authors also measured the number of mitoses per crypt (number of branches) since their study involved budding crypts with more than two branches.

Taken together, the studies summarised in this section reveal an incompetence of the standard methods to support morphological studies of crypt budding. Indicatively, 3D quantification was restricted in basic measures of the crypt shape such as volume and epithelial area. On the other hand, first steps in morphological quantification of crypt budding (the concept of symmetric vs asymmetric budding) have been confined to 2D quantification.

However, as mentioned earlier, experimentalists have not focused on morphological quantification of crypt budding. This is partly due to the fact that hypotheses that link the morphological characteristics of crypt budding with functional information, are only recently emerging (e.g. the concept of differential stiffness induced by different cell types [22, 124] discussed next). Such hypotheses are typically put forward through computational modelling, as predictions of the models that are subsequently tested experimentally. Specifically, spatial computational models attempt to explain the crypt morphology as a consequence of the underlying biological processes. Thus, spatial computational models are an invaluable tool in furthering the understanding of the earliest morphological changes in the colon during CRC development and their cause. As such, computational models that are relevant to the study of crypt budding are reviewed in the following section.

1.3.4 Computational models of the intestine

Mathematical models of the crypt can be collected in two categories, the spatial models [27] that model the crypt's morphological evolution, and the non-spatial models [152] that focus on the dynamics of crypt production and proliferation. Spatial computational models play a pivotal role in furthering the understanding of the biological processes that control crypt budding. Yet, as with every model or theoretical description of a process, experimental validation is essential to verify its accuracy. This section aims to identify the requirements of experimental procedures to provide

useful 3D morphological data for validation and further development of mathematical models of crypt budding.

Early attempts to understand the process of crypt fission have been made by Totafurno, Bjerknes *et al.* [10, 25, 26] who proposed the concept of the crypt cycle: A crypt gradually grows and undergoes fission when it reaches a critical size (volume), giving birth to two daughter crypts, and then the process starts over. Their models were based upon the observation that crypts undergoing the first steps of fission had already reached a larger volume, suggesting that crypt size was driving fission. Totafurno *et al.* [10] used the crypt fission index, the crypt volume and the branch orientations (angle between parent and daughter crypt) to formulate their crypt cycle concept. In a theoretical study of crypt dynamics, M. Bjerknes [25] developed a model that predicted the percentage of mutant crypts within aberrant crypt clusters (50 % of ACF should be monocryptal, 17 % should be composed of two crypts, and 33 % should be clusters of more than two crypts). It was then underlined that future experimental studies had to specify the exact numbers of each cluster type to properly validate the model outcomes. Taken together, the analysis of these models reveals the importance of experimental procedures that enable researchers to collectively examine the colonic tissue by looking at the entire mucosal structure at once, as opposed to focusing on a small number of isolated crypts. It is also evident in the work of Totafurno *et al.* [10] that such procedures require the spatial orientation of individual crypts to be preserved in order to derive information such as the branch angles. However, as discussed in the previous section, existing procedures either preserve the shape and orientation of individual crypts or enable the study of large number of crypts by mucosal digestion, but not both.

Spatial computational models can be divided into discrete, which model each crypt as a collection of cells, and continuum, which consider the crypt as a smooth tissue. Edwards and Chapman [153] developed a continuum model for crypt budding to model crypt shapes found in adenomas by Preston *et al.* [15]. Their model considers the epithelium as a beam, which is caused to buckle by forces exerted by its substrate (neighbouring non-epithelial cells). The Edwards and Chapman model suggests that crypt buckling, and consequently crypt budding, depends, among other factors, on the crypt size and the tissue stiffness. A continuum model approach that integrates morphological, biomechanical and functional information was applied by Pin *et al.* [90, 154] who employed organoids to study crypt fission *in vitro*. The authors used fluorescent immunostaining and confocal imaging to identify and differentiate cell types that were assumed to be stem and Paneth cells, combined with atomic force microscopy (AFM) to estimate the two cell-types' biomechanical stiffness. Their model predicted that fission occurs due to stiffness differentials introduced by neighbouring concentrations of stem and Paneth cells in the small intestine (or stem and $cKit^+$ cells in the colon), which was later experimentally confirmed and reproduced by

subsequent studies [22, 124], and computational models [23]. This notion challenged the traditional belief that crypt fission was initiated after a crypt reached a critical size (crypt cycle put forward by Totafurno *et al.*) as proposed by earlier models [10, 25, 26, 153]. The authors argue that crypt buckling is generally unrelated to absolute dimensions of the crypt compartments such as those proposed by the Edwards and Chapman model [153], and they highlighted the effect of the stiffness differentials to the initiation of crypt fission. Importantly, they proposed that the daughter crypt was formed at the base of the parent crypt and progressively migrated upwards without affecting cell proliferation and migration within the parent crypt. In other words, that the daughter crypt grows as an independent, non-parasitic entity. Pin *et al.*'s model does not consider any stimuli from signalling pathways [90], allowing the crypt to grow unrestricted. It made however a great leap forward in understanding the important role that biomechanics play in crypt budding in 3D. Their modelling of biomechanical stress in 3D highlights the necessity of appropriate 3D data (imaging and biomechanical modelling) to support and validate such newly emerging and complex models. Importantly, the experimental observations reported in organoids remain to be made *in situ* as well, and the models remain to be extended in the 3D case of the crypts within the mucosa.

3D computational models have been developed to study the effect of signalling pathways on cell-type distribution, cell migration throughout the crypt, mutation and clonality [155-162]. However, these approaches adopt an oversimplified (cylindrical) representation of the crypt structure that is notably perfectly symmetric. Thus, the effect of the 3D morphology of the crypt to cell migration has not been studied to date. With biomechanics becoming the focus of recent studies, the relationship between the 3D crypt shape and the developed biomechanical stress on the crypt epithelium is likely to become the centre of interest in the future. A pre-requisite to facilitate such studies being the availability of appropriate experimental methods. Interestingly, an underlying theme among the majority of the computational models discussed in this section is either the lack of experimental data or the difficulty to validate the mathematical models due to inability to measure certain parameters of the models [6, 27]. When it comes to crypt budding, models have been confined to the study of the healthy case, i.e. crypt fission.

As indicated by Smallwood *et al.* the development of computational models requires firstly that the models recapitulate biological functions and secondly that robust techniques of validation are available [163]. Therefore there is a need for new methods for the study of crypt budding to not only support the development of mathematical models, but also to shed light to newly emerging concepts that involve morphological quantification of the colonic crypt, specifically the notion of asymmetric budding [6], discussed earlier.

A requirement for concurrently providing functional and structural information, emerges from section '1.3.3 Morphological quantification of crypt budding', whereas a requirement for 3D imaging of a large number of crypts *in situ*, emerges from this section. Therefore, new methods and workflows that need to be developed for the study of crypt budding must fulfil these requirements. X-ray μCT is a 3D imaging technique, which has not been employed so far to image colonic crypts. Being non-destructive, it could be combined with standard histology procedures to develop methods that fulfil the above-mentioned requirements. Thus, X-ray μCT is reviewed next.

1.4 Advanced imaging approaches in research

Application areas of X-ray micro-computed tomography (μCT) have expanded since the conception of the technique in the 1970s [164] to a great variety of fields, including archaeology and palaeontology, biology, biomedical and life sciences [3-5, 165]. It exploits the physical phenomena that take place when X-rays interact with matter, in order to non-destructively reconstruct a 3D representation of the internal structure of the sample [165]. X-ray interaction with matter is a combination of attenuation and phase shift of the X-ray wave, with the attenuation being simpler and easier to exploit, both experimentally and mathematically [165]. As the application areas are profoundly broad, the involved materials vary to a great extent, posing a diverse set of challenges for X-ray imaging. For instance, fossilised samples are generally dense materials, substantially attenuating the X-ray beam [166, 167]. Soft tissue samples (e.g. brain, lung or colon) on the other hand, barely attenuate hard X-rays (in the energy range between 10 and 100 keV) [168, 169]. In this case, X-ray phase shift imparted by the tissue may be exploited for imaging instead of the X-ray attenuation [169-171]. This becomes possible because when increasing the X-ray energy, although attenuation becomes lesser, X-ray phase shift can still be observed as discussed in section '1.4.2 X-ray interaction with matter and image formation'. Technically, X-ray phase-contrast imaging is more difficult to perform than attenuation-based X-ray imaging with variations in complexity of the experimental setup and the recorded signal [169]. These include interferometric methods, diffraction enhanced imaging, crystal analyser and propagation-based methods [172]. The simplest phase-contrast imaging technique is called (free-space or inline) propagation-based phase-contrast imaging. It exploits the X-ray beam propagation from the sample to the plane of observation, to reveal low-absorbing materials and tissues [173]. Inline propagation-based phase-contrast imaging is markedly simpler to implement experimentally than other phase-contrast imaging techniques as no further optical elements, such as diffractive X-ray lenses (Fresnel zone plates), are needed compared to X-ray absorption-based imaging [169].

Being a non-destructive 3D imaging technique suitable for soft tissues, X-ray propagation-based phase-contrast imaging is appropriate for colonic tissue imaging to study the morphology of crypt budding. To this end, in this section an overview of different X-ray imaging techniques is presented aiming to elucidate key concepts and the underlying principles.

1.4.1 Computed tomography principles

Computed tomography is a technique where multiple X-ray radiographies are taken to reconstruct a 3D representation of an object. A µCT imaging system is typically composed of an X-ray source, a detector and a stage to rotate the specimen. In contrast, for medical CT and *in vivo* µCT systems used in pre-clinical research (for small animals such as mice or rats), a gantry system of the X-ray source and detector is rotated around the patient or animal that stays in place [4]. X-rays used for µCT imaging can be generated in several ways, including but not limited to, large electron accelerators (synchrotrons), micro-focus or liquid-metal X-ray sources [174]. X-rays generated in the X-ray source pass through the object or patient/animal and propagate through free space, to arrive at the detector's side, where typically, (i) a scintillator turns them into visible light that is recorded by a photodetector (indirect X-ray detection method) or (ii) the X-ray photons are counted through direct conversion into electrical charges by semiconducting materials (direct X-ray detection method) [165]. The process of combining several recorded radiographies, each obtained at a different specimen or gantry rotation angle, to create a 3D representation of the object, is termed CT reconstruction (see section '1.4.3 Attenuation-contrast X-ray imaging').

In a classic laboratory µCT system, X-rays are usually generated by micro-focus X-ray tubes, which make use of an electric potential to accelerate electrons from a cathode (filament) towards an anode (target). This X-ray tube potential is in the order of kilovolts (kV) and the corresponding electron energy is thus in the order of kiloelectronvolts (keV). When electrons hit the X-ray target, which is usually made of elements/materials with high atomic numbers, such as tungsten (W) or molybdenum (Mo) [175], an X-ray beam is generated via electron interactions with the target material. Figure 1.15 shows that the X-ray spectrum of an X-ray tube is polychromatic or spectrally continuous, which means that it covers a certain energy range (from zero keV up to a maximal energy of the applied tube voltage). The X-ray spectrum is a combination of characteristic radiation that is specific for the target material and the bremsstrahlung component [4] that spans the entire spectrum. The characteristic peaks are generated when electrons from the outer atomic shells fill in vacant space mainly at the K, L and M left after the incident beam displaces electrons. The bremsstrahlung radiation results from the loss of kinetic energy of the incoming electron beam when it is attracted by the target material's nuclei through the nuclear Coulomb

field. The highest energies in the spectrum correspond to the kinetic energy emitted when incident electrons hit the target's nucleus directly. Either the kilovoltage peak (peak voltage in kilovolts) or maximal X-ray energy of X-ray tubes is denoted as kVp or keVp, respectively. Lower energies in the spectrum typically in the soft X-ray regime (0 to 10 keV) spanning the spectrum through the lower hard X-ray regime (a few tens of keV) may be suppressed by filtering (i.e. placing a thin sheet of material, typically aluminium, copper, or rhodium in the beam path before the sample to absorb low energy X-rays) [176].

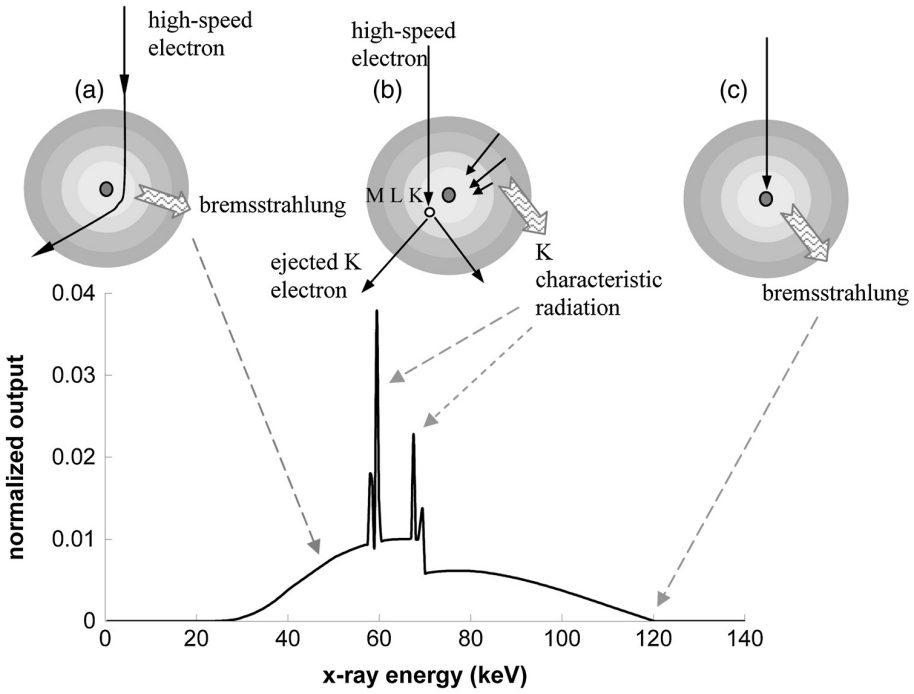


Figure 1.15: The X-ray spectrum of an X-ray tube with a target made of tungsten

X-rays are generated when accelerated electrons interact with the target's material. (a) Bremsstrahlung radiation is a result of the change of electron's speed (loss of kinetic energy) when they are attracted by the target's nuclei through the nuclear Coulomb field. (b) Electrons from the K, L and M shells are extracted from the target's atoms when hit by incident electrons. Then, free electrons from the outer shells fill in the vacant space emitting characteristic radiation. This produces peaks in the X-ray tube's spectrum at energies characteristic to the target's material. (c) The highest energies in the spectrum correspond to the kinetic energy emitted when incident electrons hit the target's nucleus directly. The spectrum of an X-ray tube with a target made of tungsten and an electron accelerating voltage of 120 kV, ranges from 0 keV to 120 keV (additional filtering was used to remove the energies below 30 keV). Image adapted from [4].

In a synchrotron radiation X-ray source shown in Figure 1.16, an electron beam is initially generated by an electron gun (e-gun) and accelerated to high energies (GeV levels) initially through a linear accelerator (LINAC) device and subsequently through a booster ring. This energy is maintained by circulating the electron beam in a 'storage ring', where X-rays are generated in so-called insertion devices (bending magnets, wigglers and undulators) when the electrons are forced to change their direction. In contrast to X-rays generated in conventional X-ray tubes, the resulting polychromatic X-ray beam spectrum emitted by bending magnets and wigglers is typically filtered by gratings and/or crystal monochromators located in the 'optics hutch' before they reach the 'experimental hutch'. This confines the spectrum to a very narrow energy band, providing a monochromatic X-ray beam characterised by a single X-ray energy, while undulators provide a quasi-monochromatic X-ray beam. The X-ray beam generated by synchrotron sources provides exceedingly higher spatial coherence and X-ray flux than conventional X-ray tubes in lab-based X-ray sources [165], allowing significantly shorter scan durations for a desired spatial resolution, generally a better image quality, and monochromatic X-rays.

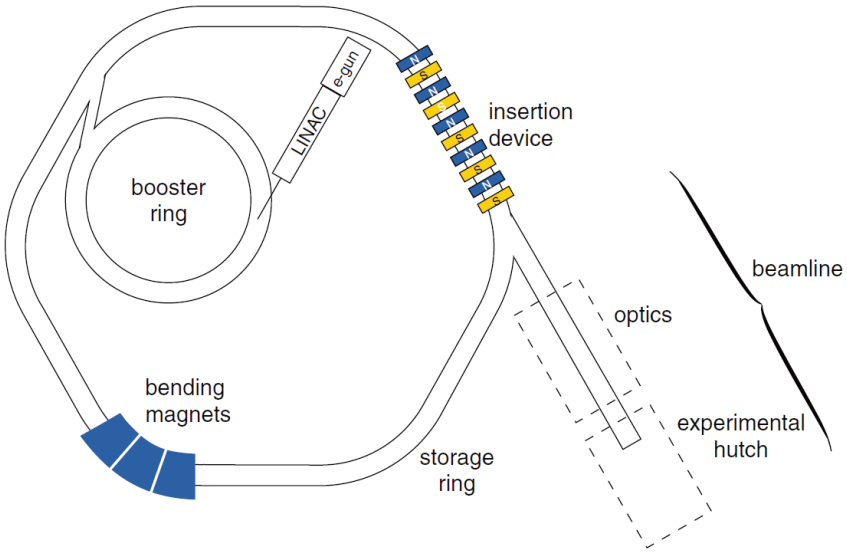


Figure 1.16: The architecture of a synchrotron source

In a synchrotron, an electron beam is generated by an electron gun (e-gun) and accelerated to high energies (GeV levels) initially through a linear accelerator (LINAC) and subsequently through a booster ring. This energy is maintained by circulating the electron beam in a 'storage ring', where X-rays are generated in so-called insertion devices (bending magnets, wigglers and undulators) when the electrons are forced to change their direction. The resulting polychromatic X-ray beam spectrum is typically filtered by gratings and/or crystal monochromators located in the 'optics hutch' before they reach the 'experimental hutch'. Image adapted from [165].

1.4.2 X-ray interaction with matter and image formation

X-ray refraction describes the effects taking place when X-rays pass from a medium through another (e.g. vacuum and sample made of some material) [169]. In the time domain, the electric field of a monochromatic or temporally coherent plane wave in vacuum moving forward in the *z*-direction (Figure 1.17) and entering an object made of a homogenous material is given by [165]:

$$E(z,t) = E_0 e^{i\left(\frac{2\pi}{\lambda}nz - \omega t\right)}, \quad (1.1)$$

where E is the electric field, z the distance from the entrance point (the object is assumed to be located at $z = 0$), otherwise referred to as X-ray propagation distance, t the time since the wave has left the source, E_0 the electric field at the source, ω the angular frequency of the wave, λ is the wavelength and n the complex refractive index:

$$n = 1 - \delta + i\beta, \quad (1.2)$$

where δ is called the decrement of the real part of the refractive index ($1 - \delta$) and β is the imaginary part of the refractive index. Replacing Equation (1.2) in (1.1) yields:

$$E(z,t) = e^{-\frac{2\pi\beta}{\lambda}z} E_0 e^{i\left(\frac{2\pi}{\lambda}(1-\delta)z - \omega t\right)}. \quad (1.3)$$

Therefore, the imaginary part of the refractive index β exponentially attenuates the electric field, while δ affects the phase, as schematically shown in Figure 1.17. In a non-homogenous material, both δ and β vary spatially, and they are directly related to the material properties of the object that is investigated. Notably, the refractive index of an object's material refers to the case where the wave moves from the vacuum into the object, where tabulated values are available [177]. A convenient tool to retrieve a material's refractive index given its composition is made available online by the *Center for X-Ray optics (CXRO)*, Berkeley, USA [178]. It is important to underline that in the more general case, the X-ray wave passes from one material to another, i.e. one of the two mediums is not necessarily vacuum. Thus, when the object is composed of several distinct materials, refraction would take place at the boundary between each pair of materials. In this case, the real part decrement δ and the imaginary part β of the refractive index for the material pair would be equal to the difference between the corresponding real part decrements and imaginary parts of the refractive indices for the individual materials [179].

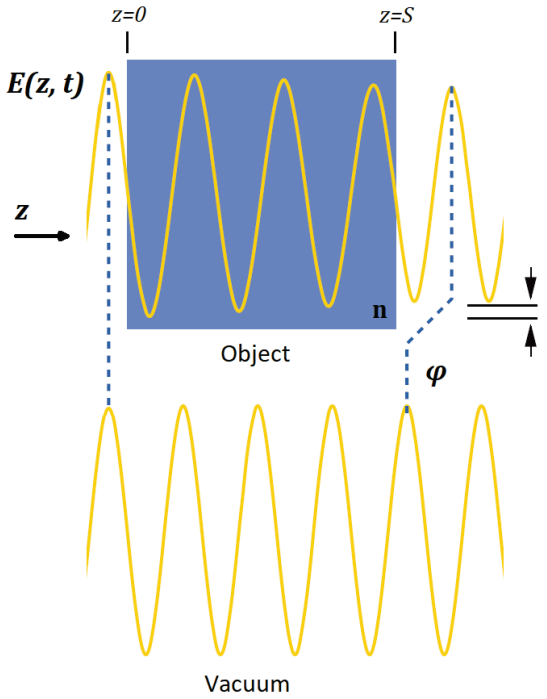


Figure 1.17: Illustration of X-ray refraction

The electric field $E(z,t)$ of a monochromatic or temporally coherent plane wave moving forward (in z-direction) in a homogenous material (object) of refractive index $n = 1 - \delta + i\beta$ and in vacuum. The imaginary part of the refractive index β exponentially attenuates the electric field $E(z,t)$ as shown by the opposing arrows. The phase φ of the electric field passing though the object is reduced compared to that passing through vacuum proportionally to the real part decrement of the refractive index δ. Image adapted from [165].

For the single-material case, the components of the refractive index may be used to define the entities of the phase shift (φ) and the linear attenuation coefficient (μ) as follows [173]:

$$\varphi = -\frac{2\pi}{\lambda}\int \delta \; dz, \qquad (1.4)$$

$$\mu = -\frac{2\pi}{\lambda}\int \beta \; dz, \qquad (1.5)$$

These two entities deliver insight into X-ray attenuation-based and phase-contrast imaging techniques: After interaction with matter, the X-ray wave can be mathematically described by the Fresnel diffraction theory in terms of the phase shift φ and the linear attenuation coefficient μ. When the features of the sample are large compared to the X-ray wavelength λ, small-angle (paraxial) approximations of the wave function may be applied. The X-ray wave at distance z from

the sample is then described by the Fresnel-Kirchhoff integral, which, after application of the Fourier transform, can be written as [173, 180]:

$$F(u) = \Delta(u) - 2m(u)\cos(x) - 2\Phi(u)\sin(x), \quad (1.6)$$

$$x = \pi\lambda z u^2. \quad (1.7)$$

where $F(u)$ is the Fourier transform of the wave function, $\Delta(u)$ is the Dirac delta function, u is the spatial frequency of the transformed radiographies at the image plane, $m(u)$ is the Fourier transform of the linear attenuation coefficient μ, and Φ is the Fourier transform of the phase shift φ. Finally, $\cos(x)$ and $\sin(x)$ are the absorption-contrast and phase-contrast term, respectively. The term x in Equations (1.6) and (1.7) modulates the contrast terms $\cos(x)$ and $\sin(x)$ as shown in Figure 1.18. The practical implications of this are important: absorption contrast and phase-contrast are complementary since the cosine and sine functions are complementary to each other. Note the difference between the absorption/phase component and the respective contrast terms. The sample's structure is portrayed by the absorption $m(u)$ and the phase $\Phi(u)$ components, which are then multiplied by the contrast terms. When either the absorption $m(u)$ or phase component $\Phi(u)$ is small, the respective contrast term needs to be exploited, by selecting x to maximise either the absorption-contrast $\cos(x)$ or the phase-contrast $\sin(x)$ term.

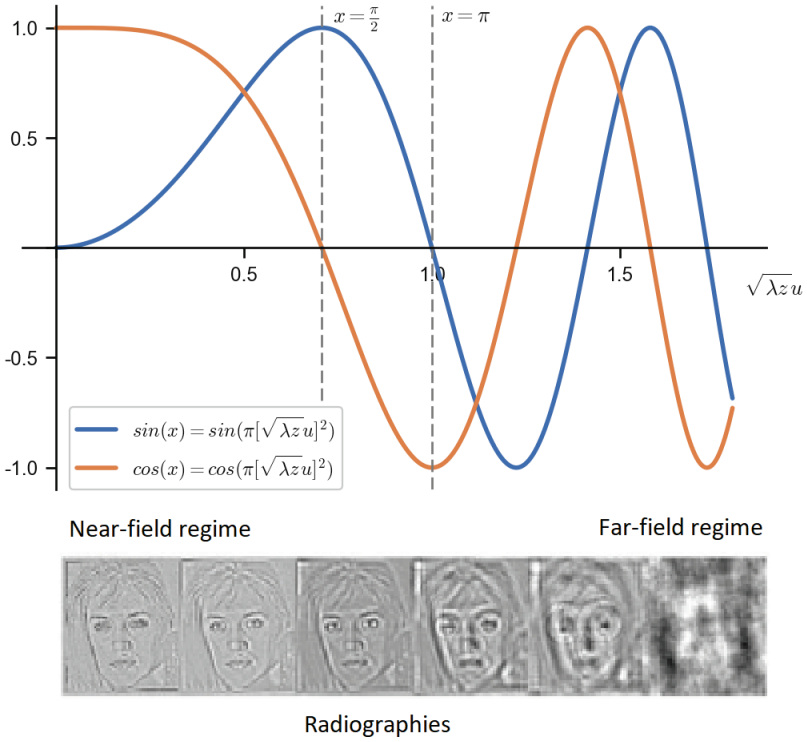


Figure 1.18: The attenuation and phase-contrast terms

The attenuation contrast $\cos(x)$ and phase-contrast $\sin(x)$ terms are plotted against the spatial frequency u scaled by the term $\sqrt{\lambda z}$ where λ is the wavelength, z the X-

ray propagation distance (for the parallel beam case) and $x = \pi\lambda z u^2 = \pi\left[\sqrt{\lambda z}\, u\right]^2$. The near-field imaging regime corresponds to $x \leq \pi$ where the maximal phase-contrast is achieved at $x = \frac{\pi}{2}$. Importantly, in the near-field imaging regime, the experimental setup (wavelength and propagation distance) can only provide either maximal attenuation contrast $\cos(x)$ or maximal phase-contrast $\sin(x)$ due to their complementary nature (cosine and sine functions). The face images (radiographies) show the effect of imaging conditions at the corresponding $\sqrt{\lambda z}\, u$ position on a pure phase object, where u is the peak spatial frequency in the image. As in the far-field regime the face (object profile) is completely obscured by diffraction effects, phase-contrast imaging is performed in the near-field regime. Image adapted with permission from [173] © The Optical Society.

In order to create a 3D image of the sample through CT, the linear attenuation coefficient μ of the material needs to be retrieved. Interestingly, the linear attenuation coefficient μ is related to both the material's electron density ρ_e and to the real part decrement of the refractive index δ, since at beam energies far from absorption edges (sudden sharp increase in absorption spectra caused by resonances between the X-ray photon energy and the minimum energy needed to ionise the atom [165], shown in Figure 1.19), these entities are related as follows [181]:

$$\mu = \sigma_{tot}^{e}\, \rho_e, \qquad (1.8)$$

$$\mu = \frac{2\pi}{r_o \lambda^2}\, \sigma_{tot}^{e}\, \delta, \qquad (1.9)$$

where r_o is the classical electron radius and σ_{tot}^{e} is the total electronic cross-section which describes how the X-ray beam is affected by an electron in the object [181]:

$$\sigma_{tot}^{e} = \sigma_{pe} + \sigma_{coh} + \sigma_{incoh}, \qquad (1.10)$$

$$\sigma_{pe} \approx const_{pe}\, \frac{Z_{eff}^{3.8}}{E^{3.2}} \propto \frac{Z_{eff}^{3.8}}{E^{3.2}}, \qquad (1.11)$$

$$\sigma_{coh} \approx const_{coh}\, \frac{Z_{eff}^{2}}{E^{1.9}} \propto \frac{Z_{eff}^{2}}{E^{1.9}}. \qquad (1.12)$$

where σ_{pe}, σ_{coh}, σ_{incoh} are the electronic cross-sections due to the photoelectric effect, coherent (Thomson) scattering and incoherent (Compton) scattering, respectively. E is the X-ray photon energy, Z_{eff} is the effective atomic number and $const_{pe}$ and $const_{coh}$ designate constant terms. The effective atomic number is the atomic number, i.e. the number of protons, of a fictitious (chemical) element that models the object's material. This fictitious element is modelled as a combination of a number of pure chemical elements in varying proportions [182].

The photoelectric and the coherent terms are inversely related to the X-ray photon energy and directly related the effective atomic number Z_{eff} (Equations (1.10), (1.11) & (1.12)), while the incoherent σ_{incoh} term only depends on the X-ray energy [181]. Although there is no analytical formula for the incoherent electronic cross-section σ_{incoh}, it can be estimated by numerical integration of its differential form and tabulated values are available [183].

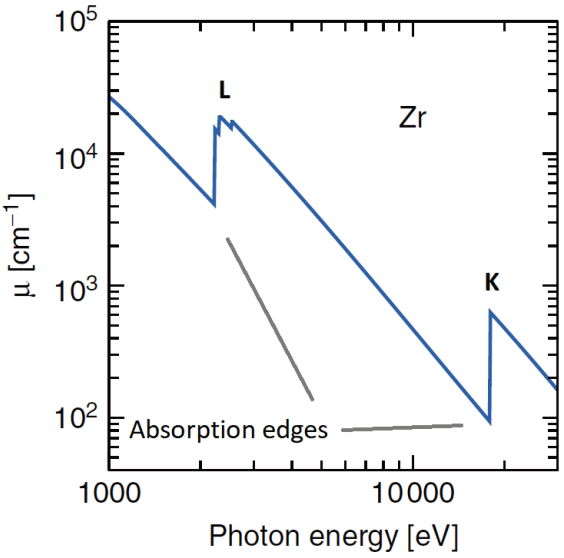


Figure 1.19: The linear attenuation coefficient (μ) of Zirconium (Zr)

The linear attenuation coefficient is inversely related to the photon energy (X-ray energy). Absorption edges are caused by resonances between the X-ray photon energy and the minimum energy needed to ionise the atom at each of its electronic shells (K, L shown here), resulting in sudden sharp increase in absorption spectra. Far from absorption edges, the linear attenuation coefficient μ has a simple relationship to both the material's electron density ρ_e and to the real part decrement of the refractive index δ, since $\mu = \sigma_{tot}^{e}\ \rho_e = \frac{2\pi}{r_o\lambda^2}\ \sigma_{tot}^{e}\ \delta$, where σ_{tot}^{e} is the total electronic cross-section, r_o is the classical electron radius and λ is the X-ray wavelength. Image adapted from [165].

The form of the total electronic cross-section σ_{tot}^{e} results in lower X-ray absorption with increasing energy (Figure 1.19), and for decreasing atomic number (e.g. for soft tissues compared to mineralised tissues). Equation (1.8) indicates that electron-dense (high ρ_e) and high-atomic number (high Z_{eff}) materials yield a high linear attenuation coefficient μ, thus high X-ray absorption. The fact that the linear attenuation coefficient μ, is directly related to X-ray absorption through the imaginary part of the refractive index β in Equation (1.5), and that the X-ray phase shift Φ is directly related to the real part decrement of the refractive index δ in

Equation (1.4), is of paramount importance. It means the internal structure of the object (mass density map or material density distribution) can be retrieved by means of either absorption-contrast (i.e. by retrieving μ) or phase-contrast imaging (i.e. by retrieving δ). Specifically, Equations (1.8) and (1.9) suggest that retrieving either of μ, δ or ρ_e is equivalent in terms of CT reconstruction and relates attenuation-based and phase-retrieval reconstruction algorithms (see section '1.4.5 X-ray phase retrieval'). Notably though, X-ray CT imagers do not record the wave function or the electric field, but the X-ray beam intensity I is measured at the detector instead, which is proportional to the square of the electric field's amplitude [184]:

$$I \propto |E|^2. \qquad (1.13)$$

where | | denotes the Euclidean norm. Therefore, the problem of X-ray CT imaging is reduced to the estimation of the linear attenuation coefficient μ at every point in space using measurements of the beam intensity I (radiographies) on the detector. X-ray CT techniques differ depending on the phenomenon they exploit in order to construct a mass density map of the scanned object. Attenuation-based or absorption-based imaging may be performed when the absorption term $m(u)$ in Equation (1.6) is large enough and the imaging experiment settings are such that the absorption contrast $\cos(x)$ is maximised. When the sample does not absorb adequately, phase-contrast imaging may be performed by exploiting the phase term $\Phi(u)$ in Equation (1.6). In this case, the phase-contrast term $\sin(x)$ is maximised by a careful selection of the experimental settings as discussed in section '3.1 Soft tissue imaging in practice'.

1.4.3 Attenuation-contrast X-ray imaging

When the absorption component $m(u)$ (the Fourier transform of the linear attenuation coefficient μ) is large compared to the phase component $\Phi(u)$ (the Fourier transform of the phase shift φ) in Equation (1.6), the phase component can be neglected [185]. For the following mathematical derivations, the classical CT problem is presented for a parallel X-ray beam, which can be solved for one section of the object or one row of the detector: The intensity $I_\theta(t)$ of a monochromatic X-ray beam at lateral detector position t, moving through an object with linear attenuation coefficient $\mu_\theta(z)$, across an X-ray path along direction z, follows the Beer-Lambert law of a non-homogenous object for a monochromatic source [3]:

$$I_\theta(t) = I_{\theta,0}(t)\, e^{-\int_0^S \mu_\theta(t,z)\, dz} = I_{\theta,0}(t)\, e^{-p_\theta(t)}, \qquad (1.14)$$

where $I_{\theta,0}(t)$ is the incident beam's intensity at lateral detector position t and at an angular position θ, while *S* denotes the object's thickness along the beam direction. $I_{\theta,0}(t)$ is also referred to as 'white image' or 'flat field image' and it is obtained by recording the beam's intensity

without an object in the X-ray path. The integral of the linear attenuation coefficient along the object $p_\theta(t)$ is called 'projection' and it is central in the process of CT reconstruction [3].

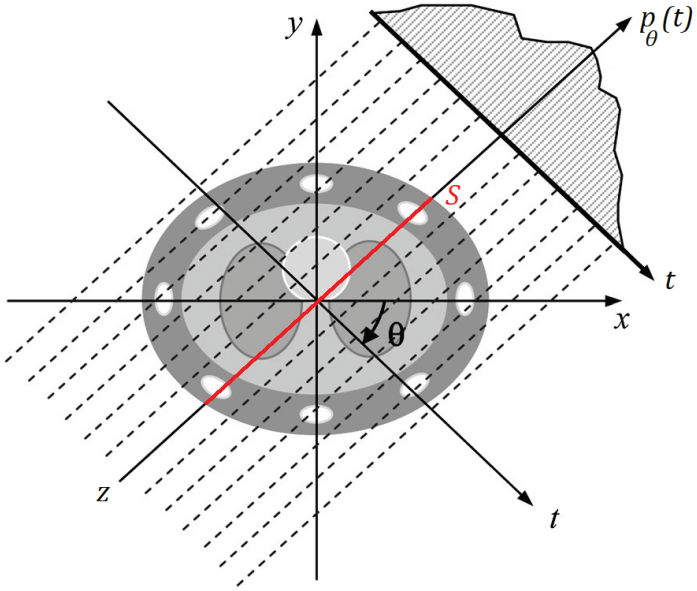


Figure 1.20: X-ray tomography principle

An X-ray beam line (dashed line) propagating across direction z and at relative (to the object) angular position θ, passes through the object and reaches the detector at lateral position t. The integral of the linear attenuation coefficient $\mu_\theta(t,z)$ at each lateral position t and over an object's thickness equal to S is called 'projection' $p_\theta(t) = \int_0^S \mu_\theta(t,z)\, dz$. Projections can be computed using radiographies recorded by the detector $p_\theta(t) = -\ln\left(\frac{I_\theta(t)}{I_{\theta,0}(t)}\right)$, where $I_\theta(t)$ is an 1D radiography at angular position θ, and $I_{\theta,0}(t)$ is the flat field image (radiography without an object in the X-ray path). Tomographic reconstruction is mathematically formulated as the derivation of the linear attenuation coefficient μ using a set of projections $p_\theta(t)$ at different angular positions θ. Image adapted from [4].

Figure 1.20 is a schematic representation of this tomographic principle. The projection $p_\theta(t)$ is calculated by recording a radiography $I_\theta(t)$ with the object at an angular position θ:

$$p_\theta(t) = \int_0^S \mu_\theta(t,z)\, dz = -\ln\left(\frac{I_\theta(t)}{I_{\theta,0}(t)}\right), \qquad (1.15)$$

CT reconstruction for a parallel beam geometry is thus mathematically formulated as the derivation of the linear attenuation coefficient μ using a set of projections, i.e. the transformation of one-dimensional (1D) radiographies (one detector row) to 2D attenuation or mass-density maps (reconstructed CT slice or section) [164]. The reconstructed 3D CT data set is eventually retrieved by putting several reconstructed CT slices together. The mathematical formulation of

the classical CT problem was given by Johann Radon [186, 187] and it is known as Radon transform. A computationally efficient way to calculate the inverse Radon transform and thus to solve the classical CT problem [4] is formulated using the Fourier slice theorem (FST) or central slice theorem [188], which, in the context of CT reconstruction, can be explained as follows [4]: The one-dimensional (1D) Fourier transform of the projection p_θ of an object at angle θ equals a linear intersection $M_\theta(u, v)$ of the 2D Fourier transform of the object's attenuation μ_θ at the same angle (Figure 1.21).

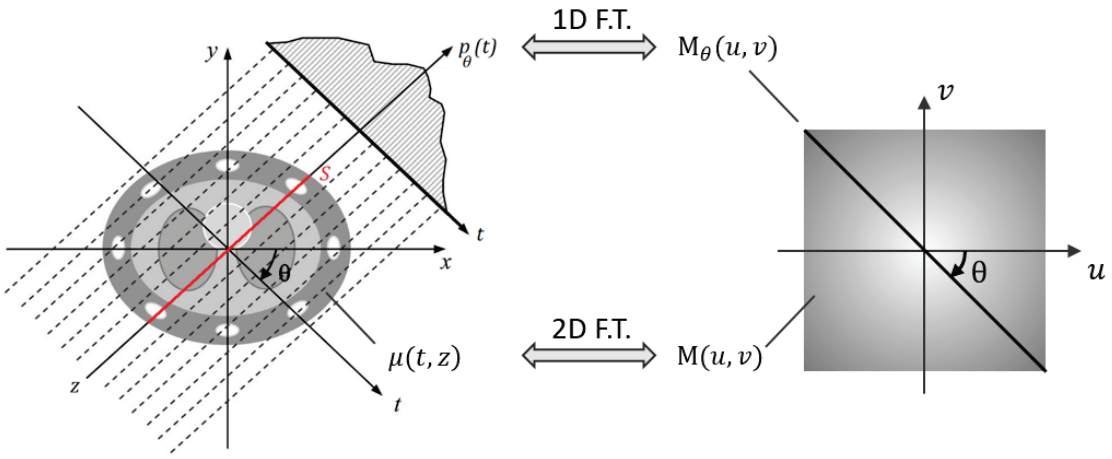


Figure 1.21: The Fourier slice theorem (FST)

A computationally efficient way to solve the classical CT problem is formulated using the Fourier slice theorem (FST) or central slice theorem, which can be explained as follows: The one-dimensional (1D) Fourier transform of the projection $p_\theta(t)$ of an object at angle θ equals a linear intersection $M_\theta(u, v)$ of the 2D Fourier transform of the object's attenuation $\mu(t, z)$ at the same angle, where (u, v) is the Fourier coordinate system (2D frequency). Therefore, by constructing the 2D Fourier transform at angle θ of the linear attenuation M_θ from the 1D Fourier transform of the experimentally recorded projections p_θ for many angular positions θ, allows sampling the Fourier space of the objects' linear attenuation map $M(u, v)$. The object function can then be calculated through an inverse 2D Fourier transform of the object function's sampled Fourier space. Image adapted from [4].

Therefore, by constructing the 2D Fourier transform of the linear attenuation μ_θ from the 1D Fourier transform of the experimentally recorded projections p_θ (FST) for many angular positions θ allows sampling the Fourier space of the linear attenuation map $M(u, v)$ of the object, where (u, v) is the Fourier coordinate system (2D frequency) as shown in Figure 1.22. The object function (the linear attenuation coefficient at each point (x, y) in the 2D plane $\mu(x, y)$) can then be calculated through an inverse 2D Fourier transform of the object function's sampled Fourier

space. This manipulation can be performed with high computational efficiency using the 1D and 2D fast Fourier transform (FFT). However, the Fourier transforms of the projections lie in a polar coordinate system, which densely samples low frequencies and samples higher frequencies in a sparser manner (Figure 1.22). The filtered backprojection (FBP) algorithm accounts for this non-uniform sampling by using a weighting (filtering) function in the frequency domain, namely a high-pass filter, suppressing lower frequencies and reinforcing higher frequencies [4].

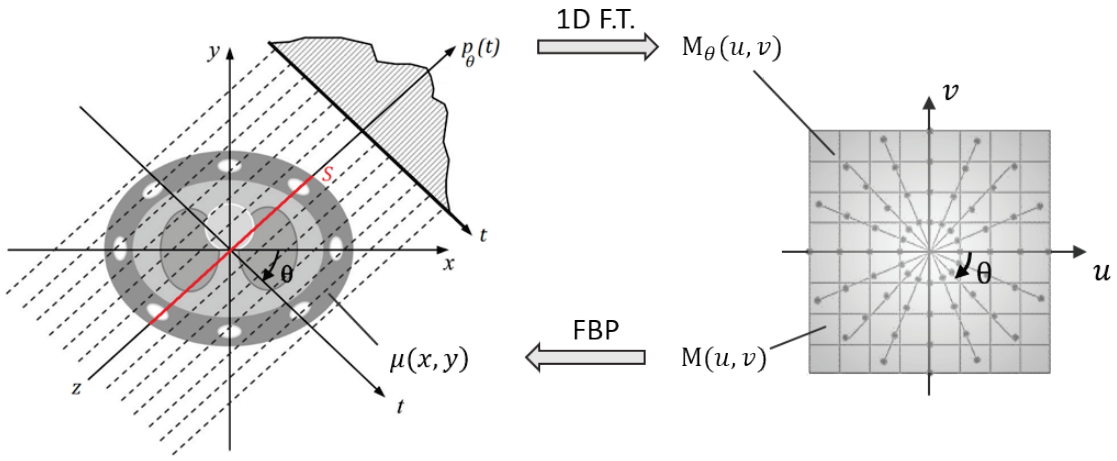


Figure 1.22: The filtered backprojection (FBP) algorithm

The object function (the linear attenuation coefficient at each point (x, y) in the 2D plane $\mu(x, y)$) is derived by applying the following steps: The projections $p_\theta(t)$ are initially experimentally recorded at many angles θ and their 1D Fourier transforms are calculated. These lie in a polar coordinate system, which densely samples low frequencies (closer to the origin) and samples higher frequencies (far from the origin) more sparsely. The filtered backprojection (FBP) algorithm accounts for this non-uniform sampling by using a weighting (filtering) function in the frequency domain, namely a high-pass filter, suppressing lower frequencies and reinforcing higher frequencies. After filtering the linear attenuation map $M(u, v)$ with a high-pass filter in the frequency domain, the FBP applies the inverse 2D Fourier transform to obtain the linear attenuation coefficient (object function) at each point on the 2D plane $\mu(x, y)$. Image adapted from [4].

1.4.4 Phase-contrast X-ray imaging

When the absorption component $m(u)$ is small compared to the phase component $\Phi(u)$ in Equation (1.6), the absorption component can be neglected [185]. The imaginary part of the refractive index β (see Equation (1.2)) drops roughly with the fourth power of the X-ray photon energy $E = \frac{hc}{\lambda} = h\nu$, where h is the Planck' s constant, c is the speed of light, λ is the wavelength

and ν is the frequency of the X-ray photons. Therefore, β drops much faster than the real part decrement δ of the refractive index n, which drops roughly with the square of the energy [165], as schematically shown in Figure 1.23. This quickly renders attenuation-based imaging (dependent on large values of β) difficult for high X-ray energies, particularly for weakly absorbing materials (e.g. soft tissues). In this case, δ can be exploited through phase-contrast imaging.

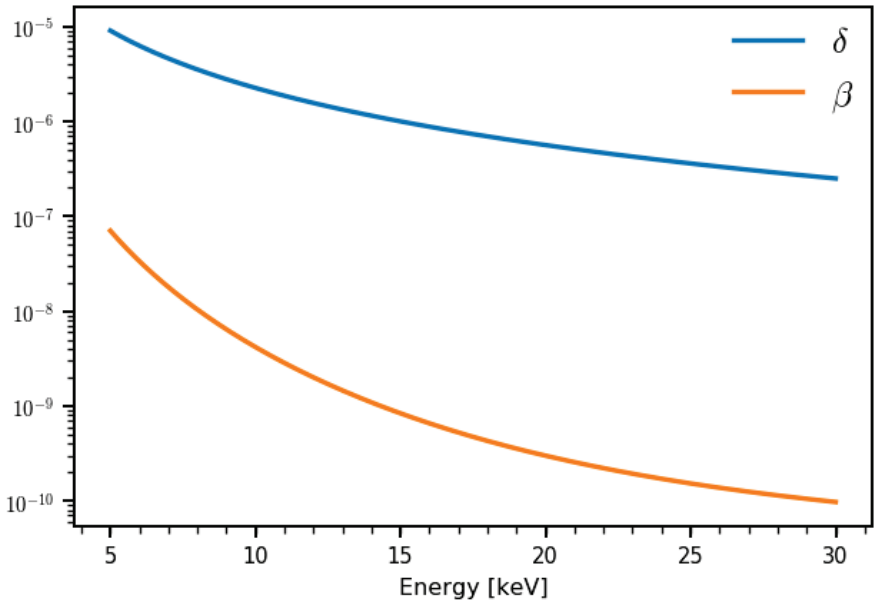


Figure 1.23: Refractive index of soft tissues

Simulated (tabulated values retrieved from [178]) real part decrement δ and imaginary part β of the refractive index for soft tissue at X-ray energies ranging from 5.0 to 30.0 keV. Both entities are unitless. The chemical formula for soft tissue was taken as $C_{12}H_{60}O_{25}N_5$ [165] and the density as 1 g cm^{-3}. The imaginary part β drops much faster than the real part decrement δ, which drops roughly with the square of the energy. This renders attenuation-based imaging (dependent on large values of β) difficult for high X-ray energies, particularly for weakly absorbing materials (e.g. soft tissue). In this case, δ can be exploited through phase-contrast imaging.

The phase shift induced upon the X-ray wave by the sample is recorded as sharp edges in the radiographies, called Fresnel fringes, which may be understood through a Huygens-Fresnel construction [165], shown in Figure 1.24. This is a corpuscle view of the phenomenon that interprets each point on a wave front in the direction of the wave propagation as a point emitter of spherical wavelets. These propagating wavelets interfere with each other, causing all directions other than the wave propagation direction z to cancel out (destructive interference), resulting in a new parallel wave front (Figure 1.24b). In theory, for the destructive interference to happen, it is crucial that the wave is monochromatic, i.e. all point emitters on the wave front are creating

secondary spherical wavelets of the same wavelength as the initial wave. Regardless, this requirement may be relaxed for waves with Gaussian spectrum (Gaussian intensity distribution of the X-ray beam emitted by the source) [180]. In other words, since for laboratory sources the energies of the generated X-rays span a large spectrum (section '1.4.2 X-ray interaction with matter and image formation'), their energy spectrum can be approximated with a single value referred to as the 'effective energy' E_{eff} (weighted mean over the spectrum), as described elsewhere [171, 180, 189]. When an obstacle (a feature in the sample) interrupts the propagation, the wave front cannot be represented anymore by point-emitter wavelets (Figure 1.24b). This withdraws the destructive interference, resulting in a characteristic Fresnel fringe that appears as a sharp edge on the beam intensity on the detector (radiography), as shown in Figure 1.24c.

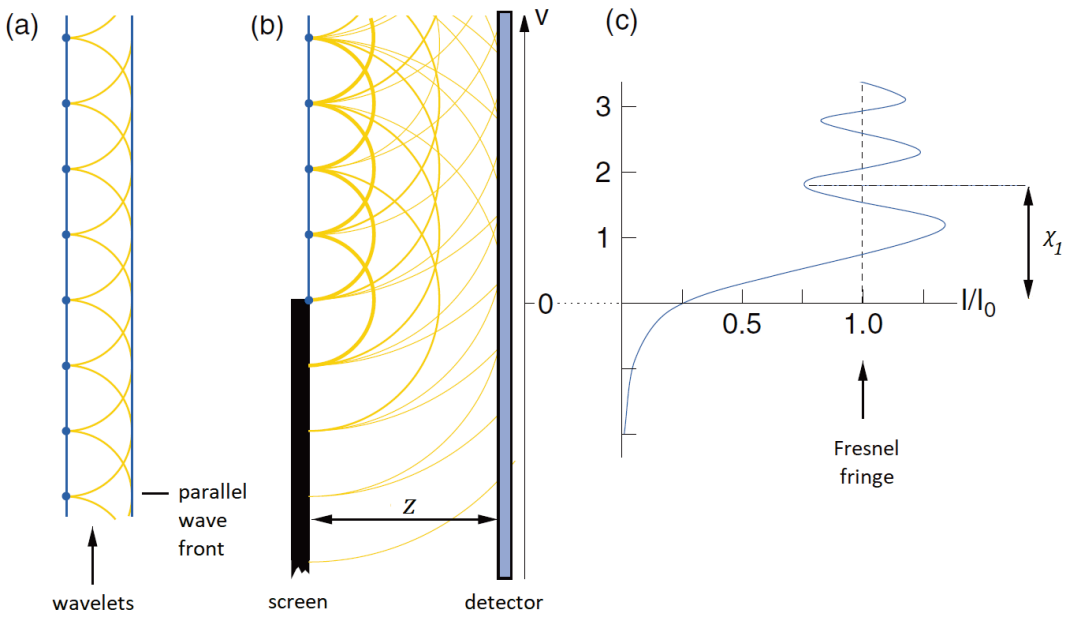


Figure 1.24: The Huygens-Fresnel construction models the Fresnel fringes

(a) The Huygens-Fresnel construction interprets each point on a wave front in the direction of the wave propagation as a point emitter of spherical wavelets. These propagating wavelets interfere with each other, causing all directions other than the wave propagation direction z to cancel out (destructive interference), resulting in a new parallel wave front. (b) When an obstacle (a feature in the sample shown as 'screen' here) interrupts the propagation, the wave front across the obstacle cannot be represented anymore by point-emitter wavelets. (c) This withdraws the destructive interference resulting in a characteristic Fresnel fringe that appears as a sharp edge on the beam intensity recorded by the detector (radiography), and thus as an oscillation around unity on the normalised intensity image $\frac{I}{I_0}$. The width of the first and largest Fresnel fringe x_1 can be calculated given the wavelength λ and the X-ray propagation distance z.

An essential requirement for the destructive interference to happen is uniform distribution of the point-emitter spherical wavelets (i.e. high X-ray spatial coherence). When this assumption breaks, the wavelets do not cancel out entirely, the wave propagates in many different directions and the wave phenomenon that would cause the Fresnel fringes to form cannot be observed anymore. Since the absorption component is typically small in the case of phase-contrast X-ray imaging, absence of Fresnel fringes would yield the object completely invisible for the detector. Therefore, the Huygens construction provides an insight why high X-ray spatial coherence is key for enabling phase-contrast imaging [169]. In practice, full X-ray spatial coherence can only be achieved to a certain extent, mainly due to the finite spot size of the X-ray source. This is mathematically formulated by the transverse coherence length l_c^t (transverse to the propagation direction), which quantitatively characterises the spatial coherence of the beam [180, 190, 191]:

$$l_c^t = \frac{\lambda R_1}{\pi\sigma}, \qquad (1.16)$$

where R_1 is the distance between the X-ray source and the object (source-to-object distance) and σ is the standard deviation of a Gaussian distribution describing the intensity of the X-ray source spot (also referred to as 'source size'). Equation (1.16) describes an X-ray source with a finite spot size of size σ emitting X-ray waves that begin from the source boundaries and propagate in free space (Figure 1.25). Note that alternative definitions of the transverse coherence length exist in the literature (scaled versions of Equation (1.16)) according to each author's coherence criteria [180], e.g. see [165, 192]).

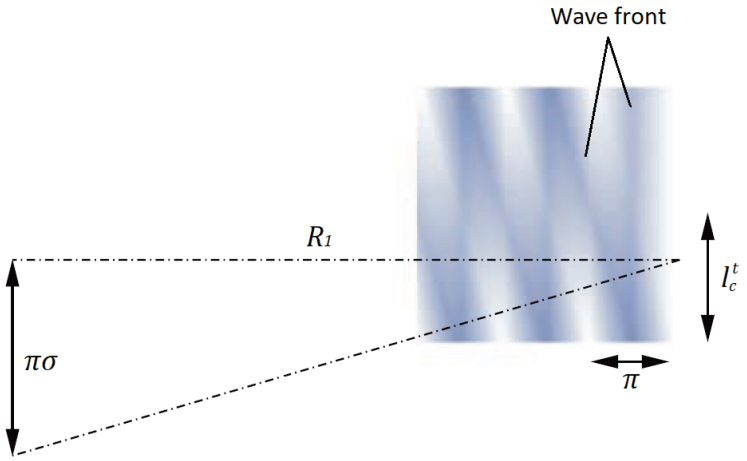


Figure 1.25: The transverse coherence length of an X-ray beam of finite source

An X-ray source with a finite spot size of size σ (lateral extent equal to $\pi\sigma$) generates X-ray waves that begin from the source boundaries and propagate in free space. These waves of identical wavelength λ and slightly different directions of propagation interfere with each other resulting in a transverse coherence length at distance R_1

equal to $l_c^t = \frac{\lambda R_1}{\pi\sigma}$. This is the space required for the phase of a beam to change by π radians, i.e. for the beam to become (fully) incoherent. For constructive interference to take place and Fresnel fringes to be visible at the detector plane, it is necessary that the transverse coherence length is larger than the size of the smallest feature of interest α in the object. This leads to the need for larger source-to-object distances R_1 and/or smaller source sizes σ in order to increase the spatial resolution of an imager (record features of smaller sizes α).

These waves of identical wavelength λ and slightly different directions of propagation interfere with each other resulting in a transverse coherence length at distance R_1 equal to l_c^t. The transverse coherence length is the space required for the phase of a beam to change by π radians [165, 192], i.e. for the beam to become (fully) incoherent. For constructive interference to take place and Fresnel fringes to be visible at the detector plane, it is necessary that the transverse coherence length is larger than the size of the smallest feature of interest α in the object [180]:

$$l_c^t > \alpha. \qquad (1.17)$$

The smallest feature of interest α is the width of a square-pulse-shaped feature as revealed by the analysis below (see also Equation (1.19)). Pogany *et al.* underline that, although Equation (1.17) serves as the lower limit for the smallest feature size of interest, it is not necessary for the whole object or field of view to be coherently illuminated. This is because in reality features do not resemble square pulses, but have arbitrary shapes. Hence, X-ray refraction happens on their boundaries, irrespective of their size or shape. Therefore, an infinitesimally sharp edge (side of a square pulse) would in practice be recorded in the radiography as a (smooth) curve with a width equal to $\frac{\alpha}{2}$.

As l_c^t is proportional to the source-to-object distance R_1, increasing the coherence length is easy in synchrotron sources, where the experimental hutch may be located several meters away from the insertion devices of a synchrotron source where X-rays are generated. For example, at beamline I13-2 of the Diamond Light Source (DLS), Didcot, UK, the X-ray source is about 250 m away from the sample. As this is not possible for laboratory systems, the source size needs to be reduced to lower the transverse coherence length, which is the reason why phase-contrast imaging became possible in the lab only after the development of micro-focus X-ray lab sources.

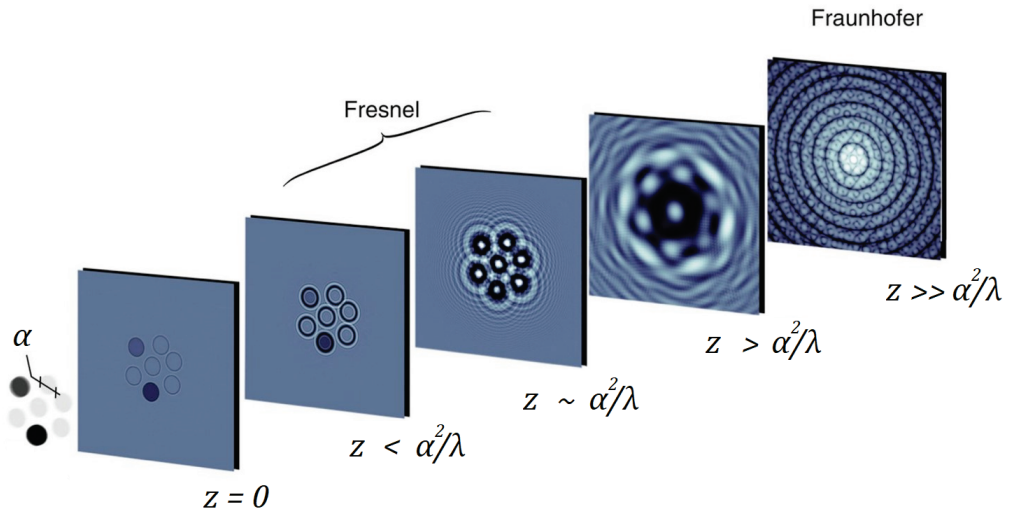


Figure 1.26: Phase-contrast imaging regimes

When X-rays of wavelength λ pass through a pure-phase (non-absorbing) object, Fresnel fringes are developed at the boundaries of materials with different refractive index (features inside the object). The width of Fresnel fringes depends on the feature size α and the propagation distance z. In the Fresnel regime (from $z > 0$ to $z \sim \frac{\alpha^2}{\lambda}$), fringes become increasingly more pronounced with larger distances and the object remains recognisable by its outline due to sharpened edges. The edges become increasingly fuzzy for propagation distances $z \geq \frac{\alpha^2}{\lambda}$ until the object completely disappears at the Fraunhofer regime ($z \gg \frac{\alpha^2}{\lambda}$) where the far-field diffraction pattern is recorded by the detector. Image adapted from [165].

Fresnel fringes form due to interactions between the X-ray wave and the boundary and internal structural features of the object, and evolve as the X-ray wave propagates, beginning as enhanced edges that become wider until they completely overlap, as shown in Figure 1.26. Depending on the X-ray propagation distance (distance between the sample and the plane of observation), the phenomenon is modelled either by the near-field (Fresnel) approximation or the far-field (Fraunhofer) approximation. In the Fraunhofer regime the object profile is completely obscured by diffraction effects and the recorded radiography is proportional to the Fourier transform of the object [165]. In this regime, holographic methods may be used to perform high-resolution imaging [173]. Free-space or inline propagation-based phase-contrast X-ray imaging is performed in the Fresnel (near-field) regime, where the object remains recognisable by its outline due to sharpened edges: In brief, for sufficiently small x (near-field approximation) the phase-contrast term in Equation (1.6) may be approximated by $\sin(x) \approx x = \pi\lambda z u^2$ as detailed by Pogany and colleagues [180]. Therefore when $x \leq \pi$, (corresponding to the regime up to the first zero

crossing of the phase-contrast term $\sin(x)$) Equation (1.7) (see also Figure 1.18) results in a required X-ray propagation distance of [173]:

$$z \leq \frac{1}{\lambda u^2}, \quad (1.18)$$

where the feature size α and the spatial frequency u are related by (by definition):

$$\alpha \equiv \frac{1}{u}, \quad (1.19)$$

This relationship between the spatial frequency u and the feature size α can be understood by looking at the Fourier transform of a square pulse of width α, which has its fundamental frequency $u = \frac{1}{\alpha}$. Therefore, the feature size α can be thought of as the width of a square-pulse-shaped feature of width equal to α. When the definition of the feature size in Equation (1.19) is combined with Equation (1.18), this yields the relationship between the propagation distance and the feature size for the near-field approximation:

$$z \leq \frac{\alpha^2}{\lambda}, \quad (1.20)$$

For optimal phase-contrast the term $2\Phi(u)\sin(x)$ in Equation (1.6) must be maximised. Since $x \leq \pi$ needs to hold in order to meet the near-field imaging conditions, x needs to be equal to $\frac{\pi}{2}$ in Equation (1.6), as schematically shown in Figure 1.18. Re-arranging the approximation $\sin(x) \approx x = \pi\lambda z u^2$ and replacing u in Equation (1.19) shows the implications of this, i.e. that only features of the following size will be at maximal phase-contrast [180]:

$$\alpha = \sqrt{2\,\lambda\,z}, \quad (1.21)$$

Focusing on the most common, indirect X-ray detection method through scintillators (see section '1.4.1 Computed tomography principles'), the X-ray wave intensity is recorded at the scintillator plane. The width of the first (and largest) Fresnel fringe on the plane of observation, which is the scintillator plane, is given by [165] (Figure 1.24):

$$x_1 = \sqrt{\frac{\lambda z}{2}}. \quad (1.22)$$

Note that Equation (1.22) describes the Fresnel fringe width, which is physical quantity. In practice, this signal is recorded by a (digital) detector, which suffers from imperfections such as spatial averaging of the photon sums around each pixel and (measurement, thermal, etc.) noise [193]. The net effect of the imperfections in the system can then be modelled by the system's (Gaussian) Point Spread Function (PSF), which is how a system would record an infinitesimally small point [193]. This effectively means that the physical quantity, i.e. the X-ray wave intensity, is

blurred or (mathematically) convoluted with a Gaussian function. One can imagine a painter trying to draw a hairline with a thick brush. This process results in a further 'widening' of the Fresnel Fringes, which relaxes the requirement for the pixel size of the system to be able to record them. However, the PSF cannot be directly modelled, but is typically experimentally derived [193]. Thus, the effects of the PSF are disregarded for the purposes of the work presented here.

Shannon's sampling criterion [194], which specifies the conditions necessary for adequate sampling of a given signal, states that the pixel size at the scintillator must be at least two times smaller than x_1 in order to properly record the Fresnel fringes (disregarding the PSF). Markedly, x_1 is typically very small for hard X-ray rays. For instance, at an X-ray beam of 20.0 keV ($\lambda = 0.062\ \mathrm{nm} = 0.62\ \text{Å}$) after 20 mm of propagation this would result in a width of the first Fresnel fringe $x_1 = 1.24\ \mu m$. Therefore, recording Fresnel fringes is particularly difficult at pixel sizes of typical scientific camera systems, such as the pco.edge 5.5 (PCO AG, Kelheim, Germany) for instance, which is 6.5 µm [195], without additional optical magnification after the scintillator.

Free-space or inline propagation-based phase-contrast X-ray imaging takes advantage of the proportionality between the Fresnel fringe width x_1 and the (square root of the) X-ray propagation distance (Equation (1.22)), in order to allow the Fresnel fringes to widen (increase x_1) through free-space propagation (over the X-ray propagation distance z). On this account, it becomes easier to record the Fresnel fringes, i.e. alleviates the need to use further optical elements or to modify the experimental setup [169, 173]. Further to the capability of recording the Fresnel fringes, inline propagation-based phase-contrast X-ray imaging is facilitated by the development of single-plane phase retrieval algorithms. That makes it possible to retrieve the real part decrement δ of the refractive index (see Equation (1.4)) from the X-ray wave intensity signal, recorded at a single propagation distance. In practice, this enables the same imager geometry to be used to perform both X-ray attenuation-based and phase-contrast imaging experiments. The phase-retrieval step is performed before standard CT reconstruction in order to remove the Fresnel Fringes from the recorded radiographies as explained in the following section.

1.4.5 X-ray phase retrieval

After imaging is performed with the conditions necessary to record the Fresnel fringes in the near-field regime, the so-called 'phase problem' needs to be solved [173]. That is, either of the linear attenuation coefficient μ, the real part decrement of the refractive index δ or the object's electron density ρ_e needs to be retrieved from X-ray intensity measurements, because these entities are proportional to each other far from absorption edges as discussed in section '1.4.2 X-ray interaction with matter and image formation'. One of the first phase retrieval algorithms

designed for propagation-based phase-contrast imaging was developed by Andrei Bronnikov [196]. It assumes that the object does not absorb X-rays (pure phase object), solving the Radon transform in the Fourier domain, in order to compute the real part decrement of the refractive index δ from a set of intensity measurements at a single X-ray propagation distance. In the original publication [196], Bronnikov presented the mathematical derivation of a frequency filter in the Fourier domain, yet without providing simulations or application on real data. In practice, this algorithm is unstable for low spatial frequencies in the radiographies due to division by zero (frequency term in the filter's denominator). The Modified Bronnikov Algorithm (MBA) [197, 198] attempts to remove the instability by introducing a regularisation term in the denominator. An alternative approach was presented by Wu *et al.* [199], who exploited the complementary nature of phase and attenuation. Their algorithm emanates from a quantum mechanics background and it requires prior knowledge of the atomic number of the object's material (composition) to estimate the electron density ρ_e. The authors did not prove their derivation in their original work nor did they apply their algorithm on real data, but they presented results on simulated data instead. A similar approach to the MBA was taken by David Paganin *et al.* [200]. The starting point of this derivation is the Transport-of-Intensity Equation (TIE), which models the X-ray beam propagation under near-field conditions [173]. The authors assumed a monochromatic X-ray beam and a low-absorbing, homogeneous object, to provide a closed-form solution of the TIE, in order to achieve what they describe as simultaneous phase and amplitude extraction from a single defocused image. Their method revolves around the so-called projected thickness T, which is an estimation of the length of the path through the sample that the X-rays travelled at each point of the radiography:

$$T = -\frac{1}{\mu} ln\left\{F^{-1}\left\{\frac{F\left\{M_g^2 \frac{I}{I_0}\right\}}{1+\frac{\delta}{\mu}\frac{z}{M_g}(u^2+v^2)}\right\}\right\} = -\frac{1}{\mu} ln\left\{F^{-1}\left\{\frac{F\left\{M_g^2 \frac{I}{I_0}\right\}}{1+\frac{\delta}{\mu}Z_{eff}(u^2+v^2)}\right\}\right\}, \quad (1.23)$$

where u and v are the lateral and vertical spatial frequencies in the Fourier domain, F and F^{-1} are the forward and inverse Fourier transform respectively, M_g is the magnification from the object's plane to the scintillator's plane (otherwise called geometric magnification), z is the X-ray propagation distance (i.e. the distance from the object's to the scintillator's plane), I is the intensity of the recorded radiography and I_0 the flat field image (intensity of radiography without the sample in the beam path). The term $z_{eff} = \frac{z}{M_g}$ is alternatively called effective propagation distance [173]. The geometric magnification M_g appears to downscale the actual X-ray propagation distance, which influences the degree of smoothing applied to the radiographies by the Paganin algorithm, as discussed next. Additionally, the square of the geometric magnification M_g^2 multiplies the normalised intensity $\frac{I}{I_0}$ in the nominator of Equation (1.23). This means that

imaging an object (with a given projected thickness T) in a setup with a non-unit geometric magnification (e.g. cone-beam geometries typical in laboratory sources have $M_g > 1$ as discussed in section '3.1.1 Generic imager geometry') would yield radiographies M_g^2 times darker compared to imaging the same object in a setup with $M_g = 1$ (e.g. parallel-beam geometries typical in synchrotron sources). The projected thickness T conveniently takes the role of the X-ray path in the Beer-Lambert law [200]:

$$\check{I} = \check{I}_0 e^{-\mu T}, \qquad (1.24)$$

filtering the normalised radiography $\frac{I}{I_0}$ as follows:

$$\frac{\check{I}}{\check{I}_0} = F^{-1}\left\{\frac{F\left\{M_g^2 \frac{I}{I_0}\right\}}{1+\frac{\delta}{\mu} z_{eff}(u^2+v^2)}\right\}. \qquad (1.25)$$

The filtered normalised radiography $\frac{\check{I}}{\check{I}_0}$ is then provided as input to the reconstruction algorithm to compute the linear attenuation coefficient μ. Note that the linear attenuation coefficient μ is varying spatially in the radiography image plane, and thus, the ratio $\frac{\delta}{\mu}$ is constant only when δ and μ are proportional. This generally holds (see Equation (1.9)) far from absorption edges as discussed in section '1.4.2 X-ray interaction with matter and image formation'. Markedly, Bronnikov' s and Paganin' s algorithms estimate δ and μ, respectively. Hence, they are directly comparable since δ and μ are proportional (Equations (1.8) and(1.9)). The linear attenuation coefficient μ and the imaginary part of the refractive index β are related by the optical theorem [201, 202]:

$$\mu = \frac{4\pi\beta}{\lambda}, \qquad (1.26)$$

which, when replaced in equation (1.25), shows (denominator: $1 + \frac{\lambda}{4\pi}\frac{\delta}{\beta}\, z_{eff}(u^2 + v^2)$) that the filtering of the normalised intensity depends on the wavelength λ and the ratio of both components of the (complex) refractive index $\frac{\delta}{\beta}$. A close look at equation (1.25) reveals that in essence, the Paganin algorithm acts as a low pass filter in the frequency domain, with its cut-off frequency equal to $|s| = \sqrt{u^2 + v^2} = \sqrt{\frac{\mu}{\delta}\frac{M_g}{z}} = \sqrt{\frac{\mu}{\delta}\frac{1}{z_{eff}}}$. This represents a disk centred at the origin of the frequency space with radius equal to $\sqrt{\frac{\mu}{\delta}\frac{1}{z_{eff}}} = \sqrt{\frac{4\pi}{\lambda}\frac{\beta}{\delta}\frac{1}{z_{eff}}}$. The filter suppresses frequencies outside the disk, resulting in smoothed radiographies (or blurred radiographies for strong filtering), effectively suppressing the Fresnel fringes. This cut-off frequency depends on both the geometry of the imaging experiment $\frac{M_g}{z} = \frac{1}{z_{eff}}$ and the ratio $\frac{\mu}{\delta}$ (or the ratio $\frac{\beta}{\delta}$ emanating

directly from the refractive index), which requires prior knowledge of the material properties of the studied object. It also means that only one single material, or materials with the same $\frac{\mu}{\delta}$ ratio to be exact, can be optimally reconstructed at a time. As a given sample may be composed of several materials, the Paganin algorithm was later extended for multiple materials [179]. This extended algorithm retrieves the phase for all pairs of materials and digitally combines the resulting images to create the phase image of the scanned sample. Thus, it is exceedingly complex to apply in practice, since it requires the knowledge of the properties of all materials the object is composed of, which are generally not known. Regardless of its low-pass filtering effect, the Paganin phase retrieval algorithm is computationally efficient since it utilises the Fourier transform, and pleasingly stable due to the presence of the unity in the denominator of Equation (1.25) [200]. It is thus well established across the X-ray imaging community, with X-ray imaging experimental stations at major synchrotron facilities (e.g. TOMCAT beamline at the Swiss Light Source (SLS) [203] or beamline I13-2 at Diamond Light Source (DLS) [204]) that provide an implementation in their respective in-house software suites.

1.4.6 X-ray imaging modes

Depending on the experimental settings and setup (beam energy, X-ray propagation distance, etc.) and image processing needs (phase retrieval, etc.), different X-ray imaging modes can be defined, which are summarised here.

i. Object imaged using a low-energy and incoherent X-ray beam:

The simplest case is the absorption-contrast or attenuation-based X-ray imaging mode [169]. When the experimental setup is such that the absorption component $m(u)$ is large compared to the phase component $\Phi(u)$ (see Equation (1.6)), the sample appears as a smooth dark shadow, with no pronounced edges or fringes in the radiographies. Figure 1.27 shows an exemplar profile plot across such a radiography, where the object's profile appears as a square with no overshoot or undershoot at its boundaries. Direct CT reconstruction (without prior phase retrieval) of these radiographies yields the typical CT images where the object appears brighter than the background, while the object's edges are not significantly pronounced.

ii. Object imaged using a low-energy, coherent and monochromatic X-ray beam:

The second case considers an object imaged with a coherent and monochromatic beam. In this case, both components of X-ray refraction (attenuation and phase shift) are observed and depending on the experimental setup the phase $\Phi(u)$ and absorption $m(u)$ terms might both be non-zero. This results in shadows in the radiographies with Fresnel fringes at the boundaries, yielding profiles with overshoot and undershoot at the object's edges (Figure 1.27). At this point,

if these radiographies are reconstructed without prior phase-retrieval, the sample will appear darker (than in i.) with pronounced edges. This mode is called 'edge enhancement'. Alternatively, phase retrieval can be applied prior to CT reconstruction, in order to remove the Fresnel fringes yielding profiles with smoother feature boundaries. After CT reconstruction of these phase-retrieved radiographies the object appears brighter than the edge-enhanced case (CT reconstruction without prior phase retrieval) and its edges are smoother. This mode is called 'phase-contrast imaging' or 'phase-retrieved imaging'.

iii. Object imaged using a high-energy, coherent and monochromatic X-ray beam:

The third case considers an object imaged with a coherent, monochromatic beam and an X-ray energy high enough so that the sample is not attenuating significantly (i.e. pure phase object). An exemplar profile of such a case is shown in Figure 1.27, where similarly to ii., there are two alternative approaches in terms of CT reconstruction (with and without prior phase retrieval). If the phase is not retrieved prior to CT reconstruction, the sample appears as a set of edges in the reconstructed data. As in this case no shadows are visible in the radiographies due to the negligible X-ray absorption, the inside of the sample in the reconstructed data is similar to the background. Alternatively, if the phase is retrieved prior to CT reconstruction, the sample will appear brighter than the background, with smooth edges. As in ii., this mode is called 'phase-contrast imaging' or 'phase-retrieved imaging'.

A close look at these three cases and section '1.4.2 X-ray interaction with matter and image formation' reveals the crucial difference between attenuation-contrast and phase-contrast imaging: The experimental conditions have to be optimised for the chosen imaging mode to maximise either the absorption-contrast term $\cos(x)$ to exploit the absorption component $m(u)$, or the phase-contrast term $\sin(x)$ to exploit the phase component $\Phi(u)$, according to the preferred imaging mode (see Equation (1.6)). As the absorption-contrast term and the phase-contrast term are complementary, it is in principle not possible to maximise both attenuation-contrast and phase-contrast, at the same time.

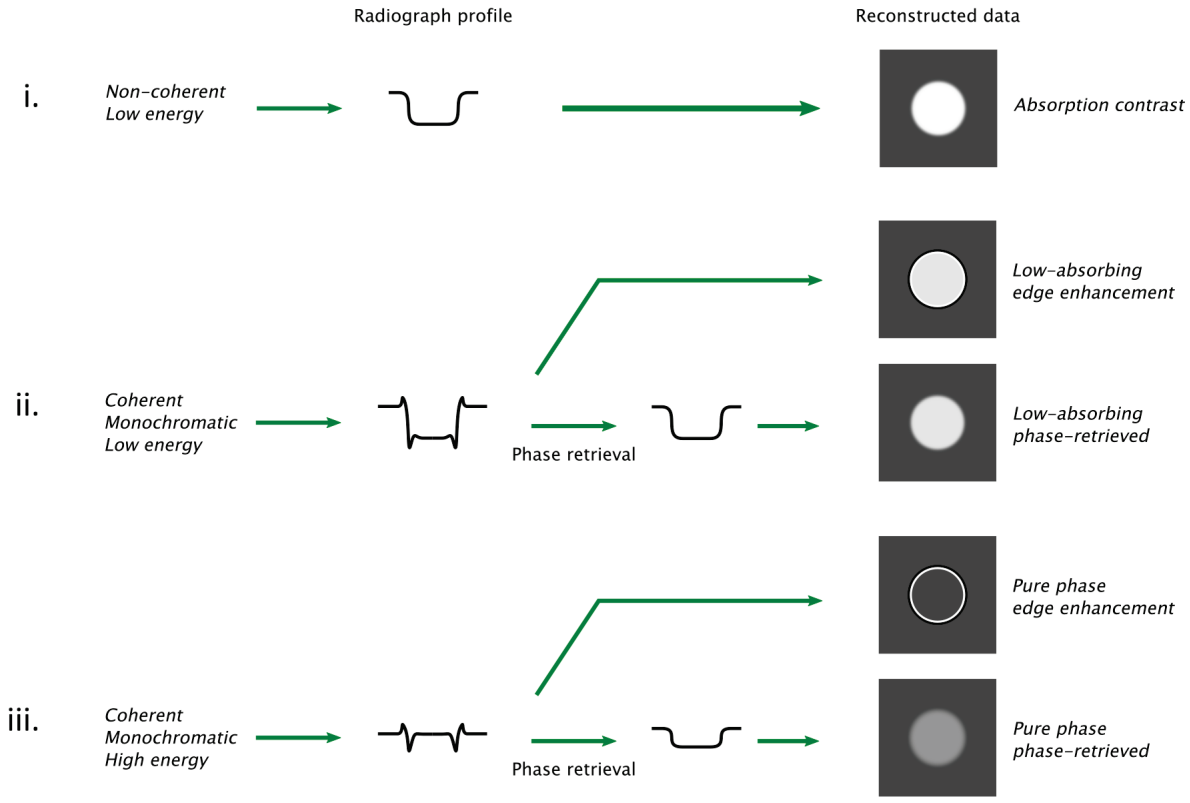


Figure 1.27: X-ray imaging modes

i. When a sample is imaged with a non-coherent and low energy beam it appears as a smooth dark shadow with no pronounced edges or fringes in the radiographies. The object's profile plot across such a radiography appears as a square with no overshoot or undershoot at its boundaries. Direct CT reconstruction (without prior phase retrieval) yields the typical CT images where the object appears brighter than the background, while its edges are not significantly pronounced (absorption contrast). ii. When an object is imaged with a coherent and monochromatic beam, both components of X-ray refraction (attenuation and phase shift) are observed. This results in shadows in the radiographies with Fresnel fringes at the boundaries, yielding profiles with overshoot and undershoot on the object's edges. If these radiographies are reconstructed without prior phase-retrieval, the sample will appear darker (than in i.) with pronounced edges (low-absorbing 'edge enhancement'). If phase retrieval is applied prior to CT reconstruction, the object appears brighter than the edge-enhanced case and its edges are smoother (low-absorbing phase-retrieved). iii. When an object is imaged with a monochromatic beam and an X-ray energy high enough so that it is not attenuating (i.e. pure phase object) there is no shadow on the radiographies. The radiograph profile, similarly to ii., has Fresnel fringes at the

boundaries but it is otherwise flat. If the phase is not retrieved prior to CT reconstruction, the sample appears as a set of edges in the reconstructed data, and the inside of the sample is similar to the background due to negligible X-ray absorption (pure phase, edge enhancement). Alternatively, if the phase is retrieved prior to CT reconstruction, the sample will appear brighter than the background, with smooth edges (pure phase, phase-retrieved).

1.5 Thesis overview

As discussed in section '1.2 Colorectal cancer and crypt budding', crypt budding precedes the earliest lesions on the colonic mucosal surface (ACF) during the development of CRC. As such, crypt budding is lately receiving increased attention in CRC research, with recent breakthroughs attempting to link functional and structural information discussed in section '1.3 Standard approaches for crypt budding research'. However, our understanding of crypt budding remains limited. This is partly due to the lack of experimental methods that enable *in situ* 3D imaging of concurrent structural and functional information of a large numbers of crypts.

Preparing samples following standard histology procedures, i.e. in the form of formalin-fixed and paraffin embedded (FFPE) tissue blocks, enables standard 2D histology and serial paraffin sectioning methods to be performed on the same sample. Thus, developing a workflow for non-destructive imaging of the internal structure of such samples and subsequent application of standard histology procedures would provide both structural and functional information, facilitating the study of crypt budding during the early-stage CRC.

To this end, the work presented in this thesis deals with the development of new engineering tools that enable users with non-technical backgrounds to use advanced 3D imaging and signal processing techniques for the structural quantification of budding crypts within FFPE colonic tissue samples.

1.5.1 Objective

The main objective of this thesis is to develop the engineering tools necessary for 3D imaging of FFPE colonic tissue using inline X-ray phase-contrast μCT and morphological quantification of budding crypts in 3D.

1.5.2 Specific aims

i. **Application of an animal model for CRC**
   To employ a well-established AOM/DSS murine model in order to produce FFPE colons to be subsequently imaged and quantified. The model needs to recapitulate aspects of CRC so that crypt budding can be observed.

ii. **X-ray phase-contrast μCT imaging of the budding crypt microstructure**
   To develop a method that enables X-ray phase-contrast μCT imaging of murine colon samples that is easy to apply by non-experts in physics or engineering.

iii. **Digital colon unrolling to aid budding crypt identification**
   To develop a method that facilitates the identification of budding crypts within the μCT scans in an efficient and easy way.

iv. **Morphological quantification of crypt budding in 3D**
   To develop a method that enables objective quantification of the structure of identified budding crypts.

Chapter 2 Application of an animal model for colorectal cancer

In Chapter 1 the need for new tools for the study of colorectal cancer (CRC) was identified. Specifically, the need for 3D imaging approaches that enable *in situ* 3D morphological quantification of crypt budding and subsequent functional assays via histological or immunohistochemical methods. X-ray micro-computed tomography (µCT) was identified as a suitable candidate for non-destructive 3D imaging. Performing functional assays after imaging becomes possible when µCT imaging is combined with sample preparation methods compatible with standard histology workflows, i.e. µCT imaging of colons embedded in paraffin wax blocks. This chapter discusses the materials and methods employed to obtain animal colon tissues induced with CRC in the form of formalin-fixed and paraffin-embedded (FFPE) wax blocks, which are suitable for non-destructive µCT imaging.

Section 2.1 discusses the materials and methods employed to induce CRC to mice through an appropriate animal model (section '2.1.1 Azoxymethane and dextran sodium sulphate mouse model') and the relevant FFPE tissue sample preparation method (section '2.1.2 Tissue sample acquisition and preparation'). Next, application of standard sectioning and histology staining (section '2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining') serves as a proof of concept for the complete workflow of sample preparation followed by µCT imaging and conventional histology. Finally, section '2.2 Animal model results' discusses the measurements taken during the application of the animal model to ensure that CRC was effectively induced.

2.1 Animal model materials and methods

All mice were bred and maintained at the biomedical research facility (BRF), University of Southampton in accordance with the United Kingdom animals scientific procedures act (ASPA) 1986. This study has received ethical approval from the local ethics committee animal welfare ethics research board (AWERB) and was operated under an existing Home Office animal project licence. The regulatory protocols were performed according to the animal research reporting of *in vivo* experiments (ARRIVE) guidelines [205] and in compliance with the replacement refinement & reduction (3R) policy [206]. Animals were kept in an environment with a 12h light/dark cycle, and at a constant temperature of 22 ± 2 ^{o}C with sterilised food available *ad libitum*.

2.1.1 Azoxymethane and dextran sodium sulphate mouse model

The azoxymethane and dextran sodium sulphate (AOM/DSS) mouse model for inflammation-associated colorectal cancer was used to produce murine colon samples at three progressive stages of CRC. The model uses AOM, which is a procarcinogen that inflicts DNA damage before it is excreted into the bile and taken up by the colonic epithelium [207], and DSS, which is a polysaccharide that inflicts epithelial damage in the colon and inflammatory response [208]. The specific protocol applied was adapted from Neufert *et al.* [103] and Parang *et al.* [209] and summarised in Table 2.1. The treatment schedule was applied to thirty-six C57BL/6J male mice that were 8 weeks old at the beginning of the treatment. The inflammation-associated CRC model was applied on a group (AOM/DSS) of 18 individuals, housed in 3 cages with 6 mice in each cage. Three more groups were used as controls, a no-treatment control group or untreated (6 mice, 1 cage) and two carcinogenic substance control groups, AOM only (6 mice, 1 cage) and DSS only (6 mice, 1 cage). Mice in the AOM/DSS group were treated with one intra-peritoneal (I.P.) injection of 10 mg/kg AOM solution on day 1. Starting on day three, they were treated with up to three cycles of inflammation and healing: During each cycle they were provided with 2 % weight/volume (wt/vol) DSS solution in sterile water for 7 days followed by 14 days of sterile water (inflammation and healing phases, respectively) provided *ad libitum* via a single water drinking bottle per cage.

Table 2.1: The AOM/DSS treatment schedule

Mice in the AOM and AOM/DSS groups were given one injection of 10 mg/kg AOM solution on day 1. Mice in the DSS and AOM/DSS groups were subsequently treated with up to three cycles of: 7 days of 2 % wt/vol DSS solution and 14 days of sterile water. Red lines represent the culling days (24, 45 and 66). The minus ('-') signs indicate absence of AOM administration, whereas the plus ('+') signs indicate administration of AOM via one I.P. injection per mouse.

	AOM injection		**Cycle 1**		**Cycle 2**		**Cycle 3**	
Week			1	2, 3	4	5, 6	7	8, 9
Day(s)	1	1 - 3	3 - 9	10 - 23	24 - 30	31 - 44	45 - 51	52 - 65
Untreated group	-	Water	Water	Water	Water	Water	Water	Water
AOM group	+	Water	Water	Water	Water	Water	Water	Water
DSS group	-	Water	DSS	Water	DSS	Water	DSS	Water
AOM/DSS group	+	Water	DSS	Water	DSS	Water	DSS	Water

When the animals are treated with the inflammatory agent (DSS), their body weight is expected to drop. The effect is acute (within one or two days) and it indicates the loss of bodily fluids due to decreased appetite, diarrhoea and bad welfare [209]. In order to minimise animal suffering, a disease activity scoring system and an animal welfare scoring system have been employed as described in detail in the supplementary standard operating procedure (SOP) (see supplementary material S1). A relative weight loss limit was set: No more than 20 % of weight loss (with respect to the weight on the first day of DSS cycle 1 for each individual, i.e. day 3 of the experiment) was allowed. DSS administration was withdrawn once critical weight loss (>17 %) was observed. Animals were euthanized in cases where their weight loss exceeded 20 % with no signs of health improvement (as scored by the animal welfare scoring system) after DSS withdrawal. Unless early termination was necessary due to welfare reasons as detailed below, animals were culled via cervical dislocation according to the plan of Table 2.2: 6 mice of the AOM/DSS group and 2 mice of each of the rest of the groups at the end of each cycle. Diverging from this schedule, out of the 18 animals in the AOM/DSS group, 5 had to be euthanized early (two on day 10, one on day 11, one on day 14, and one on day 52). The three different points in time (days 24, 45 and 66) represent progressive stages of CRC (cycle 1: early, cycle 2: mid, cycle 3: late).

Table 2.2: Culling plan for the AOM/DSS murine model

Animals were culled via cervical dislocation at three time points (days 24, 45 and 66), representing progressive CRC stages (cycle 1: early, 2: mid, 3: late). Diverging from this schedule, out of the 18 animals in the AOM/DSS group, 5 were euthanized for welfare reasons (two on day 10, one on day 11, one on day 14 and one on day 52).

Cycle	1	2	3
Day	24	45	66
Untreated	2	2	2
AOM	2	2	2
DSS	2	2	2
AOM/DSS	6	6	6

2.1.2 Tissue sample acquisition and preparation

For accurate morphological quantification of crypt budding it is vital to preserve the native colonic shape and minimise tissue deformation due to the sample preparation procedures. Therefore, the sample acquisition and preparation protocols were designed to inflict minimal damage to the

colons while ensuring that they remained intact, so that the natural tubular shape was preserved. Longitudinal cutting was avoided and the use of sharp tools was minimised. The freshly excised colons were turned into FFPE cylindrical blocks suitable for µCT imaging, which may subsequently be re-embedded in standard FFPE wax blocks appropriate for paraffin cutting and histology staining. This section describes the process to turn freshly excised colons to cylindrical and standard FFPE wax blocks.

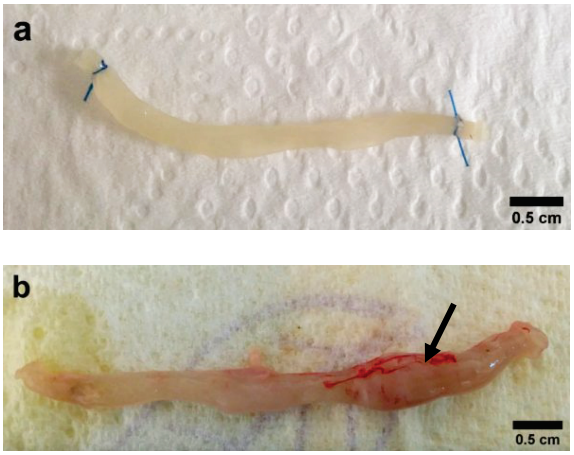


Figure 2.1: Photographs of excised colons

(a) Healthy murine colons of the untreated control group are uniform in shape with a slender profile. The sutures at the two ends were used during the sample preparation procedure. (b) Colons from the AOM/DSS group developed gross tumours (arrow) and deformed shape.

The mice treated according to the plan of section 2.1.1 were culled via cervical dislocation and their colons were excised. Immediately after excision, the colons were turned into 'sausages', as shown in Figure 2.1 following the procedure developed by Whittem *et al.* [210]. This process is described in detail in the supplementary SOP - S1, and summarised here: The faecal content was flushed out with phosphate buffered saline (PBS) using a syringe fitted with a plastic murine catheter. The lumen was then irrigated with 10 % neutral buffered formaldehyde (NBF), commercially known as formalin, and the two ends were sealed with surgical sutures. Sausages were immersed in 10 % NBF and left to fix for 24 hours. By fixing the tissues in the form of sausages, the risk of damaging the internal surface of the colonic lumen was minimised while the colons retained their natural shape. After fixation, the samples were stored in 70 % industrial methylated spirit (IMS), a lower purity equivalent to ethanol (EtOH), until further analysis.

Table 2.3: A standard tissue processing protocol run on a VIP5 Jr processor

Tissue dehydration and infiltration with paraffin wax was achieved by immersing the tissue samples in a sequence of reagents. The tissue was initially dehydrated with

increasing concentrations of IMS (steps 1 to 7), before clearene (steps 8 to 10) prepared it for paraffin wax infiltration (steps 11 to 14).

Step	Reagent	Temperature (°C)	Time (hours)
1	70 % IMS	37	1
2	80 % IMS	37	1
3	90 % IMS	37	1
4	100 % IMS	37	1
5	100 % IMS	37	1
6	100 % IMS	37	2
7	100 % IMS	37	2
8	Clearene	37	1.5
9	Clearene	37	1.5
10	Clearene	37	1.5
11	Paraffin wax	60	1.5
12	Paraffin wax	60	1
13	Paraffin wax	60	1
14	Paraffin wax	60	1

Each colon was dissected into pieces of length approximately equal to 10 mm while taking special care not to cut through tumours. The samples were then infiltrated with paraffin wax by applying the standard overnight tissue processing protocol of Table 2.3 on a Shandon Hypercenter tissue processor (VIP5 Jr, Thermo Fisher Scientific, Loughborough, UK), available at the histochemistry research unit (HRU), University of Southampton, UK.

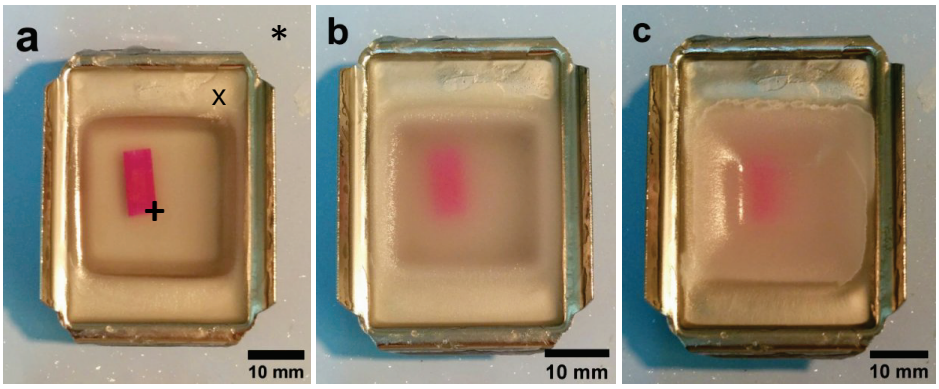


Figure 2.2: Wax-embedding of colonic samples using a piece of plastic drinking straw

The tissue was firstly submerged in the molten wax and the drinking straw (+) was slipped over it, while making sure that any air bubbles trapped in the wax were released by gently tapping the metallic troughs (x). The troughs were placed on an ice bed (*) and wax shrank as it solidified with time (a: 1 min, b: 15 min, c: 30 min).

To provide mechanical support and protect the tissue, the samples were then embedded in cylindrical wax blocks following a custom procedure: Standard metallic troughs were firstly filled with molten paraffin wax and the tissue was laid at the bottom. A 10 mm long and 5 mm in diameter piece of plastic drinking straw was slipped over the colon sample using plastic forceps to avoid introducing any metallic shreds (see Figure 2.2). The troughs were then tapped gently to release any air bubbles trapped in the wax, and left on an ice bed for the wax to solidify. Rapid cooling on the ice bed introduced rapid shrinkage of the paraffin wax, which cracked with a characteristic sound. To avoid this, the blocks should be left to solidify at room temperature, which would probably take several hours. Next, the blocks were removed from the troughs, the excess wax was trimmed away and the straw was removed with a scalpel, resulting in uniform cylindrical blocks with 5 mm diameter shown in Figure 2.3.

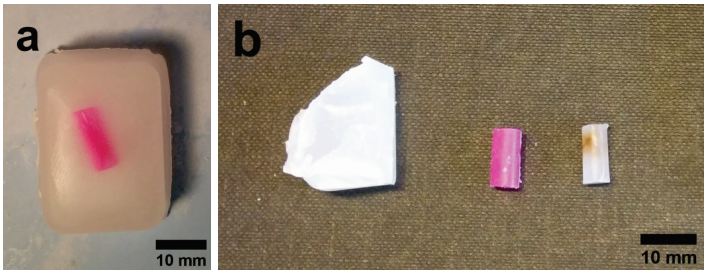


Figure 2.3: Preparation of cylindrical wax blocks

(a) A solid wax block right after it was extracted from the metallic trough. (b) Excess wax on the left, the drinking straw in the middle and the cylindrical FFPE tissue sample on the right were removed with a scalpel from the block shown in a.

The tissue samples may be re-embedded in paraffin wax blocks after µCT imaging, to allow for paraffin sectioning and immunohistochemistry, as shown in Figure 2.4. This was done via pouring a layer (about 2 mm) of molten wax into a standard metallic trough and placing the cylindrical wax block in an upright position. The metallic trough was then rapidly cooled on an ice bed to secure the sample in place. The trough was subsequently topped-up with molten wax and a standard plastic cassette was placed on top. After the entire block solidified on the ice bed, it was removed from the trough. In this form, the wax block is firmly attached to the plastic cassette that serves as a mounting bracket for paraffin sectioning (see section '2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining'). During the re-embedding process the temperature of the cylindrical wax block does not rise enough to re-melt the wax. As a result, the tissue within the cylindrical block does not lose mechanical support and its shape remains unchanged. This allows consistent sectioning of the re-embedded block, with no added practical issues (i.e. the cylindrical block does not detach from the surrounding wax during sectioning). As the orientation of the cylindrical block is consistent between µCT imaging and paraffin sectioning, this re-embedding method allows for co-registration of the two imaging modalities (see Figure 3.11).

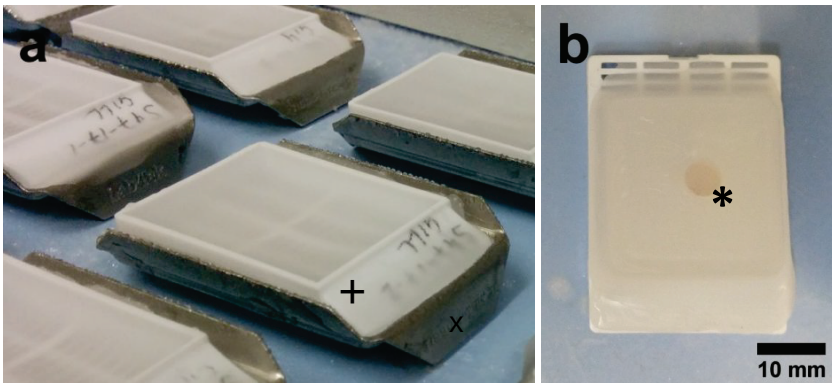


Figure 2.4: Re-embedding into a standard wax-block for paraffin sectioning

(a) Plastic cassettes (+) were placed on top of the metallic troughs (x) during the re-embedding process to create standard FFPE blocks for paraffin sectioning. (b) A standard FFPE block. The cylindrical FFPE block (*****) is supported and held on the plastic cassette by paraffin wax.

2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining

The sample preparation procedures outlined in this chapter aim to enable 3D imaging of murine colons in their natural shape while at the same time allow functional assays to be performed. The cylindrical FFPE tissue samples produced as described in section '2.1.2 Tissue sample acquisition and preparation' are suitable for X-ray µCT imaging. This section outlines a standard histology procedure to image 2D colonic sections using the light microscope after 3D µCT imaging.

After µCT imaging, the cylindrical blocks were re-embedded into standard wax blocks mounted on plastic cassettes as described in section '2.1.2 Tissue sample acquisition and preparation' and shown in Figure 2.4. In this form, the blocks are ready to be sectioned on the microtome, which yields slices of approximately 4 µm in thickness. In order to increase the block's stiffness and minimise sectioning artefacts (tearing of the tissue, streaks and cracks) [123], the blocks were placed on dry ice every few minutes before they were mounted on the microtome available at the biomedical imaging unit (BIU), University of Southampton, UK. The cut slices were subsequently picked with a pair of forceps and floated in a warm water bath, which melted the thin layer of wax. The floating tissue slices were then collected on glass histology slides (ClariTex Pro, Cellpath, Whales, UK) and left to dry at room temperature.

Table 2.4: The haematoxylin & eosin staining protocol on a Varistain slide stainer

The first two steps prepare the tissue so that it becomes transparent to the visible light. Steps 3 to 6 rehydrate the tissue so that it is miscible with the dye. Steps 7 to 10 are the H&E staining protocol, followed by a de-hydration process (steps 11 to 15).

Step	Reagent	Time (minutes)
1	Clearene	10
2	Clearene	10
3	100 % IMS	5
4	100 % IMS	5
5	70 % IMS	5
6	Distilled water	5
7	Mayer's haematoxylin	5
8	Running tap water	5
9	2 % eosin in 2 % calcium chloride	0.5
10	Running tap water	2
11	100 % IMS	2
12	100 % IMS	2
13	Clearene	2
14	Clearene	2
15	Clearene	2

The tissue sections captured on the glass slides were subsequently dyed with the standard haematoxylin and eosin (H&E) histology staining protocol shown in Table 2.4. The protocol ran on an automatic slide staining machine (Varistain 24 – 4, Thermo Fisher Scientific, Loughborough, UK). The first two steps (clearing) prepare the tissue so that it becomes transparent to the visible light [211]. Steps 3 to 6 re-hydrate the tissue so that it is miscible with the dye. Steps 7 to 10 are the H&E staining protocol which is followed by a de-hydration procedure (steps 11 to 15). The haematoxylin dye stains basophilic structures (e.g. the nuclei) purple. In contrast, eosin stains the eosinophilic structures (e.g. cytoplasm and muscles) pink. Thus, the H&E stain effectively creates colour contrast between the several tissues within a sample. When staining is finished, a drop of xylene-based mounting medium (PERTEX®, Pioneer research chemicals, Colchester, UK) was put over the slide to protect the tissue by mounting a plastic coverslip (vfm, Cellpath, Whales, UK). When the mounting medium dried, the slide was examined under the light microscope (Olympus VS110 light microscope) revealing the tissue components as shown in Figure 1.2b.

Paraffin sectioning, histology staining and imaging under the light microscope is a standard procedure that yields 2D images of exceedingly high quality (high resolution and high contrast). It is however a destructive technique, so it has to be applied after X-ray imaging in order to obtain images of the same tissue from both modalities.

2.2 Animal model results

The effectiveness of the animal model outlined in section ‘2.1.1 Azoxymethane and dextran sodium sulphate mouse model’ was quantified via three indices: the animal’s body weight loss over time, the colon length and colon weight/length ratio immediately after excision. The body weight of the animals treated with DSS (i.e. the DSS and AOM/DSS groups) follows a characteristic trend over time as a result of fluid loss and gain during inflammation and recovery phases [209, 212]. This characteristic pattern readily differentiates the DSS-treated groups from the AOM and untreated control groups. The presence of inflammation and tumours results in shorter and heavier colons due to increased mucosal thickness [212]. This was quantified by measuring the colon’s length and weight upon excision, and deriving the colonic weight/length ratio.

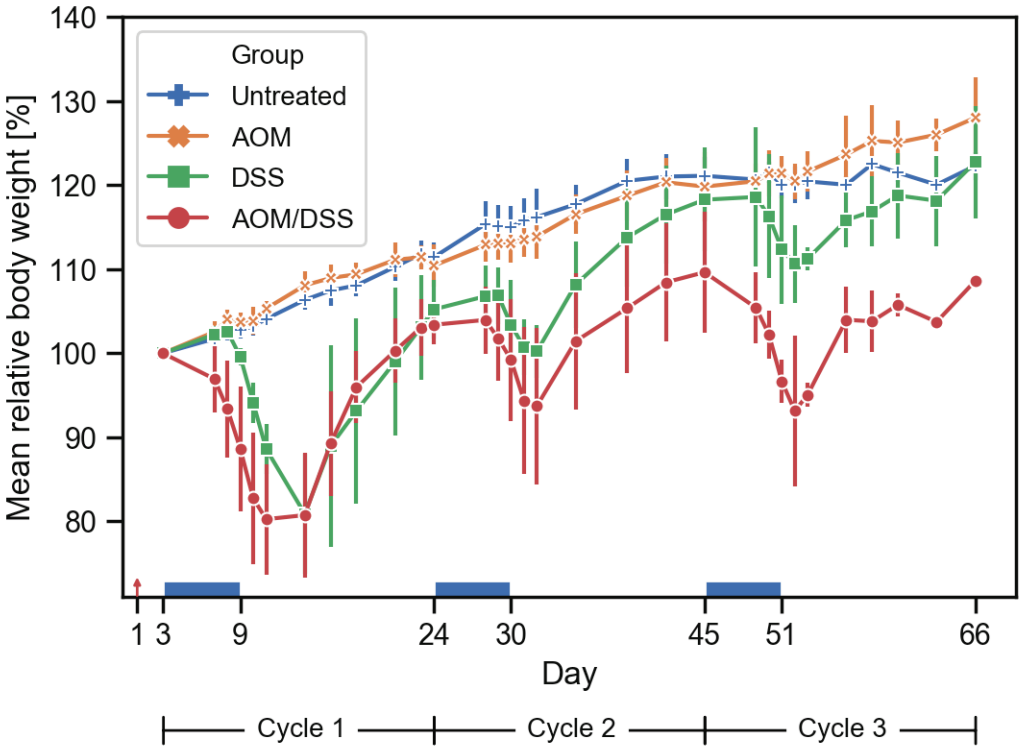


Figure 2.5: Mean relative body weight over time

Each point on the plot shows the arithmetic mean of each treatment group's relative weights (relative to the body weight on the first day of the first cycle, i.e. day 3 of the experiment) at the specific time point. Error bars correspond to the standard deviation of the sample distribution for each time point. The AOM injection on day 1 (arrow ↑) marks the start of the experiment. DSS was administered *ad libitum* in drinking water to the DSS and AOM/DSS groups during the first seven days of each of the three cycles, while sterile water was provided during the remaining days of the cycle. The initiation of DSS administration on days 3, 24 and 45 marks the start of the three cycles (cycle 1: day 3 to 24, cycle 2: day 24 to 45, cycle 3: day 45 to 66).

The mean relative body weight (relative to the body weight on the first day of the first cycle, i.e. day 3 of the experiment) is shown in Figure 2.5 for each group during the experiment. The animal's reference weight w_r is the weight of each mouse on the first day of DSS administration (day 3). The relative body weight $w_{\hat{n}}$ at time n corresponds to the mouse weight w_n at that time n as a fraction of its reference weight w_r and it is computed as follows: $w_{\hat{n}} = \frac{w_n}{w_r} \cdot 100\ \%$. Each point on the plot shows the arithmetic mean of each treatment group's relative body weights at the given time point. Note that culling at the end of each experimental cycle resulted in reduction of mouse numbers and thus increased statistical variability in the measurements. This is apparent in the ever-changing standard deviations of the untreated control group that begins with averages

of 6 in the first cycle, proceeds with averages of 4 and concludes with averages of 2 in cycles 2 and 3, respectively. Therefore, although the mean relative body weight of the untreated control group was relatively stable, the variability in the measurements is increased in cycle 2 and the curve displays an unexpected drop in cycle 3, as a result of insufficient animal numbers. Apart from the increased variability due to animal number reduction, the two DSS-treated groups display an increased variability from as early as cycle 1 onwards. This originates in differences in DSS susceptibility for each individual mouse. Firstly, the DSS dose for individual mice was not strictly controlled (provided *ad libitum* through a common drinking bottle for all mice in each cage). Secondly, the inflammatory response to DSS of each individual depends on a number of factors including the microflora of the colon, environmental conditions (housing) and eating habits [99, 110, 209]. As the 18 animals of the AOM/DSS group were housed in three different cages, these factors are also relevant and contribute to increased variability in the animal body weight measurements.

The AOM/DSS group exhibited a dramatic weight loss due to DSS administration (see Figure 2.5) which agrees with observations of others [110, 209, 212]. Immediate weight loss was observed upon DSS administration in cycle 1 (acute response) and persisted for about two days after DSS withdrawal. Conversely, the animals developed resistance to the DSS during cycle 2, where weight loss was not observed until day 28 (four days after DSS administration began). Weight gain (healing) during cycle 2 was, however, slower compared to cycle 1, which is a sign of animal health decline and fatigued immune system. This resulted in virtually no resistance to the DSS administration during cycle 3, thus immediate weight loss from day 45 onwards. The presence of inflammation and the formation of tumours over the duration of treatment for the AOM/DSS group resulted in difficulty to pass faeces and thus a loss of appetite. This resulted in a larger weight loss for the AOM/DSS group compared to the DSS group, beginning from cycle 2 and peaking during cycle 3. Out of the 18 animals in the AOM/DSS group, 5 had to be euthanized early (two on day 10, one on day 11, one on day 14 and one on day 52) due to bad welfare, as scored via the animal welfare scoring system detailed in supplementary SOP - S1. These animals were withdrawn from the DSS when critical weight loss was observed (>17 %) and closely monitored for health improvement until culling via schedule 1 was carried out. Upon investigation of the reasons driving this aggressive response of the AOM/DSS group to DSS, it was found out that regular cage cleaning by the biomedical research facility (BRF) staff occurred during the time of the daily welfare scoring and animal weighing. As a result, the animal cages were not regularly cleaned during cycle 1, which led to rapid health deterioration (exaggerated inflammatory response to DSS due to environmental factors, i.e. dirty housing [99, 110, 209]). Once this issue was identified, the cages were being replaced daily after the welfare scoring and animal weighing procedure during

cycle 2 and cycle 3. The samples from the 5 euthanized mice were excluded from the analysis and the presented results. Therefore the measurements for the AOM/DSS group include 6 mice in cycle 1, 5 mice in cycle 2, and 2 mice in cycle 3.

The DSS group initially exhibited similar weight loss pattern to the AOM/DSS group, which was later clearly differentiated during cycles 2 and 3. During cycle 1, mice in the DSS group began losing weight around 6 days after DSS administration began. This delay in weight loss initiation compared to the AOM/DSS group is a sign of the effect of AOM to the AOM/DSS group, which being a procarcinogen, facilitates the development of CRC. In contrast, the DSS group was only affected by the inflammatory action of DSS and exhibited greater resistance before it started losing weight. Weight gain after DSS withdrawal in cycle 1 followed the same rate as for the AOM/DSS group with a trajectory that, when projected in time, would reach the untreated control group by day 32. The weight gain was however stopped early with DSS administration during cycle 2. As with the AOM/DSS group, the DSS group developed resistance to DSS during the second cycle with only 7 % weight loss in cycle 2 compared to 20 % in cycle 1. The resistance to DSS was also exhibited with the complete weight regain over the healing phase of cycle 2, which lead the DSS group weight to reach the untreated control group by day 48. Another 7 % to 8 % weight was lost during cycle 3 that was regained once more during the final healing phase in cycle 3, leading the DSS group weight to reach the untreated control group by day 66.

The weight gain of the untreated and the AOM groups follow a similar pattern, increasing with a constant rate of 10 % every 20 days. The constant weight gain is normal for mice fed *ad libitum* [212]. It appears that the single AOM injection did not have a distinctive effect on the weight gain or animal health in general, which is in agreement with the observations of others [213].

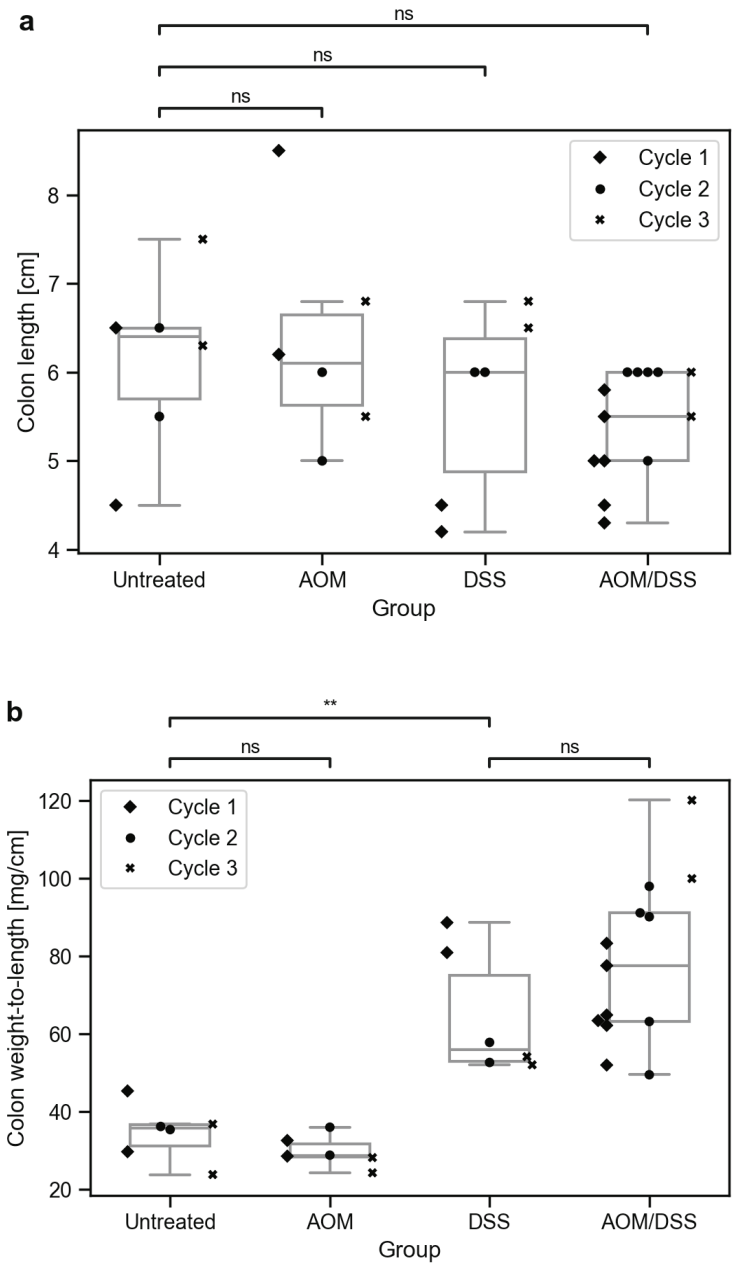


Figure 2.6: Colon length and colon weight-to-length ratio at the time of excision

The degree of inflammation was scored using the colon length (a) and colon weight-to-length ratio (b) measured at the time of excision. Each marker corresponds to one mouse colon, with diamond (◆), disk (●) and cross (✕) representing cycles 1, 2 and 3, respectively. The box-and-whiskers plots in the background represent the statistical distribution of all samples including all cycles for each group. The middle horizontal line of the box shows the median value of the sampling distribution and the bottom and top vertical lines show the first (q_1) and third (q_3) quartiles, respectively. The inter-quartile range (IQR) is the difference between the two quartiles ($q_3 - q_1$). The length of the whiskers is equal to 1.5 times the IQR. The colon length (a) cannot

differentiate between the groups due to the large variability. Statistical comparison of the colon length distributions of the untreated control group with each of the rest of the groups with the *Kruskal-Wallis H-test* shows that the difference between their median values is not statistically significant (p-value > 0.05). The colon weight-to-length ratio (b) macroscopically evaluates the colonic wall thickness. Statistical comparison of the colon weight-to-length distributions with the *Kruskal-Wallis H-test* shows that the difference between the median values of the DSS and AOM/DSS groups is not statistically significant (p-value > 0.05), similarly the median values of the AOM and untreated groups is not statistically significant. The difference between the untreated and DSS group is statistically significant with a p-value <= 0.01.

The colon length and weight was measured just after the colons where excised and cleaned, and before they were turned into sausages (see section '2.1.2 Tissue sample acquisition and preparation'). The degree of inflammation for each mouse colon was scored using the colon length and colon weight-to-length ratio shown in Figure 2.6. Each marker corresponds to one mouse colon, with diamond (◆), disk (●) and cross (✕) representing cycles 1, 2 and 3, respectively. The box-and-whiskers plots in the background represent the statistical distribution of all samples including all cycles for each group. Figure 2.6a shows that there is a tendency for shorter colon lengths in the DSS and AOM/DSS groups compared to the untreated and the AOM groups, however the difference was not statistically significant (*Kruskal-Wallis H-test* between the untreated control group and each of the other groups yield p-values > 0.05). The large variability in the non-DSS-treated control groups (the untreated group range is equal to 3 cm and the AOM group range is equal to 3.5 cm) indicates that the colon length alone is not a good measure for comparison between the groups.

Instead, the colon length was used to normalise the colon weight and compute the colon weight-to-length ratio (see Figure 2.6b), which is a measure proportional to the tissue mass density (if the colon wall thickness is assumed constant). Thus, the colon weight-to-length ratio can be used to macroscopically evaluate the colonic wall thickness. Both AOM/DSS and DSS groups exhibited increased weight-to-length ratio compared to the AOM and untreated groups, which is a result of increased mucosal thickness due to the inflammatory response to DSS [212]. The DSS group exhibited distinct clustering according to the treatment cycle. This is because the inflammatory response to DSS is acute [209], thus directly related to the culling day. The AOM/DSS group on the other hand exhibited increased variability, which is related to the individual susceptibility to DSS of each mouse as discussed earlier.

As for accurate quantification of crypt budding it is essential to preserve the natural colonic shape, the employed tissue preparation protocol requires that the colons remain intact. Therefore, longitudinally cutting the colons was avoided which made it impossible to count the number of gross tumours. However, qualitative observations of the excised colons revealed tumours forming only on the AOM/DSS group at cycles 2 and 3 shown in Figure 2.1.

Taken together, the relative body weight profile, the colon weight-to-length ratio and the macroscopic evaluation of the excised colons suggest that CRC was successfully induced to the AOM/DSS group, inflammation varying in time was induced to the DSS control group, and no pathology was induced to the AOM and untreated control groups. After excision and quantification of the measures discussed above, the colons were processed as detailed in section '2.1.2 Tissue sample acquisition and preparation' in order to prepare samples appropriate for μCT imaging ('Chapter 3 Soft tissue X-ray μCT imaging').

Chapter 3 Soft tissue X-ray µCT imaging

In Chapter 1 the need for new tools for *in situ* 3D imaging of a large number of crypts has been identified. These tools should allow functional assays to be performed after imaging in order to support the study of crypt budding. A workflow comprised of in-line phase-contrast X-ray µCT imaging of FFPE tissue blocks followed by traditional histology has been introduced as a candidate solution to this problem. In Chapter 2, murine colons were induced with early-stage colorectal cancer and embedded in paraffin wax, producing FFPE murine colon samples ready to be imaged by µCT. This chapter deals with practical aspects of soft-tissue imaging, providing a step-by-step recipe on how to perform in-line phase-contrast X-ray µCT imaging of such samples. The tools developed here are applicable not only for imaging the particular samples produced in section '2.1.2 Tissue sample acquisition and preparation', but also to the more general case of phase-contrast imaging in synchrotrons or in the lab.

Firstly, in section '3.1 Soft tissue imaging in practice' a minimal set of technical requirements for in-line phase-contrast X-ray µCT imaging is derived and a flowchart is provided to perform this technique in a synchrotron or laboratory source. The flowchart is validated experimentally through a synchrotron experiment utilising a sample from the animal study of Chapter 2. Next, in section '3.3 Phase-contrast imaging of murine colons for the study of crypt budding' the developed flowchart is put in practice, and high-resolution µCT imaging is validated through one-to-one comparison with H&E histology (section '3.3.1 High-resolution synchrotron phase-contrast CT of murine colons'). Finally, in section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons', a µCT imaging protocol is developed to image all the samples produced in Chapter 2, and to use the reconstructed data for further study of crypt budding in the following chapters.

3.1 Soft tissue imaging in practice

Although X-ray phase-contrast imaging of soft tissues is gradually becoming more common, it generally requires a certain level of knowledge and/or expertise in engineering, material science, mathematics and physics. Thus, X-ray propagation-based phase-contrast imaging can quickly become too difficult to follow for readers with different academic backgrounds, particularly when the literature becomes too technical. The main concepts necessary to understand the physics behind this advanced imaging technique have been covered in section '1.4 Advanced imaging approaches in research'.

Yet, further technical requirements need to be satisfied such that the radiographies are properly captured by a digital detector, especially when it comes to phase-contrast and edge-enhancement imaging. These technical requirements are detailed here, and a step-by-step protocol design flowchart is presented to empower users to define optimal X-ray propagation-based phase-contrast imaging protocols at the desired spatial resolution, and for their specific samples.

3.1.1 Generic imager geometry

The generic geometry of an X-ray imager for the most common, indirect X-ray detection method through scintillators (see section '1.4.1 Computed tomography principles'), is shown in Figure 3.1. X-rays are generated by an X-ray source, they propagate through free space for a distance R_1 (source-to-object distance) until they reach the object or sample, where they interact with the matter it is made of. X-ray refraction takes place at the object boundary and internal structural features. X-rays then propagate through free space for a distance R_2 (X-ray propagation distance [173, 200]) until they reach the scintillator, where they are converted to visible light.

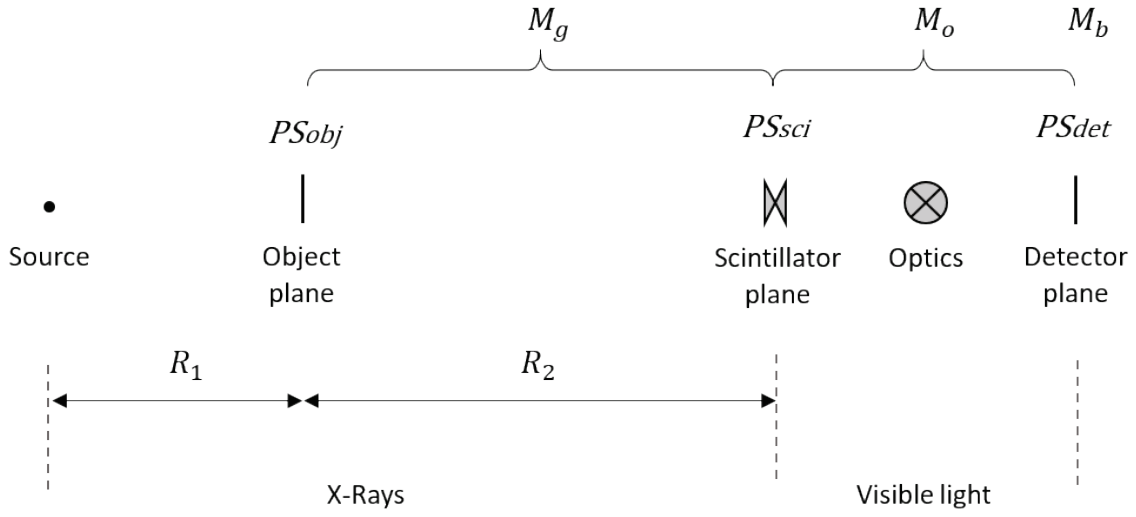


Figure 3.1: The generic geometry of an X-ray imager

X-rays are generated by a source and propagate through free space for a distance R_1 (source-to-object distance), until they reach the object. X-ray refraction takes place at the object features, and X-rays then propagate through free space for a distance R_2 (object-to-scintillator distance) until they reach the scintillator, where they are converted into visible light. Geometric magnification M_g happens between the object and the scintillator, optical M_o magnification is achieved by lenses between the scintillator and the detector, and binning M_b may be applied on the detector. As a result, the pixel size on the detector PS_{det} is reduced to a smaller pixel size on the scintillator PS_{sci}, and an even smaller pixel size on the object PS_{obj}.

Sometimes, R_2 is referred to as object-to-detector distance, which is only correct for systems with their scintillator directly attached to the detector. Otherwise, for X-ray imaging systems with optical magnification and spatial separation of the scintillator from the detector, R_2 denotes the source-to-scintillator distance. It is important to note that for X-ray systems with optical magnification, the expression 'X-ray propagation' only makes sense for the path where the light waves are X-rays, and not for the visible light path between the scintillator and the detector.

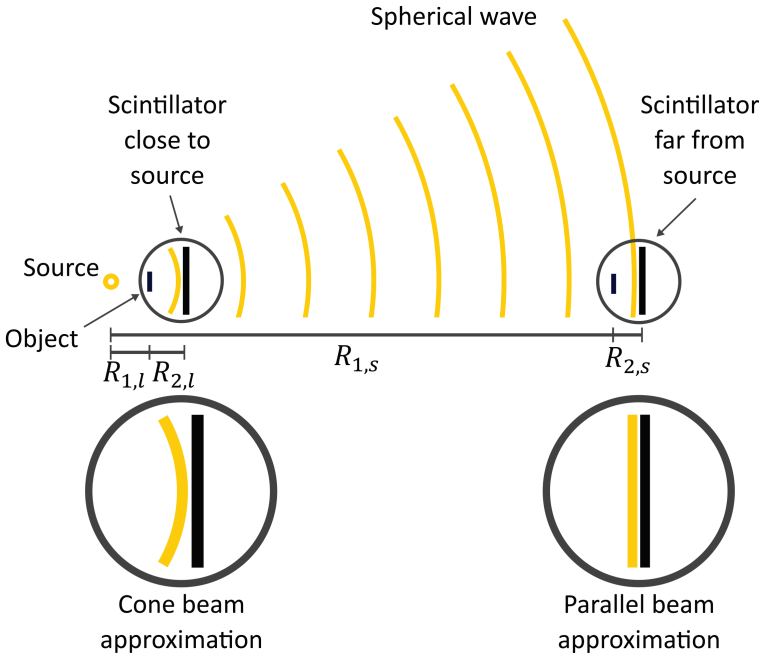


Figure 3.2: The cone and parallel beam approximations

In synchrotrons the source-to-object distance $R_1 = R_{1s}$ is significantly larger than the object-to-scintillator distance $R_2 = R_{2s}$ (or $R_{1s} \gg R_{2s}$), resulting in geometrical magnification M_g approximately equal to one. Thus, the spherical wave originating in the source can be approximated by a plane wave or equivalently the X-ray beam can be considered being parallel. In lab X-ray sources $R_1 = R_{1l}$ and $R_2 = R_{2l}$ are comparable to each other resulting in geometric magnification M_g greater than one. The scintillator in this case receives a section of the spherical X-ray wave, and hence the X-ray beam appears as conical. Mathematically, the X-ray wave in the cone geometry is modelled by a spherical wave which strictly holds when $R_{2l} \gg R_{1l}$.

The generic X-ray imaging geometry comprises two stages of magnification, the geometric magnification M_g, and the optical magnification M_o. The geometric magnification happens between the object and the scintillator as a consequence of the fact that a spherical wave expands from a point source through space:

$$M_g = \frac{R_1+R_2}{R_1}. \qquad (3.1)$$

This results in a feature of size α at the object plane to be depicted with size $M_g\alpha$ at the scintillator plane, hence M_g times larger. This phenomenon happens for both laboratory and synchrotron imagers, with the geometrical magnification M_g being approximately equal to one for synchrotron setups because there, R_1 is significantly larger than R_2 ($R_1 \gg R_2$). In this case, the spherical wave that originates in the X-ray source, can be approximated by a plane wave that can be considered being parallel to the scintillator plane (see Figure 3.2). For laboratory X-ray sources on the other hand, the geometric magnification is greater than one since R_1 and R_2 are comparable. The scintillator in this case receives a section of the spherical X-ray wave, and hence the X-ray beam appears as conical (cone beam geometry).

An important consequence of the geometric magnification provided by the cone beam geometry is the effect it has on the propagation distance during the phase retrieval step. As discussed in section '1.4.5 X-ray phase retrieval', the geometric magnification M_g downscales the propagation distance in Equation (1.23), which leads to the term effective propagation distance ($z_{eff} = \frac{z}{M_g}$) that is frequently used [171, 189] for convenience. In essence the effective propagation distance stems from an argument stating that both the planar and spherical waves can be considered as a planar wave with the propagation distance z and the term x (see Equation (1.7)) downscaled by M_g, as discussed in detail by Pogany *et al.* [180]. As a result, in the planar wave approximation (where $R_1 \gg R_2$), the geometric magnification becomes $M_g \approx 1$, leading to an effective propagation distance $z_{eff} = \frac{z}{M_g} \approx \frac{z}{1} = \frac{R_2}{1} = R_2$. Conversely, in the spherical wave approximation (where $R_2 \gg R_1$) the geometric magnification becomes $M_g \approx \frac{R_2}{R_1}$, leading to an effective propagation distance of $z_{eff} = \frac{z}{M_g} \approx \frac{R_1 z}{R_2} = \frac{R_1 R_2}{R_2} = R_1$. This results in an optimal feature of size α (feature size where the phase-contrast term is maximised) given by adapting Equation (1.21) accordingly:

$$\alpha \approx \sqrt{2\,\lambda\,R_1}, \qquad (3.2)$$

for cone beam geometry (typical for laboratory systems) and

$$\alpha \approx \sqrt{2\,\lambda\,R_2}. \qquad (3.3)$$

for parallel beam geometry (synchrotron setups). The second stage of magnification is performed by optical lenses at an optical magnification M_o. A last factor that affects the feature's size is demagnification due to the binning factor M_b. Binning is a process of combining square areas of pixels and replacing them with their average value (performed either on the detector or through digital processing after imaging), thus increasing the photon count that equivalently falls onto

each pixel while reducing the noise. This results in increased signal-to-noise ratio (SNR) and hence better image quality. Combining all sources of magnification yields the total magnification:

$$M = \frac{M_g M_o}{M_b}, \quad (3.4)$$

In brief, the pixel size at the detector plane PS_{det} (a fixed technical characteristic otherwise referred to as the detector's pitch) is reduced to a smaller pixel size at the scintillator plane PS_{sci} (due to the optical magnification M_o), and an even smaller pixel size at the object plane PS_{obj} (due to the geometric magnification M_g) as follows:

$$PS_{sci} = PS_{det} \frac{M_b}{M_o}, \quad (3.5)$$

$$PS_{obj} = PS_{sci} \frac{1}{M_g} = PS_{det} \frac{M_b}{M_o M_g} = \frac{PS_{det}}{M}. \quad (3.6)$$

In other words, features in the object are magnified by a factor M at the detector plane: $PS_{det} = M \cdot PS_{obj}$. This allows detecting sample features that are many (~M) times smaller than the actual pitch of the detector. The presence of multiple magnification stages in the imager geometry has a significant implication. Dimensions of features are not the same across the different planes on the beam path. It is thus important to quote the plane upon which a given feature size is defined in order to avoid confusion. As X-ray refraction and propagation only makes sense when dealing with X-rays, the size of Fresnel fringes, built up when the X-rays travel for the propagation distance z, is defined at the scintillator plane. This gives rise to the alternative definition of the scintillator pixel size PS_{sci} as the effective pixel size. Equation (3.4) shows that the geometric and optical magnification factors increase the total magnification, whilst binning has the opposite effect. One must thus be cautious when using binning during phase-contrast imaging, as it increases the effective pixel size (see Equation (3.5)), making it more difficult to record Fresnel fringes. As discussed in the following section, it is important to fulfil a set of minimal imaging requirements in order to perform phase-contrast imaging and properly capture Fresnel fringes. Following phase retrieval and CT reconstruction, the images may then be downscaled (e.g. by binning) to finally achieve the required SNR and/or pixel size.

3.1.2 Minimal requirements for phase-contrast imaging

In order to perform phase-contrast imaging, not only the physical phenomenon of Fresnel fringe development must occur, it must also be technically possible to record these fringes. Therefore the requirements emanating from section '1.4 Advanced imaging approaches in research' that are necessary to perform a successful phase-contrast imaging experiment are collected here.

As mentioned in section '1.4.4 Phase-contrast X-ray imaging', an infinitesimally sharp edge would in practice be recorded in the radiography as a (smooth) curve with a width equal to $\frac{\alpha}{2}$. This can be considered as the spatial resolution of the system, which is applicable to both absorption-contrast and phase-contrast imaging. As the smallest discernible edge would have a width equal to $\frac{\alpha}{2}$, it follows that the smallest feature that can be recorded would have a width equal to α (see also Figure 3.6), called here feature size (size of the smallest feature of interest or smallest discernible feature as defined at the object plane).

Note that the analysis presented here disregards the effects of the detector's imperfections, which may be captured by the system's point spread function (PSF). Hence, the requirements derived here (particularly Req 5) are more stringent than practically necessary to capture the Fresnel fringe width.

Req 1. Coherent beam (Equations (1.16) and (1.17) combined) [165]:

$$R_1 > \frac{\pi\alpha\sigma}{\lambda}, \qquad (3.7)$$

where R_1 is the source-to-object distance, α the feature size, σ the source size and λ the wavelength. Equation (3.7) describes the 'coherence resolution limit' [180], i.e. the minimum feature size α allowed before the beam becomes incoherent. Note that this refers to the resolution limit, and not all of the existing features in the image. In other words, larger features will still be visible as "it is not necessary for the whole object or field of view to be coherently illuminated" [180].

Req 2. Imaging far from absorption edges:

The X-ray energy needs to be selected such that imaging is performed far from the absorption edges of the material that the sample is made of (see section '1.4.2 X-ray interaction with matter and image formation' and Figure 1.19). Then, the X-ray wavelength is:

$$\lambda = \frac{hc}{E_{eff}}. \qquad (3.8)$$

where E_{eff} is the effective energy, i.e. a single-value weighted mean over the entire energy spectrum of the X-ray beam as discussed in section '1.4.4 Phase-contrast X-ray imaging'. For synchrotron sources this can be obtained through discussion with the beamline scientist. For laboratory X-ray sources, when technical details cannot be acquired by the manufacturer, the

effective energy may be experimentally estimated [171, 189] or computed as the weighted spectral average from simulated values [214-217]. A rule-of-thumb that can be used to estimate the effective energy E_{eff}, given the source tube voltage, is that the effective energy ranges from ¼ to ⅓ of the applied tube voltage, depending on the target material [171, 218].

Req 3. Mathematical model assumptions:

3.1 The 'near-field' ('Fresnel regime') condition [165]:

3.1.i For cone beam (typical for laboratory systems):

$$R_1 \leq \frac{\alpha^2}{\lambda}, \quad (3.9)$$

3.1.ii For parallel beam (synchrotron setups):

$$R_2 \leq \frac{\alpha^2}{\lambda}. \quad (3.10)$$

3.2.i For cone beam (typical for laboratory systems) the spherical wave approximation [180]:

$$R_2 \gg R_1, \quad (3.11)$$

3.2.ii For parallel beam (synchrotron setups) the plane wave approximation [180]:

$$R_1 \gg R_2. \quad (3.12)$$

Req 4. The phase-contrast term must be maximised [180]:

4.i For cone beam (typical for laboratory systems) rearranging Equation (3.2) yields:

$$R_1 \approx \frac{\alpha^2}{2\lambda}, \quad (3.13)$$

4.ii For parallel beam (synchrotron) rearranging Equation (3.3) yields:

$$R_2 \approx \frac{\alpha^2}{2\lambda}. \quad (3.14)$$

Req 5. Shannon's sampling criterion must hold:

5.1 Features of interest must be properly sampled [194]:

$$PS_{obj} < \frac{\alpha}{2}, \quad (3.15)$$

Combining equations (3.6) and (3.15) yields:

$$M_o > \frac{2M_b}{\alpha M_g} PS_{det}. \quad (3.16)$$

5.2 The Fresnel fringes must be properly sampled [165]:

$$PS_{sci} \leq \frac{x_1}{2}, \quad (3.17)$$

where x_1 is the width of the first (and largest) Fresnel fringe at the scintillator plane (see Figure 1.24). Combining Equations (1.21), (1.22), (3.5) and (3.17) yields:

$$R_2 \geq \frac{8\left(PS_{det}\frac{M_b}{M_o}\right)^2}{\lambda}. \quad (3.18)$$

The requirements outlined above, while being fairly loose for the synchrotron case, are exceedingly stringent for the case of laboratory imagers. Excessively increasing the source-to-object distance R_1 to satisfy Req 1 is difficult in laboratory imagers due to limitations in the available space and photon flux. On the other hand, decreasing the source-to-object distance R_1, in order to increase the spatial resolution by reducing α via Req 4, constrains the pixel size at the object plane PS_{obj} as per Req 5. On X-ray imagers that comprise optical magnification capabilities, an appropriate choice of the optical magnification provides extra versatility by decoupling R_1, R_2 and the total magnification M (Equations (3.1) and (3.4)), making it easier to satisfy Req 5. However, low photon count, and hence low SNR, is typically a constraining factor, especially for highly magnifying microscope objectives (e.g. 20× or 40×), leading to the need for long exposure times, and thus long scanning times.

Req 5 ensures that the Shannon's sampling theorem holds in order to properly record the theoretical width of the Fresnel fringes. As the effect of the detector's imperfections (PSF) has been disregarded here, which would 'widen' the recorded Fresnel fringes, Req 5 is stricter than what is actually necessary in practice. As a result, Equation (3.18) tends to overestimate the minimum X-ray propagation distance R_2. Regardless, a close look at Req 5.2 (Equation (3.18)) underlines the pivotal role of the detector pixel size as it is being raised to the power of two before it serves as a lower threshold for the X-ray propagation distance R_2 needed to properly sample the Fresnel fringes. In a laboratory system such as the ZEISS Xradia Versa scanners, an

effective energy of 14.0 keV (applied tube voltage ~42 keVp), a detector's pixel size of 13.5 µm and an optical magnification objective of 20× would require a propagation distance R_2 of at least 41 mm. Although imaging with this setup is actually feasible, the exceedingly low photon count due to the high-magnification objective and the long propagation distance requires long exposures and exceedingly long scanning times for good image quality data (12-24 h or more) [171].

3.1.3 Protocol design flow chart for X-ray propagation-based phase-contrast imaging

To concurrently satisfy the requirements of section '3.1.2 Minimal requirements for phase-contrast imaging' one must solve an optimisation problem (for maximal spatial resolution, i.e. minimal feature size). This is particularly difficult, especially when the user has control over several (optical) microscope objectives, the choice of the binning factor and a free selection of the X-ray propagation distance, as it is typically the case with synchrotron setups. Optimal protocol design is even more difficult in the laboratory case due to further and excessive constraints such as small X-ray propagation distances and low photon counts that prohibit the use of large magnification objectives. In order to facilitate the design of X-ray propagation-based phase-contrast imaging protocols that satisfy the minimal requirements discussed in the previous section, a flowchart for each of the two cases (synchrotron setup and laboratory system) is presented here. They implement an iterative solution of the optimisation problem at hand, which begins with the ideal protocol (one that yields the best possible spatial resolution) and incrementally relaxes its conditions (lowers the spatial resolution) in each iteration step, until all (minimal) requirements are satisfied.

A closer look at Req 4 shows that for a minimal feature size α, one has to select $R_1 = 0$ for cone beam, or $R_2 = 0$ for parallel beam. This yields the theoretical optimal spatial resolution (imaging infinitely small features $\alpha = 0$), which is not practically achievable due to experimental limitations such as the finite pixel size of the detector or an inadequate photon count due to excessive optical magnifications (Req 5). Therefore, in order to design an optimal imaging protocol, one has to select the experimental settings such that the smallest technically achievable feature size α can be recorded while respecting the assumptions for the X-ray source, the imaged object and the mathematical modelling of the X-ray propagation in the near-field regime (Req 1 – Req 3, respectively). One way to achieve this is to begin with the theoretical optimal feature size and gradually increase R_1 and/or R_2 until all requirements are satisfied.

The nominal feature size α is the size of the smallest feature that can be practically recorded by the imaging setup (at the object plane in the radiographies and before phase retrieval) with a

particular set of experimental settings. In other words, the width of a practically recorded step-like feature is equal to $\frac{\alpha}{2}$ (spatial resolution of the system). The desired feature size α_{des} is defined as the smallest feature size of interest in the radiographies at the object plane. Therefore, the nominal feature size α needs to be smaller or equal to α_{des} so that the features of interest can be properly recorded. When phase retrieval is applied, the spatial resolution of the phase-retrieved radiographies deteriorates due to the smoothing imposed by the phase-retrieval algorithm (see section '1.4.5 X-ray phase retrieval'). Therefore, the desired smallest feature size of interest in the radiographies at the object plane and after phase-retrieval is defined as $\alpha_{p,des}$. This is identical to the desired smallest feature size of interest in the reconstructed CT slices with prior phase retrieval. Note that, as $\alpha_{p,des}$ is defined for the final output of the phase-contrast imaging experiment (on the reconstructed CT slices with prior phase retrieval), it is a convenient design parameter, i.e. it is a parameter defined by the user and it is determined by the size of the features of interest in the sample being imaged.

The ratio between the phase retrieved and the non-phase-retrieved desired smallest feature size $\frac{\alpha_{p,des}}{\alpha_{des}}$ depends on the phase retrieval filter's cut-off frequency (section '1.4.5 X-ray phase retrieval'), with a smaller cut-off frequency (stronger smoothing) resulting in larger ratio $\frac{\alpha_{p,des}}{\alpha_{des}}$. The ratio was estimated experimentally for the synchrotron case and for $\frac{\delta}{\beta} = 218$, $E = 27$ keV (see section '3.2 Experimental validation of the protocol design flow chart') and found that it was not larger than three. Therefore setting $\frac{\alpha_{p,des}}{\alpha_{des}} = 3$ results in an additional condition for the protocol design flowcharts below. This relates the nominal feature size α and the user-defined desired smallest feature size of interest for the reconstructed CT slices with prior phase retrieval $\alpha_{p,des}$, with the condition that $\alpha < \frac{\alpha_{p,des}}{3}$.

Protocol design flowchart for parallel beam X-ray propagation-based phase-contrast imaging (Figure 3.3):

- **i.** Assuming Req 1 holds (coherent beam), pick the X-ray beam energy following Req 2 and calculate the wavelength using Equation (3.8). Set the geometric magnification $M_g = 1$ and select a desired feature size $\alpha_{p,des}$ equal to the smallest feature size of interest at the object plane (feature size for reconstructed and phase-retrieved CT slices). Select the smallest available optical objective lens (M_o) and binning (avoid binning if possible, i.e. set $M_b = 1$).
- **ii.** Increase the optical magnification M_o by selecting an objective lens at a higher magnification level and/or decrease the binning factor M_b (step x.) until the following

inequality holds: $\frac{M_b}{M_o} \leq \frac{\alpha_{p,des}}{6PS_{det}}$. This emerges from Req 5.1 when imposing the experimentally derived condition $\alpha < \frac{\alpha_{p,des}}{3}$.

iii. Select a minimum/starting propagation distance $R_{2,min}$, e.g. the closest practically attainable object-to-scintillator distance.

iv. Calculate the nominal feature size α resulting from the currently selected energy and propagation distance R_2 using Equation (3.3).

v. If Req 5.1 (Equation(3.16)) does not hold, either:
Proceed to step x. This will result in a smaller pixel size and better spatial resolution while sacrificing photon count, thus longer exposure will be required to maintain the SNR.
Or:
Proceed to step ix. This will result in an unchanged photon count (X-ray flux), while sacrificing spatial resolution due to increased Fresnel fringe width x_1.

vi. If Req 5.2 (Equation (3.18)) does not hold, proceed to step ix.

vii. If the nominal feature size α calculated at step iv is not at least three times smaller than the desired spatial resolution $\alpha_{p,des}$, proceed to step x.

viii. Reaching this step means that the currently selected settings (M_o, M_b, R_2) are optimal for phase-contrast imaging and yield a nominal spatial resolution equal to α. This would be the nominal spatial resolution of the reconstructed CT slices in the case that the phase was not retrieved. The feature size of the phase-retrieved and reconstructed CT slices α_p will be increased ($\alpha_p > \alpha$) due to the low-pass filtering effect of the phase retrieval, but not worse than the desired feature size defined by the user ($\alpha_{p,des} < \alpha_p$).

ix. Increase the propagation distance R_2 by a fixed distance $R_{2,step}$ and continue from step iv.

x. Increase the optical magnification M_o and/or decrease the binning factor M_b and continue from step ii.

Note that in this protocol design flowchart for parallel beam, the ratio between the phase-retrieved and non-phase-retrieved radiographies was experimentally estimated $\frac{\alpha_{p,des}}{\alpha} = \overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)} < 3$, which resulted in a denominator of $2\overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)}PS_{det} = 6PS_{det}$ in ii. and $\alpha < \frac{\alpha_{p,des}}{\overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)}}$ in vii. These choices may be adjusted if a better estimation of the ratio $\frac{\alpha_{p,des}}{\alpha}$ becomes available.

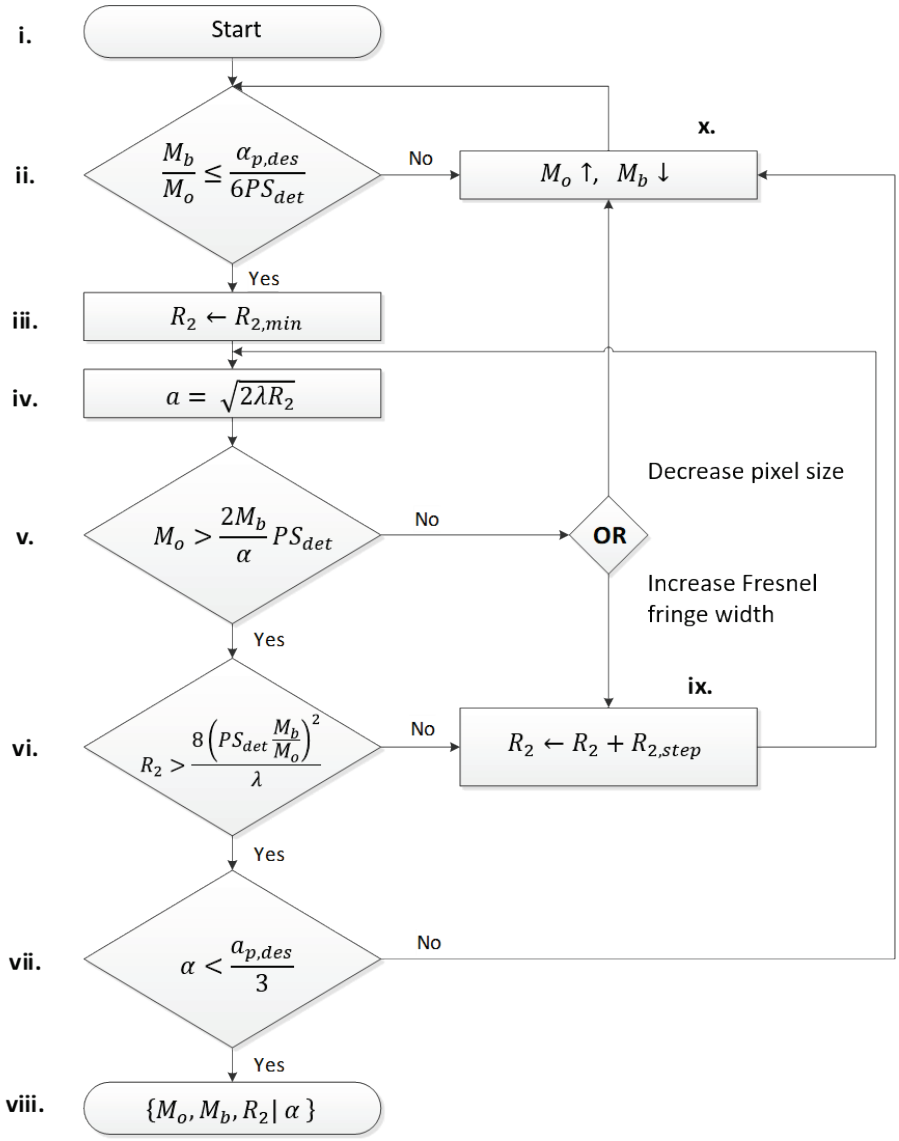


Figure 3.3: X-ray propagation-based phase-contrast imaging protocol design flowchart for parallel beam setups (synchrotron)

A generic flowchart that enables the user to define an optimal X-ray propagation-based phase-contrast imaging protocol at the desired spatial resolution and for their specific samples customised for synchrotron X-ray sources ($M_g = 1$).

Protocol design flowchart for cone beam X-ray propagation-based phase-contrast imaging (see Figure 3.4):

i. Assuming Req 1 holds (coherent beam), pick the X-ray beam energy following Req 2 and calculate the wavelength using Equation (3.8). Select a desired feature size $\alpha_{p,des}$ equal to the smallest feature size of interest at the object plane (feature size for reconstructed and phase-retrieved CT slices). Select the smallest available optical objective lens (M_o) and binning (avoid binning if possible, i.e. set $M_b = 1$).

ii. Increase the optical magnification M_o by selecting an objective lens at a higher magnification level and/or decrease the binning factor M_b (step x.) until the following inequality holds: $\frac{M_b}{M_o} \leq \frac{\alpha_{p,des}}{6PS_{det}}$. This emerges from Req 5.1 when imposing the experimentally derived condition $\alpha < \frac{\alpha_{p,des}}{3}$.

iii. Select a minimum/starting propagation distance $R_{2,min}$, e.g. the closest practically attainable object-to-scintillator distance.

iv. Select R_1 following Req 3.2.i. (Equation (3.11)), e.g. $R_1 = \frac{R_2}{2}$. Calculate the nominal feature size α resulting from the currently selected energy and propagation distance R_2 using Equation (3.2). Calculate the geometric magnification M_g using Equation (3.1).

v. If Req 5.1 (Equation(3.16)) does not hold, either:
Proceed to step x. This will result in smaller pixel size and better spatial resolution while sacrificing photon count, thus longer exposure will be required to maintain the SNR.
Or:
Proceed to step ix. This will result in both decreased photon count and worse spatial resolution due to increased Fresnel fringe width x_1.

vi. If Req 5.2 (Equation (3.18)) does not hold, proceed to step ix.

vii. If the nominal feature size α calculated at step iv is not at least three times smaller than the desired spatial resolution $\alpha_{p,des}$, proceed to step x.

viii. Reaching this step means that the currently selected settings (M_o, M_b, R_1, R_2) are optimal for phase-contrast imaging and yield a nominal spatial resolution equal to α. This would be the nominal spatial resolution of the reconstructed CT slices in the case that the phase was not retrieved. The feature size of the phase-retrieved and reconstructed CT slices α_p will be increased ($\alpha_p > \alpha$) due to the low-pass filtering effect of the phase retrieval, but not worse than the desired spatial resolution defined by the user ($\alpha_{p,des} < \alpha_p$).

ix. Increase distance R_2 by a fixed distance $R_{2,step}$ and continue from step iv.

x. Increase the optical magnification objective M_o and/or decrease the binning factor M_b and continue from step ii.

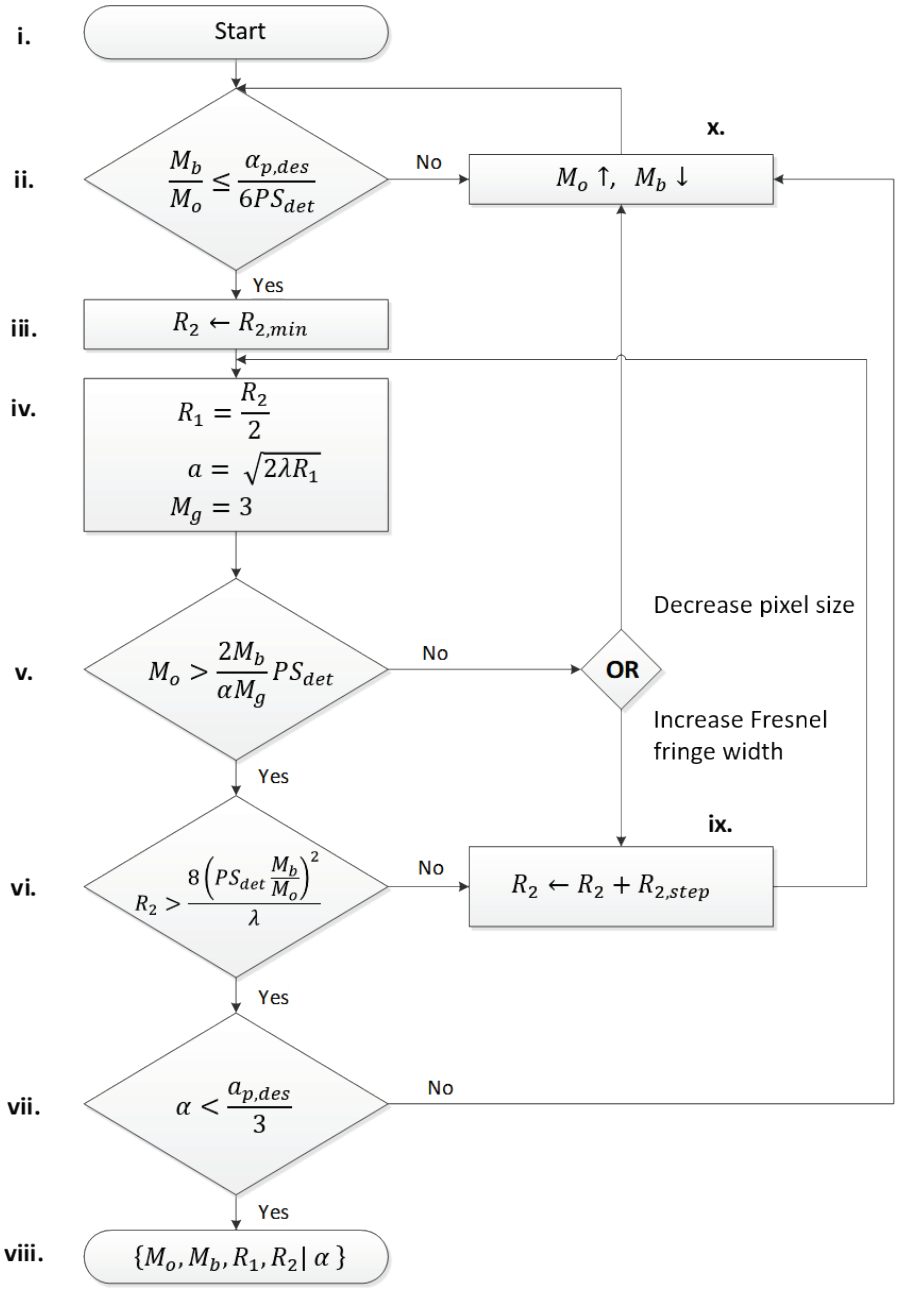


Figure 3.4: X-ray propagation-based phase-contrast imaging protocol design flowchart for cone beam setups (laboratory)

A generic flowchart that enables the user to define an optimal X-ray propagation-based phase-contrast imaging protocol at the desired spatial resolution and for their specific samples, customised for laboratory X-ray sources. In this version the source-to-object distance was set to $R_1 = \frac{R_2}{2}$ to satisfy Req 3.2.i: $R_2 \gg R_1$ resulting in a fixed geometric magnification $M_g = 3$, yet different choices could be made.

Note that in this protocol design flowchart for cone beam, the source-to-object distance was set to $R_1 = \frac{R_2}{2}$ to satisfy Req 3.2.i: $R_2 \gg R_1$ resulting in a fixed geometric magnification $M_g = 3$, yet different choices could be made. Also, the ratio between the phase-retrieved and non-phase-

retrieved radiographies was experimentally estimated $\frac{\alpha_{p,des}}{\alpha} = \overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)} < 3$, which resulted in a denominator of $2\overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)}PS_{det} = 6PS_{det}$ in ii. and $\alpha < \frac{\alpha_{p,des}}{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)}$ in vii. These choices may be adjusted if a better estimation of the ratio $\frac{\alpha_{p,des}}{\alpha}$ is available.

To understand the process of satisfying the constraints for optimal phase-contrast imaging, one has to examine the two underlying processes that are affected by the different choices of the experimental settings. Figure 3.5 shows how a profile containing Fresnel fringes (grayscale values of a radiography, selected across a straight line) may be captured by a digital detector.

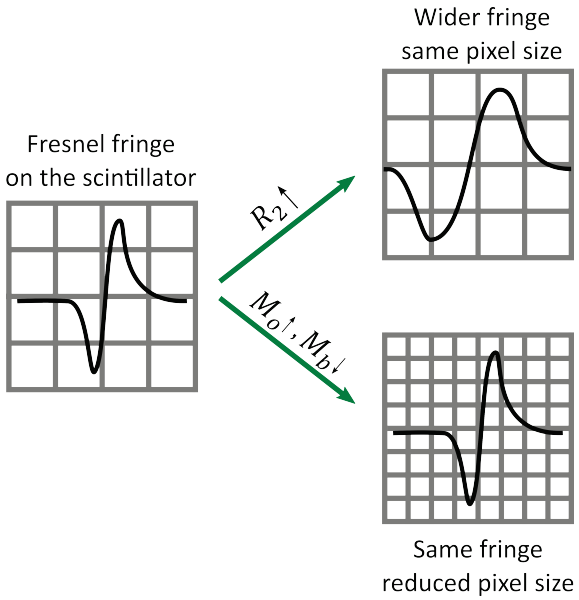


Figure 3.5: The mechanism behind propagation-based phase-contrast imaging

To properly sample the Fresnel fringes, either the fringes need to be enlarged so that they can be captured by a given pixel size on the scintillator, or the pixel size on the scintillator needs to be reduced to capture a given fringe size. The fringes' size is increased with larger X-ray propagation distance, yet with the cost of decreased spatial resolution. The scintillator's pixel size may be reduced by increasing the optical magnification ratio and/or decreasing the magnification due to binning. This is done in the expense of increased noise due to reduced photon count and reduced area per pixel on the detector, leading to increased exposure times and longer scans.

As discussed earlier, properly capturing the Fresnel fringes is imperative to performing phase-contrast imaging, which means that the pixel size at the scintillator plane has to be adequately small (Req 5). This can be achieved in two ways, either the fringes need to be enlarged (larger x_1) so that they can be captured by a given pixel size at the scintillator plane PS_{sci}, or the pixel size needs to be reduced to capture a given fringe size. The fringes' size x_1 is increased with larger

X-ray propagation distance, yet at the cost of decreased spatial resolution (larger α), as imposed by Req 4. The scintillator's pixel size PS_{sci} may be reduced by increasing the optical magnification M_o and/or decreasing the magnification due to the binning factor M_b following Equation (3.5). Reduction of the pixel size at the scintillator plane is done at the expense of increased noise or decreased SNR because of a reduced photon count (due to the increased optical magnification M_o) and reduced area on the detector per pixel when no binning is applied ($M_b = 1$). This leads to increased exposure times, thus longer scan times needed for the same image quality. Importantly, longer exposure times lead to an increased X-ray radiation dose, which might be of concern for biological soft tissue samples. Notice that the sequential increment of the propagation distance $R_{2,step}$, until all (minimal) requirements are satisfied, eventually yields a propagation distance R_2 that is tightly dependent on the rest of the imaging settings as well as the desired spatial resolution $\alpha_{p,des}$. It is crucial to understand that the whole optimisation process (for maximal spatial resolution, i.e. minimal feature size α) is a compromise between either achieving a smaller pixel size at the scintillator plane or increasing the Fresnel fringe width until the Fresnel fringes can be properly captured.

3.2 Experimental validation of the protocol design flow chart

As discussed in section '3.1.3 Protocol design flow chart for X-ray propagation-based phase-contrast imaging', the ideal theoretical nominal feature size of $\alpha = 0$ is not achievable due to experimental limitations captured by Req 5, while the intention of the proposed protocol design flowchart is to derive the experimental settings providing the best experimentally achievable nominal feature size. Req 1 – Req 3 are necessary to hold for the assumptions of the X-ray source, the imaged object and the mathematical modelling of the X-ray propagation in the near-field regime, respectively. Thus, a synchrotron experiment for free-space or inline propagation-based phase-contrast imaging was designed, which readily satisfies these first three requirements. The X-ray propagation distance R_2 was then progressively increased and the (theoretical) nominal feature size α that satisfies Req 4 was calculated. Next, the attained feature size α_{exp} was experimentally measured and the optimal propagation distance $R_{2,opt}$ was defined as the smallest propagation distance R_2 where the difference between the two feature sizes (α and α_{exp}) was not statistically significant. As the feature size α is an increasing function of R_2 (Equation (3.14)), the optimal propagation distance $R_{2,opt}$ yields the smallest practically attainable feature size. The proposed protocol design flowchart was thus tested and validated by its capability to predict $R_{2,opt}$ for two different sets of experimental settings (S_1, S_2), as detailed in section '3.2.1 Materials and methods'.

3.2.1 Materials and methods

A healthy murine colon in the form of a cylindrical FFPE wax block was prepared following the methods outlined in section '2.1.2 Tissue sample acquisition and preparation'. During this procedure, the melted wax was abruptly cooled down on an ice-cold plate, resulting in the formation of cracks throughout the wax block (see Figure 3.6). These cracks, although otherwise undesirable, conveniently serve as features suitable for the quantification of the spatial resolution as discussed later on. The sample was imaged at beamline I13-2 of the Diamond Light Source (DLS) with pink beam (beam with a broader than monochromatic X-ray spectrum). Beamline I13-2 was designed for high X-ray coherence that is mainly achieved by a large source-to-object distance $R_1 \approx 250\ m$, which results in a geometric magnification $M_g \approx 1$ for small propagation distances R_2 in the order of millimetres to centimetres (parallel beam geometry).

DLS is a synchrotron source and thus provides a highly brilliant X-ray beam, i.e. a high X-ray flux and high spatial coherence at the same time. For the current experiments, an effective X-ray energy E_{eff} of 27 keV ($\lambda = 0.046\ \mathrm{nm} = 0.46\ \text{Å}$) was selected, with a significant contribution to the available X-ray flux at 18 to 32 keV. These experimental settings readily satisfy Req 1, Req 2 and Req 3. The pco.edge 5.5 camera was used as detector (PCO AG, Kelheim, Germany) [219], which has a pitch of $PS_{det} = 6.5\ \mu m$. Two different sets of imaging protocols (S_1, S_2), shown in Table 3.1, were designed by varying the microscope objectives M_o and the binning factor M_b, and incrementally changing the X-ray propagation distance R_2, for 1501 projections per scan. For the first set S_1 the two-times (×2) objective was selected that resulted in a four-times (×4) optical magnification (a further ×2 magnification was provided by the optical lenses that convey the visible light from the (optical) microscope objectives to the detector [219]), no binning ($M_b = 1$), and an exposure time of 100 ms per projection. Protocols $P_{1,1} \dots P_{1,13}$ within S_1 correspond to propagation distances $R_2 =$ 25, 50, 75, 100, 175, 250, 325, 400, 426, 475, 550, 625, 700 mm, respectively. For the second set S_2 the four-times (×4) objective was selected that resulted in an eight times (×8) optical magnification, two-times (×2) binning ($M_b = 2$), and an exposure time of 10 ms per projection. Protocols $P_{2,1} \dots P_{2,12}$ within S_2 correspond to propagation distances $R_2 =$ 25, 50, 75, 200, 250, 273, 300, 500, 700, 900, 1100, 1300 mm, respectively.

Table 3.1: Protocol sets for experimental validation of the phase-contrast imaging protocol design flowchart

Two different sets of protocols (S_1, S_2) were designed by alternating the optical objectives M_o and the binning M_b. Incremental X-ray propagation distance R_2, for 1501 projections per scan, and an exposure time of 100 ms per projection.

Setting	R_2[mm]	Lens	M_o[-]	M_b[-]	PS_{sci}[μm]	Exposure [ms]	Projections
S_1	25, 50, 79, 100, 175, 250, 325, 400, 426, 475, 550, 625, 700.	2×	4	1	1.625	100	1501
S_2	25, 50, 75, 200, 250, 273, 300, 500, 700, 900, 1100, 1300.	4×	8	2	1.625	10	1501

Image processing was performed with the in-house implementation of Savu 2.3 [220] at I13-2, which is a versatile tomography reconstruction pipeline that allows cascading processing modules [221] to be applied on the data acquired at the beamline. The image pre-processing steps were: (i) Dezinger filter [222] (to suppress speckle-type noise called zingers that originates in interactions of X-rays with the detector) with an averaging kernel size of 5 frames and an outlier threshold $\mu = 10$ and (ii) ring removal regularisation [223] (to suppress tomography artefacts where defective pixels on the scintillator and/or detector result in a ring-like feature in the reconstructed CT data) with a regularisation strength parameter $\alpha = 5 \cdot 10^{-5}$. The phase was retrieved using the Paganin phase retrieval algorithm [200].

The ratio $\frac{\delta}{\beta} = 218$ was experimentally derived (qualitatively, through visual inspection) to obtain the value that better suppresses the Fresnel fringes at the cracks (boundaries between tissue and air)[2]. CT reconstruction was performed by standard filtered backprojection (FBP) [4, 224] with a Ram-Lak filter using the ASTRA toolbox implementation [225-227] within Savu. For each X-ray propagation distance R_2 the theoretical nominal feature size α was calculated using Equation (1.14). Next the non-phase-retrieved feature size α_{exp} and the phase-retrieved feature size $\alpha_{p,exp}$ were experimentally derived through the following approach:

The smallest ideal feature size was modelled as a rectangular pulse of width equal to α and height equal to h:

$$\Pi(x, h, \alpha) = \begin{cases} 0, & |x| > \frac{\alpha}{2} \\ h, & |x| \leq \frac{\alpha}{2} \end{cases}. \quad (3.19)$$

[2] Note that there is a discrepancy between the $\frac{\delta}{\beta}$ ratio used at DLS (sections 3.2.1 and 3.3.2) compared to that used at SLS (section 3.3.1) for similar imaging conditions. A ratio of $\frac{\delta}{\beta} = 218$ optimally reconstructs the boundary between tissue and air at DLS, while it is also suitable to reconstruct the boundary between tissue and wax at SLS. In contrast, a ratio of $\frac{\delta}{\beta} = 23$ was suitable to reconstruct the boundary between tissue and wax at DLS. Unfortunately, it was not possible to discern the source of this discrepancy, though it is likely to stem from an erroneous scaling in the implementation of the Paganin algorithm in either of the facilities.

where x is the pixel number across a line. The rectangular pulse can be mathematically represented as two opposing ideal step functions of width $\frac{\alpha}{2}$ (see Figure 3.6a). Since it is impossible to experimentally record the ideal step function for the reasons discussed in section '3.1.3 Protocol design flow chart for X-ray propagation-based phase-contrast imaging', the step functions constituting the rectangular pulse of Equation (3.19) would be have a sigmoid shape in an experiment (Figure 3.6). The minimal feature size α_{exp} could then be estimated by measuring the width $\frac{\alpha_{exp}}{2}$ of a step-like feature, in this case the cracks, in the reconstructed and non-phase-retrieved CT slice for α_{exp} or in the reconstructed and phase-retrieved CT slice for $\alpha_{p,exp}$. In practice, the width of the curved (or sigmoid) step $\frac{\alpha_{exp}}{2}$ was measured as the 10 % - 90 % rise time (x_R) [228], i.e. the number of pixels ($x_R = |x_b - x_a|$) between the points x_b where the curved step is at 10 % of its final height ($0.1 \cdot h$) and x_a where the curved step is at 90 % of its final height ($0.9 \cdot h$), as shown in Figure 3.6.

Practically, to quantify the spatial resolution on the middle slice of each of the reconstructed CT volumes, Fiji [229] (ImageJ version 2.0.0-rc-71/1.52p; Java 1.8.0_66), a scientific image processing software package written in Java that is distributed as an extended version of ImageJ [230, 231] was used. CT slices were kept as real (32-bit) grayscale images (output of Savu) to avoid introduction of errors due to bit-depth reduction. Firstly, cracks were identified on the middle slice of each reconstructed scan, which are ideal for this purpose as they represent abrupt transitions from paraffin wax to air. Linear regions of interest (ROIs) with 10 pixels (16.3 µm) width were then drawn across 10 random sites perpendicularly to the crack. The grayscale profile (averaged across the ROI width) was plotted and the width was measured as the 10 % - 90 % rise time. The feature size at each ROI was estimated as twice the width $x_{R,i}$ of ROI i as follows:

$$\alpha_{exp,i} = 2x_{R,i}, i = 1 \dots 10, \quad (3.20)$$

Then the feature size α_{exp} was estimated as the mean of the individual ROI feature sizes:

$$\alpha_{exp} = \overline{\alpha_{exp,i}}, i = 1 \dots 10. \quad (3.21)$$

where $\bar{\cdot}$ denotes the mean of the ten measurements. Statistical distributions of 10 measurements $\alpha_{exp,i}$ at each propagation distance R_2 were then compared with the corresponding theoretically predicted feature size α using a two-tailed *t*-test. The null hypothesis that α_{exp} is equal to α was rejected for $p < 0.05$. Thus, statistically significant differences ($p < 0.05$) were taken as evidence that the experimentally estimated feature size α_{exp} and the theoretically predicted feature size α do not match.

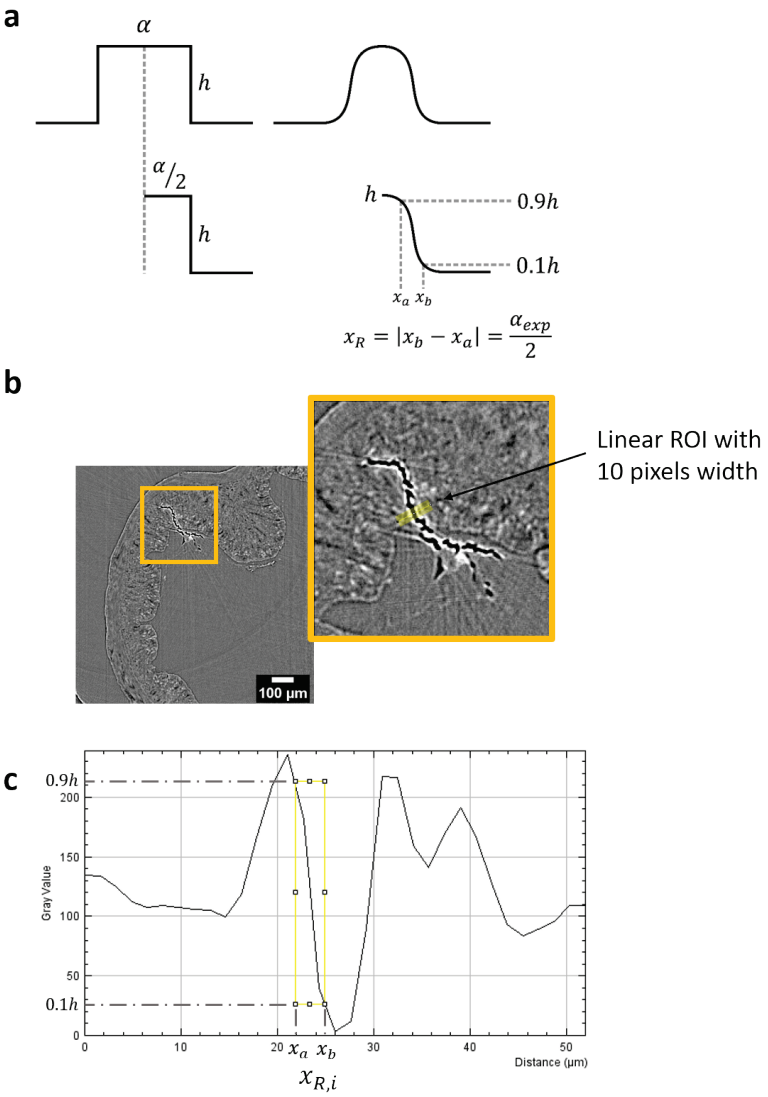


Figure 3.6: Definition and measurement of the spatial resolution

(a) To experimentally measure the feature size α_{exp} the smallest ideal feature size $\alpha = \alpha_{exp}$ was modelled as a rectangular pulse with width equal to α and height equal to h. The rectangular pulse can be mathematically represented as two opposing ideal step functions of width $\frac{\alpha}{2}$ and height h. In practice, the profile across a step-like feature in the reconstructed CT data has as a sigmoid shape and its width $\frac{\alpha_{exp}}{2}$ was measured as the 10 % - 90 % rise time x_R. (b) Cracks on the middle slice of each reconstructed scan were used as step-like features as they represent abrupt transitions from paraffin wax to air. Linear regions of interests (ROIs) with 10 pixels (16.3 µm) width were then drawn across 10 random sites perpendicularly to the cracks. (c) The grayscale profile was plotted for each ROI and the feature size was estimated as twice the 10 % - 90 % rise time $\alpha_{exp,i} = 2x_{R,i}$. The experimentally measured feature size α_{exp} for each CT slice was eventually estimated as the mean of the ten measurements $\alpha_{exp} = \overline{\alpha_{exp,i}}$.

3.2.2 Results

Figure 3.7 shows the theoretically predicted feature size α and the experimentally derived α_{exp} and $\alpha_{p,exp}$ for the non-phase-retrieved and phase-retrieved CT image (the middle slice of the CT volume), respectively, corresponding to the different propagation distances of set S_1 and set S_2 (see Table 3.1). As discussed in section '3.1.2 Minimal requirements for phase-contrast imaging', the nominal feature size α increases with increasing propagation distance R_2 (Equation (3.14)).

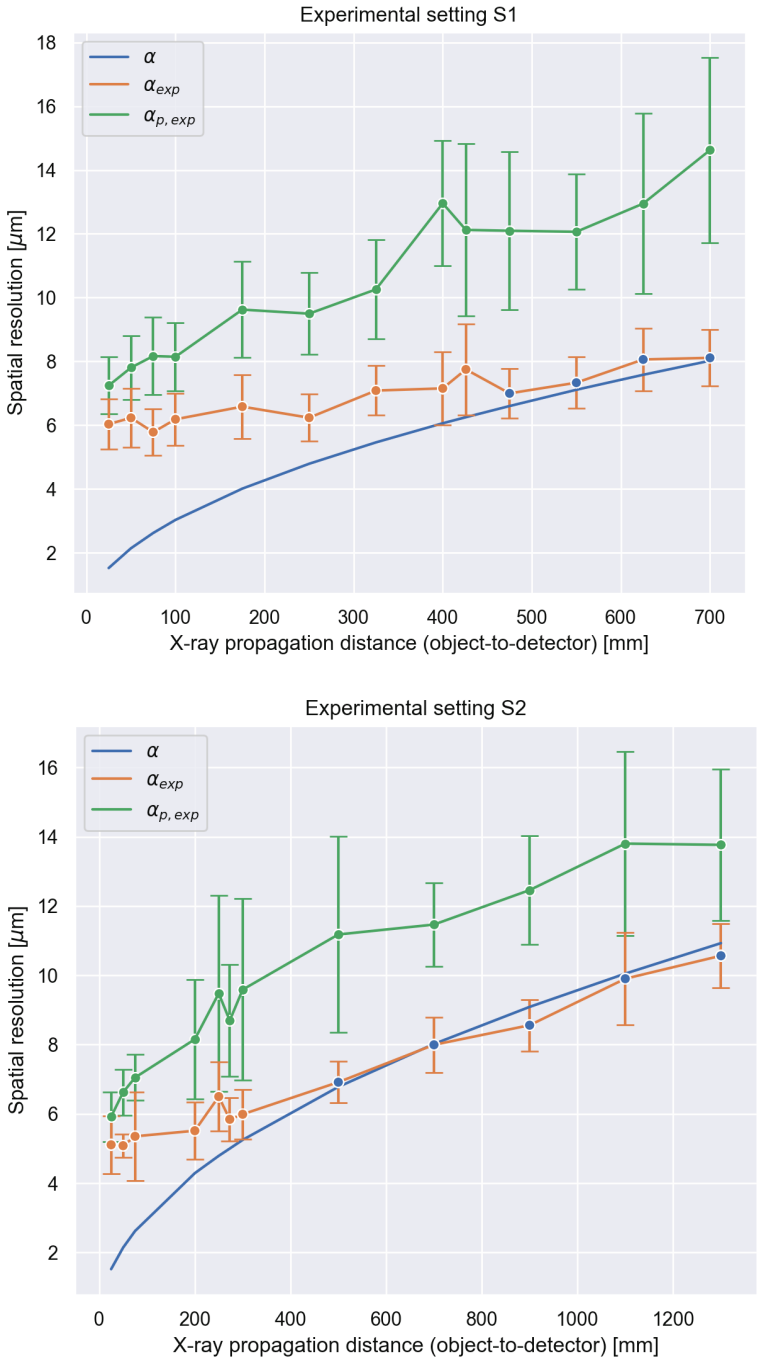


Figure 3.7: Theoretically predicted and experimentally measured spatial resolution

The theoretically predicted feature size α and the experimentally measured α_{exp} and $\alpha_{p,exp}$, respectively for the non-phase-retrieved and phase-retrieved CT image and for different propagation distances of set S_1 and set S_2 (see Table 3.1). Statistical distributions of 10 measurements $\alpha_{exp,i}$ at each propagation distance were compared with the corresponding theoretically predicted feature size α with one-tailed t-test and $p < 0.05$. Measurements with no statistically significant difference from the theoretical prediction are shown as blue points. The experimentally best spatial resolution with the settings of S_1 and S_2 was achieved at the optimal

propagation distance $R_{2,opt} = 475$ mm and $R_{2,opt} = 500$ mm, respectively, which were the smallest propagation distances where all of the requirements for phase-contrast imaging were concurrently satisfied. The experimentally measured feature size for the phase-retrieved image $\alpha_{p,exp}$ was larger than α_{exp} for all propagation distances R_2due to the low-pass filtering effect of the phase retrieval algorithm.

In contrast, the experimentally derived feature size α_{exp} remained constant for small propagation distances. A close look at Table 3.2 shows that the imaging protocols of set S_1 with $R_2 < 475$ mm do not satisfy the minimal set of requirements, resulting in Fresnel fringes being too small to be properly captured by the pixel size at the scintillator plane PS_{sci}. In this case, the spatial resolution (the minimum feature size) is limited by the pixel size at the scintillator plane. For $R_2 \geq 475$ mm there was no statistically significant difference ($p > 0.05$) between the experimentally retrieved feature size α_{exp} and the theoretically predicted feature size α.

Measurements where there was no statistically significant difference between the theoretical prediction α and the experimentally retrieved feature size α_{exp}, are shown as blue points in Figure 3.7, with α increasing for larger R_2 (diminishing spatial resolution). Therefore, when using the experimental settings of the first imaging protocol set S_1, the best spatial resolution was achieved at the optimal propagation distance $R_{2,opt} = 475$ mm. The experimentally derived feature size for the phase-retrieved image $\alpha_{p,exp}$ was larger than α_{exp} for all X-ray propagation distances R_2. The reduction in the spatial resolution of the phase-retrieved reconstruction is a result of the low-pass filtering effect of Paganin's algorithm. It depends on the cut-off frequency $|s| = \sqrt{u^2 + v^2} = \sqrt{\frac{\mu}{\delta}\frac{M}{z}} = \sqrt{\frac{4\pi}{\lambda}\left(\frac{\delta}{\beta}\frac{z}{M}\right)^{-1}} = \sqrt{\frac{4\pi}{\lambda}\left(\frac{\delta}{\beta}\frac{R_2}{M}\right)^{-1}} = \sqrt{\frac{4\pi}{\lambda}\left(\frac{\delta}{\beta}\cdot z_{eff}\right)^{-1}}$ (Equations (1.25) and (1.26) in section '1.4.5 X-ray phase retrieval'). This means that for a fixed ratio $\frac{\delta}{\beta} = 218$, an increasing (effective) propagation distance $z_{eff} = \frac{R_2}{M} = \frac{R_2}{\left(\frac{M_g M_o}{M_b}\right)}$ yields a smaller cut-off frequency, which results in a more aggressive smoothing of the phase-retrieved images. Smoothing is necessary to 'reverse' the increased width of the Fresnel fringes that is proportional to the propagation distance R_2 as per Equation (1.22). It does, however, inadvertently increase the phase-retrieved feature size $\alpha_{p,exp}$. While the phase retrieval algorithm was smoothing more aggressively, the difference between α_{exp} and $\alpha_{p,exp}$ for propagation distances $R_2 < R_{2,opt} = 475$ mm was increasing for larger propagation distances R_2, because the increased width of the Fresnel fringes was not properly captured due to the non-optimal experimental settings. On the other hand, the gap remained constant when the Fresnel fringes were properly captured ($R_2 \geq R_{2,opt} = 475$ mm) with a mean ratio of $\overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)} = 1.70$.

Table 3.2: Phase-contrast imaging requirements evaluated for imaging protocols within settings S_1 and S_2

For each of the protocols $P_{1,1}$ … $P_{1,13}$ within S_1 and $P_{2,1}$ … $P_{2,12}$ within S_2 the nominal feature size α for the corresponding X-ray propagation distance R_2 and an energy of 27 keV was computed with Equation (3.3). The mean value α_{exp} and standard deviation of the experimentally measured feature size for the non-phase-retrieved reconstructed data (evaluated at 10 different sites) is shown here. The difference between α_{exp} and α was evaluated for each protocol with two-tailed *t*-test and deemed not statistically significant for p-values $p < 0.05$ (shown here green). Finally the two conditions of Req 5 for optimal phase-contrast imaging was evaluated for each protocol with cross (×) signifying that the condition is not met and tick (√) signifying that the condition is met. The measurements suggest that the theoretically predicted nominal feature size α and the experimentally retrieved feature size α_{exp} are in agreement only for protocols where all requirements are satisfied (Req 1-Req 3 are readily satisfied via the synchrotron source, Req 4 is imposed via computation of α and Req 5 is evaluated for each protocol).

			Req 4				Req 5.1	Req 5.2
	Protocol	R_2 [mm]	α [µm]	α_{exp} [µm]	Standard deviation	t-test p-value	$R_2 > \frac{8 \cdot \left(PS_{det}\frac{M_b}{M_o}\right)^2}{\lambda}$	$M_o > \frac{2 \cdot M_b}{\alpha} \cdot PS_{det}$
S_1	$P_{1,1}$	25	1.52	6.03	0.79	2.16E-08	×	×
	$P_{1,2}$	50	2.14	6.23	0.92	1.96E-07	×	×
	$P_{1,3}$	75	2.62	5.78	0.73	2.38E-07	×	×
	$P_{1,4}$	100	3.03	6.18	0.82	6.84E-07	×	×
	$P_{1,5}$	175	4.01	6.58	1.00	1.94E-05	×	√
	$P_{1,6}$	250	4.79	6.23	0.73	1.56E-04	×	√
	$P_{1,7}$	325	5.46	7.09	0.78	9.67E-05	×	√
	$P_{1,8}$	400	6.06	7.15	1.15	1.50E-02	×	√
	$P_{1,9}$	426	6.25	7.75	1.42	8.83E-03	×	√
	$P_{1,10}$	475	6.60	7.00	0.78	1.41E-01	√	√
	$P_{1,11}$	550	7.11	7.33	0.81	4.06E-01	√	√
	$P_{1,12}$	625	7.58	8.06	0.98	1.55E-01	√	√
	$P_{1,13}$	700	8.02	8.11	0.88	7.49E-01	√	√
S_2	$P_{2,1}$	25	1.52	5.11	0.83	2.56E-07	×	×
	$P_{2,2}$	50	2.14	5.08	0.33	4.62E-10	×	×
	$P_{2,3}$	75	2.62	5.35	1.28	8.35E-05	×	×
	$P_{2,4}$	200	4.29	5.52	0.83	1.16E-03	×	√
	$P_{2,5}$	250	4.79	6.50	1.00	4.24E-04	×	√
	$P_{2,6}$	273	5.01	5.84	0.63	2.24E-03	×	√
	$P_{2,7}$	300	5.25	5.99	0.72	9.73E-03	×	√
	$P_{2,8}$	500	6.78	6.92	0.60	4.92E-01	√	√
	$P_{2,9}$	700	8.02	8.00	0.80	9.23E-01	√	√
	$P_{2,10}$	900	9.09	8.56	0.74	5.07E-02	√	√
	$P_{2,11}$	1100	10.05	9.90	1.33	7.34E-01	√	√
	$P_{2,12}$	1300	10.93	10.57	0.93	2.45E-01	√	√

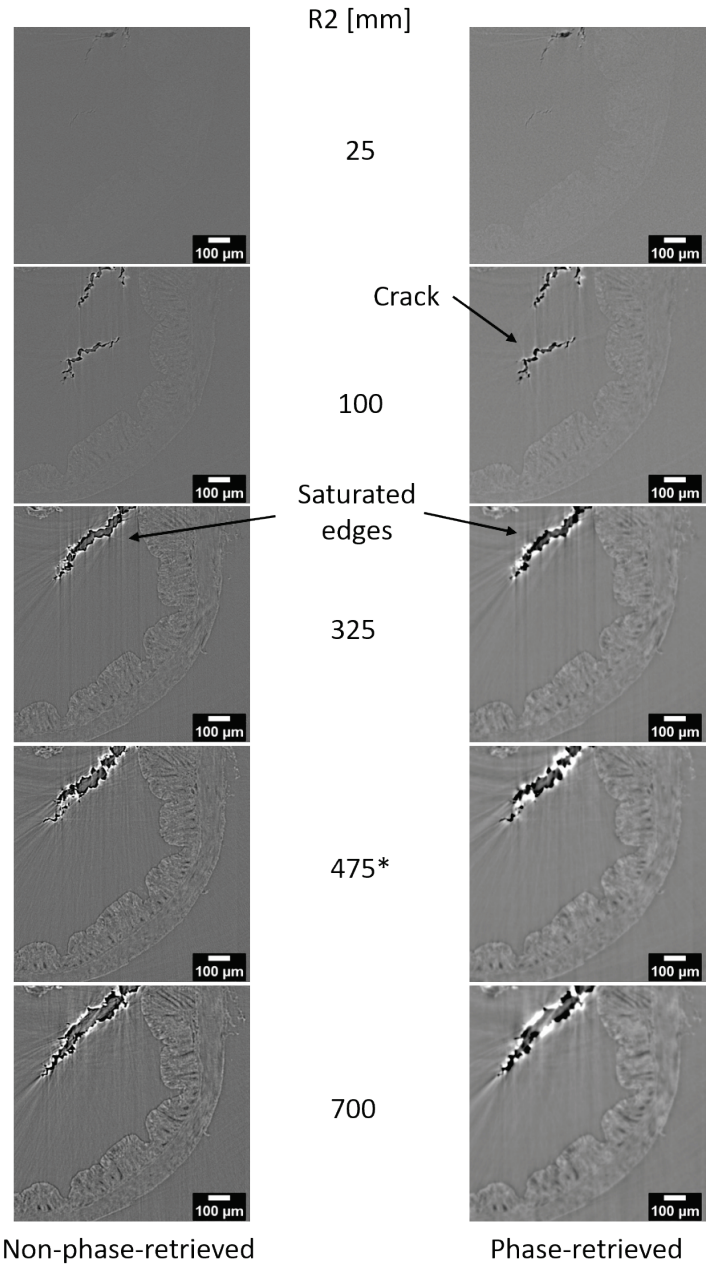


Figure 3.8: Reconstructed slices of experimental setting S1

The non-phase-retrieved and the phase-retrieved reconstructed CT slices corresponding to the protocols of S_1. For small propagation distances R_2 the sample is barely visible, while it becomes more prominent when R_2 approaches the optimal value $R_{2,opt} = 475\ \mathrm{mm}$. In the phase-retrieved slices the phase is sub-optimally retrieved on the boundaries between cracks and wax resulting in sharp (saturated) black/white edges due to exceedingly large Fresnel fringes that were not suppressed. This is not the case on the boundaries between tissue and wax, where the phase was optimally retrieved. This happens because the $\frac{\delta}{\beta}$ value was purposefully optimised for the interface between the tissue and wax in order to reveal the sample.

Figure 3.8 shows the non-phase-retrieved and the phase-retrieved reconstructed CT slices corresponding to the protocols of set S_1. For small propagation distances R_2 the sample is barely visible, while it becomes more prominent when R_2 approaches the optimal value $R_{2,opt} = 475\ \text{mm}$. When judged qualitatively, it is difficult to decide whether the reconstructed CT data corresponding to $R_2 = 475\ \text{mm}$ is better than that of $R_2 = 700\ \text{mm}$. In fact, the common misconception that an increasing propagation distance R_2 (always) improves image quality might lead in this case to the decision to select a bigger propagation distance that yields a sub-optimal feature size (worse spatial resolution). It is therefore important to either quantitatively evaluate or theoretically estimate the image quality in terms of spatial resolution in order to select the optimal propagation distance $R_{2,opt}$. Notice also that in the phase-retrieved slices the phase is sub-optimally retrieved on the boundaries between cracks and wax resulting in sharp black/white edges due to exceedingly large Fresnel fringes that were not suppressed. In contrast, this is not the case on the boundaries between tissue and wax, where the phase was optimally retrieved. This happens because the ratio $\frac{\delta}{\beta}$ was selected particularly for the interface between the tissue and wax in order to reveal the sample.

Table 3.3: Application of the protocol design flowchart for optimal soft-tissue X-ray phase-contrast imaging at a synchrotron facility

Shows how the flowchart of Figure 3.3 can be applied to predict the optimal propagation distance $R_{2,opt}$ (the smallest propagation distance R_2 where the minimal set of requirements is satisfied), and design the optimal imaging protocol. Beginning at step i., the initial values $PS_{det} = 6.5\ \mu m$, $R_1 = 250 m$, $E_{eff} = 27$ keV ($\lambda = 0.046\ \text{nm} = 0.46$ Å), $\alpha_{p,des} = 20$ µm, $R_{2,min} = 25$ mm, $M_b = 1$, $M_o = 4$ and $R_{2,step} = 225$ µm were selected. Next, following the steps of the flowchart of Figure 3.3 yields the optimal protocol at step viii.

Step	$\mathbf{M_o}$ [-]	$\mathbf{R_2}$ [mm]	$\boldsymbol{\alpha}$ [µm]	$\frac{\mathbf{M_b}}{\mathbf{M_o}} \leq \frac{\boldsymbol{\alpha}_{\mathbf{p,des}}}{\mathbf{6 \cdot PS_{det}}}$	$\mathbf{M_o} > \frac{\mathbf{2 \cdot M_b}}{\boldsymbol{\alpha}} \cdot \mathbf{PS_{det}}$	$\mathbf{R_2} > \frac{\mathbf{8 \cdot \left(PS_{det} \frac{M_b}{M_o}\right)^2}}{\boldsymbol{\lambda}}$	$\boldsymbol{\alpha} < \frac{\boldsymbol{\alpha}_{\mathbf{p,des}}}{\mathbf{3}}$
ii.				0.25 ≤ 0.51 | Yes			
iii.		25					
iv.			1.52				
v.					4 > 8.55 | No		
ix.		250					
iv.			4.80				
v.					4 > 2.71 | Yes		
vi.						250 > 459 | No	
ix.		475					
iv.			6.61				
v.					4 > 1.97 | Yes		
vi.						475 > 459 | Yes	
vii.							6.61 < 6.67 | Yes
viii.	Finished: $M_0 = 4,\ R_2 = 475\ \text{mm},\ E_{eff} = 27\ \text{keV},\ M_b = 1$						

Table 3.3 shows how the flowchart of Figure 3.3 can be applied to predict the optimal propagation distance $R_{2,opt}$, and to design the optimal imaging protocol. Beginning at step i., the initial values were selected as follows: The pco.edge 5.5 detector that was available at beamline I-13, DLS, has a $PS_{det} = 6.5\ \mu m$, while the source-to-object distance R_1 at I-13 is about 250 m. A high X-ray energy ($E_{eff} = 27$ keV) that renders the sample almost transparent was selected, which corresponds to a wavelength of $\lambda = 0.046\ \text{nm} = 0.46\ \text{Å}$ using Equation (3.8). The desired feature size of the phase-retrieved image was set at $\alpha_{p,des} = 20\ \mu\text{m}$ and the minimum achievable propagation distance $R_{2,min} = 25\ \text{mm}$. No binning was applied ($M_b = 1$) and the smallest available objective lens (2×) was selected. In conjunction with the additional 2× magnification provided by the optical lenses that convey the visible light from the microscope objectives to the detector [219], this resulted in an optical magnification $M_o = 4$. Finally, an X-ray propagation step of $R_{2,step} = 225\ \mu\text{m}$ was used to run the flowchart of Figure 3.3. Note however that for an accurate estimation of $R_{2,min}$ one needs to select $R_{2,step}$ as small as possible, at the expense of an increased number of iterations for the algorithm/flowchart to converge. The left column of Table 3.3 shows the step on the flowchart at a given time and the corresponding change in the protocol. At step viii., the flowchart finishes with the final protocol: $M_0 = 4$ (objective lens 2×), $R_2 = R_{2,opt} = 475\ \text{mm},\ E_{eff} = 27\ \text{keV},\ M_b = 1$. This protocol ($P_{1,10}$) in Table 3.2 yields a nominal feature size of the non-phase-retrieved image of $\alpha = 6.6\ \mu\text{m}$, which is indeed more than three times smaller than the desired feature size on the phase-retrieved reconstruction ($\alpha_{p,des} = 20\ \mu\text{m}$).

A close look at Figure 3.3 shows that the protocol design flowchart provides the user flexibility to design their experiment that is suitable for the application at hand. In particular, step x. allows for a free choice of the optical magnification M_o and/or binning M_b that eventually define the total magnification $M = \frac{M_g M_o}{M_b}$. These experimental settings are typically chosen in a way to take into account the size of the volume that needs to be imaged (increased total magnification M results in smaller volumes that can be imaged), image quality (higher binning M_b and/or lower optical magnification M_o result in reduced noise and thus increased SNR), and hardware availability (microscope objectives are optical lenses providing a fixed magnification level). Thus, for a given object with a certain requirement for the feature size $\alpha_{p,des}$ one could design different optimal protocols by selecting different experimental settings at step x.

Running the flowchart of Figure 3.3 again with $PS_{det} = 6.5\ \mu m$, $R_1 \approx 250\ m$, $E_{eff} = 27$ keV, $\alpha_{p,des} = 20\ \mu\text{m}$, $R_{2,min} = 25\ \text{mm}$, $R_{2,step} = 240\ \mu\text{m}$ and choosing a different magnification objective and binning values at step x. results in a different optimal protocol: $M_0 = 8$ (objective lens 4×), $R_{2,opt} = 505\ \text{mm},\ E_{eff} = 27\ \text{keV},\ M_b = 2$. The corresponding protocol ($P_{2,8}$) in Table

3.2 yields a nominal feature size of the non-phase-retrieved image of $\alpha = 6.8$ µm. Figure 3.7 shows the theoretically predicted feature size α and the experimentally determined feature size α_{exp} and $\alpha_{p,exp}$ for the non-phase-retrieved and phase-retrieved CT image (the middle slice of the CT volume), respectively, corresponding to the different protocols of set S_2 (see Table 3.3). Similarly to S_1, the experimentally derived feature size α_{exp} remained constant for small propagation distances, while for larger propagation distances, no statistical difference between α_{exp} and the nominal feature size α has been observed (blue points in Figure 3.7), with an optimal propagation distance of $R_2 = 500\ \mathrm{mm} \approx\ R_{2,opt} = 505\ \mathrm{mm}$. The mean ratio between the phase-retrieved and non-phase-retrieved feature size was in this case $\overline{\left(\frac{\alpha_{p,exp}}{\alpha_{exp}}\right)} = 1.44$, in line with the experimentally derived condition at step vii. of the flowchart, i.e. that the nominal feature size α needs to be at least three times smaller than the desired feature size after phase retrieval, which again justifies the choice of requiring the nominal feature size α to be at least three times smaller than the desired feature size after phase retrieval ($\alpha < \frac{\alpha_{p,des}}{3}$). This difference in the spatial resolution is evident in Figure 3.9, which shows the non-phase-retrieved and the phase-retrieved reconstructed CT slices corresponding to the imaging protocols of set S_2. Similarly to Figure 3.8, the object becomes more visible with increasing propagation distance.

The experimental approach applied here validated the theoretical expectation of a minimal set of technical requirements necessary for (free-space or inline) propagation-based x-ray phase-contrast imaging. It was shown that the theoretical estimation of the nominal feature size is only accurate when the minimal set of requirements described in section '3.1.2 Minimal requirements for phase-contrast imaging' were satisfied. This underlines an important, yet widely overlooked aspect of propagation-based phase-contrast imaging: there is an optimal propagation distance that yields the best experimentally achievable spatial resolution, while increasing the propagation distance does not necessarily improve image quality. It was thus demonstrated here how the protocol design flowchart for X-ray propagation-based phase-contrast imaging can be used in practice to estimate the optimal propagation distance, in order to derive iteratively an imaging protocol that satisfies the minimal set of technical requirements and a user-defined desired feature size.

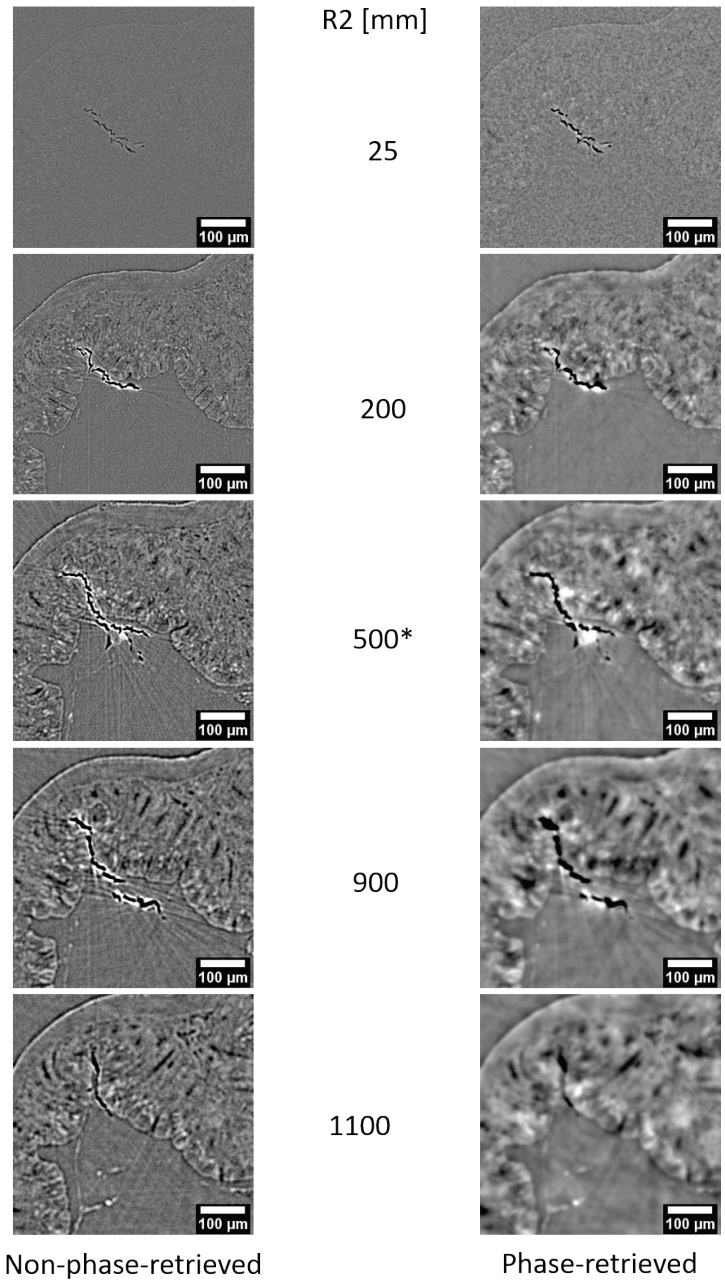


Figure 3.9: Reconstructed slices of experimental setting S2

The non-phase-retrieved and the phase-retrieved reconstructed CT slices corresponding to the protocols of S_2. For small propagation distances R_2 the sample is barely visible, while it becomes more prominent when R_2 approaches the optimal value $R_2 = 500\ \text{mm} \approx R_{2,opt} = 505\ \text{mm}$. Here it is readily apparent that the image quality deteriorates with increasing propagation distances greater than the optimal value $R_2 > R_{2,opt}$, especially in the phase-retrieved images.

3.3 Phase-contrast imaging of murine colons for the study of crypt budding

The flowcharts developed in section '3.1 Soft tissue imaging in practice' are versatile tools capable of designing optimal in-line propagation-based X-ray phase-contrast imaging protocols both in synchrotrons and in the lab. These were used to design different protocols in order to image murine colons induced with early-stage colorectal cancer at different X-ray imaging setups. This section aims to showcase the technical capabilities of X-ray phase-contrast imaging of soft tissues and specifically its applicability in the study of colonic tissue embedded in paraffin wax.

Firstly, in section '3.3.1 High-resolution synchrotron phase-contrast CT of murine colons' the experimental settings were pushed to the limit of the technical capabilities of the synchrotron facility so that the best experimentally possible spatial resolution could be achieved. This provided CT images comparable to those obtained with traditional histology in terms of spatial resolution in order to validate that X-ray μCT imaging followed by traditional histology staining, is indeed feasible. Next, in section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons' the experimental settings were selected such that the entire colonic tube section (4 mm in diameter) could fit in the reconstructed CT slices, yielding an overview of the sample. This enabled imaging a large number of samples produced during the animal study described in section '2.1 Animal model materials and methods' obtaining CT data suitable for morphological quantification of crypt budding used in the chapters that follow.

3.3.1 High-resolution synchrotron phase-contrast CT of murine colons

The flowchart of Figure 3.3 was used to design an optimal high-resolution protocol for in-line phase-contrast X-ray imaging at the TOMCAT beamline at Swiss light source, Switzerland. An effective X-ray energy E_{eff} of 21.0 keV ($\lambda = 0.059$ nm $= 0.59$ Å) was selected. The pco.edge 5.5 camera was used as detector (PCO AG, Kelheim, Germany) [219], which has a pitch of $PS_{det} = 6.5$ μm. Twenty times optical magnification ($M_o = 20$) and no binning ($M_b = 1$) was applied, with X-ray propagation distance $R_2 = 14$ mm for 1501 projections per scan, and an exposure time of 100 ms per projection, yielding a pixel size of the reconstructed data equal to $PS_{obj} = 0.325$ μm. Phase retrieval and CT reconstruction was performed using the in-house software suite available at the TOMCAT beamline. The phase was retrieved using the Paganin phase retrieval algorithm [200].

The ratio $\frac{\delta}{\beta}$ was derived following a two-step approach: Firstly, an initial estimation for the refractive index between the tissue and wax was calculated as follows: The chemical composition

for soft tissue was approximated as $C_{12}H_{60}O_{25}N_5$ [165] and that of paraffin wax was approximated as $C_{21}H_{27}NO_3$ [232], both mass densities were assumed to be equal to $1\ g/cm^3$ and the X-ray effective energy equal to 27.0 keV. The refractive indices with respect to vacuum were then obtained by tabulated values [178] for the soft tissue (st) and paraffin wax (pw). This resulted in an initial estimation of the $\frac{\delta}{\beta} = \frac{\delta_{st}-\delta_{pw}}{\beta_{st}-\beta_{pw}} = 115$ between the soft tissue and paraffin wax. This initial estimation was subsequently refined experimentally through trial and error in order to obtain a $\frac{\delta}{\beta}$ value that better suppresses the Fresnel fringes (qualitatively, through visual inspection). The final value after refinement $\frac{\delta}{\beta} = 218$ is about twice the initial estimation. This likely stems from mismatch between the assumed mass densities of the tissue and wax. Note that, as refraction in this particular case takes place on the interface between soft tissue and wax, which have similar refractive indices, the differential refractive index $\frac{\delta}{\beta} = \frac{\delta_{st}-\delta_{pw}}{\beta_{st}-\beta_{pw}} = 218$ is one order of magnitude smaller than that of soft tissue not embedded in wax.

CT reconstruction was performed by the Gridrec algorithm [233], which is a high-performance alternative to the standard filtered backprojection algorithm, with a Parzen filter. The resulting reconstructed CT slices were 2560×2560 pixels^2 or 0.83×0.83 mm^2 large, and each scan consisted of 2160 slices or 0.7 mm. This particular protocol was developed for a desired feature size on the object plane after phase retrieval of $\alpha_{p,des} = 10$ µm. This resulted in a nominal feature size prior to phase retrieval of $\alpha = 1.29$ µm, which is indeed more than 3 times smaller than $\alpha_{p,des}$, as imposed by the protocol design flow chart.

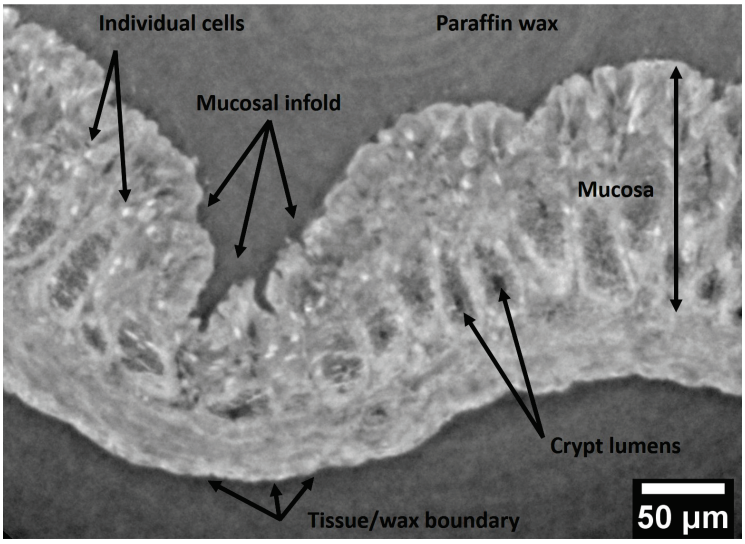


Figure 3.10: High-resolution synchrotron phase-contrast CT of a healthy murine colon

The optimal phase-contrast protocol provided enough soft-tissue contrast revealing the colon out of the paraffin wax. The gross colonic structure (e.g. mucosal folds and

tissue/wax boundaries) is readily visible as the tissue is brighter than the wax. The mucosal tissue layer can be differentiated, and crypts are visible within. Crypt lumens appear darker than the surrounding tissue.

Figure 3.10 shows a cropped CT slice of colonic tissue from a mouse in the untreated control group culled at the end of cycle 1, and prepared as described in section '2.1.1 Azoxymethane and dextran sodium sulphate mouse model'. The carefully designed protocol provides high image contrast, yielding much brighter grey values for the tissue compared to the wax, effectively revealing the sample. Due to the high spatial resolution (minimisation of the nominal feature size) the microstructure of the murine colon has been revealed. Individual cells within the mucosa are visible, as well as the different tissue layers (mucosa, muscle layers and serosa) and the crypts.

Next, in order to validate that the microstructure is accurately depicted by µCT through direct comparison between the CT reconstruction and the corresponding histology section, traditional histology was performed on one of the samples. The process described in section '2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining' was performed on one colonic tissue sample from a mouse in the DSS-treated group that was culled at the end of cycle 1. In brief, the cylindrical FFPE wax block was re-embedded into a standard wax block after µCT imaging and thin (5 µm) slices or sections were taken off the surface of the block. The paraffin sections were stained with H&E and imaged with an Olympus VS110 light microscope [234] with 20× optical magnification, resulting in a pixel size of $PS_{hist} = 0.32$ µm, which is almost identical to the pixel size of the CT data.

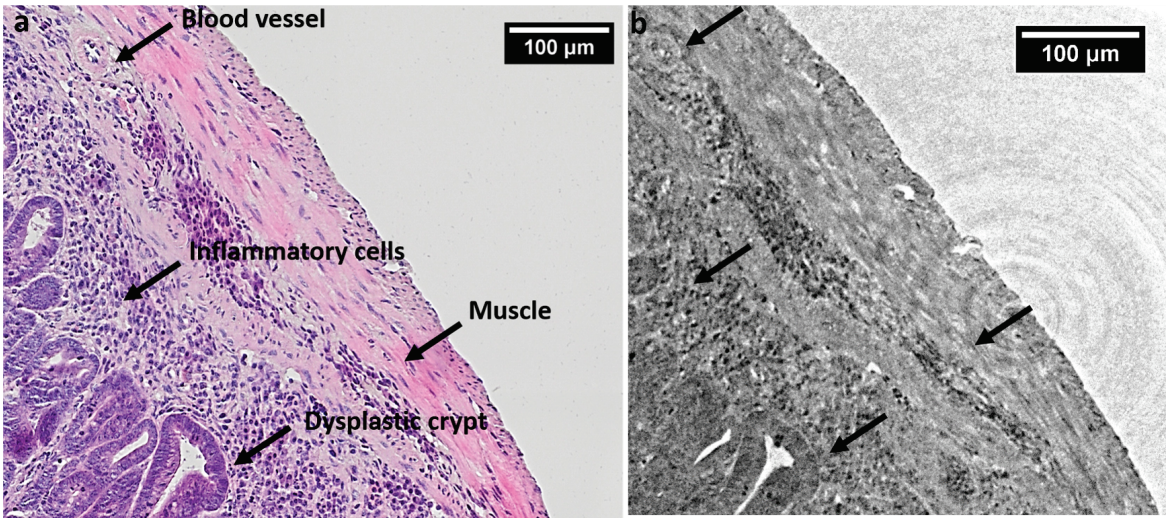


Figure 3.11: Histology vs µCT comparison of a FFPE murine colon sample

(a) H&E histology slide reveals the colonic structure including muscles and blood vessels within the submucosa, and inflammatory cells and dysplastic crypts within the mucosal layer. (b) The same features can be identified in the CT slice of the same

site (CT performed prior to paraffin sectioning and H&E histology). This confirms that X-ray CT non-destructively and accurately depicts the colonic structure while still permitting histology staining to be performed. In b. the grey values were inverted in order to match the colour brightness of the H&E stain.

As the orientation of the histology section was similar to that of the CT slice (paraffin section was taken on the surface of the standard wax block), it was possible to identify the matching CT slice by manually searching throughout the CT image stack. In-plane alignment of the two images was performed visually using the rotation tool of Fiji so that the two images are roughly aligned. As discussed in section '1.4 Advanced imaging approaches in research', the grey values in the CT reconstructions are proportional to the mass density of the material being imaged, thus denser tissues (e.g. muscle) appear brighter than less dense ones (e.g. mucosa) as shown in Figure 3.10. As a result, the colonic crypt lumens appear dark and individual cells appear bright. In contrast, tissue types are revealed with certain tints of pink/purple in H&E histology depending of the chemical bindings between the stain and the tissue (see section '2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining'). Thus, individual cells are stained darker than the cytoplasm, which is stained pink, and the muscles, which are stained pinky red. In order to accommodate for this mismatch in colour-matching, the grey values of the CT slice were inverted (a process where the grey value of each pixel is flipped so that bright pixels become dark and *vice versa*) using the invert tool in Fiji. The resulting image is shown in Figure 3.11b.

Visual inspection of the histology section and the corresponding (inverted) µCT slice in Figure 3.11 shows that the two modalities reveal the same underlying structure of the tissue. Individual cells within the mucosal layer can be identified in the CT data, as well as muscles and structures such as blood vessels. Both the lumen and the epithelium of colonic crypts can be identified and the sizes are identical. This shows that X-ray µCT imaging of FFPE colonic tissue can indeed accurately reveal the microstructure of the tissue and it does so non-destructively. Therefore, the workflow proposed in '2.1 Animal model materials and methods' for *in situ* 3D imaging of a large number of crypts at once while at the same time enabling functional assays (histology staining, including H&E) to be performed is indeed feasible.

3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons

As discussed in section '1.1.2 The need for new tools in the study of crypt budding' it is important to obtain an overview image of the colonic samples in order to capture a large number of crypts in each scan. To achieve this, one has to sacrifice spatial resolution in order to fit a larger volume of the sample in the detector's field of view.

Thus, a low-resolution protocol was developed for in-line phase-contrast X-ray imaging at beamline I13-2 of the Diamond Light Source, Didcot, UK. An effective X-ray energy $E_{eff} = 25$ keV ($\lambda = 0.05$ nm $= 0.5$ Å) and an X-ray propagation distance of $R_2 = 230$ mm were selected. The pco.4000 camera was used as detector (PCO AG, Kelheim, Germany) [219], which has a pitch of $PS_{det} = 9$ µm. The four-times (×4) objective was selected that resulted in eight-times (×8) optical magnification (a further ×2 magnification was provided by the optical lenses that convey the visible light from the optical magnification objectives to the detector [219]), two-times (×2) binning ($M_b = 2$) and an exposure time of 85 ms per X-ray projection for 2501 projections. This resulted in a pixel size on the object plane of $PS_{obj} = 2.2$ µm with 1844×1844 pixels^2 or 4.1×4.1 mm^2 large reconstructed CT slices, and 1176 slices or 2.6 mm per CT image stack. Phase retrieval was performed using the Paganin algorithm [200] with ratio $\frac{\delta}{\beta} = 23$ (derived experimentally, qualitatively, through visual inspection), prior to standard filtered back projection (FBP) [4] for CT reconstruction using the in-house implementation Savu 2.2 [220].

The colonic samples prepared during the animal study described in section '2.1.1 Azoxymethane and dextran sodium sulphate mouse model' were imaged with this protocol. A representative CT slice from an untreated control sample and a sample threated with one cycle of AOM&DSS are shown in Figure 3.12. The experimental settings were such that the entire colonic cross-section could fit in the field of view while the spatial resolution was enough to capture the gross colonic microstructure. Specifically, the tissue layers can be differentiated and the colonic crypt lumens can be identified as darker spots in the CT data. However, as the colonic structure is naturally complex, with varying mucosal thickness, folds and ever-changing curvature, the colonic crypts are not well aligned with the CT image plane. Therefore, not all crypts are dissected longitudinally in each CT slice, making their identification and the quantification of their 3D shape particularly challenging. To this end, the murine colons were digitally unrolled using a technique that was developed for this particular purpose (see 'Chapter 4 Colonic 3D cyclorama: a method for digital volume unrolling'), enabling the identification of budding crypts within the samples that were subsequently phenotyped in terms of their 3D structure (see 'Chapter 6 Quantitative morphometry of crypt budding in a mouse model').

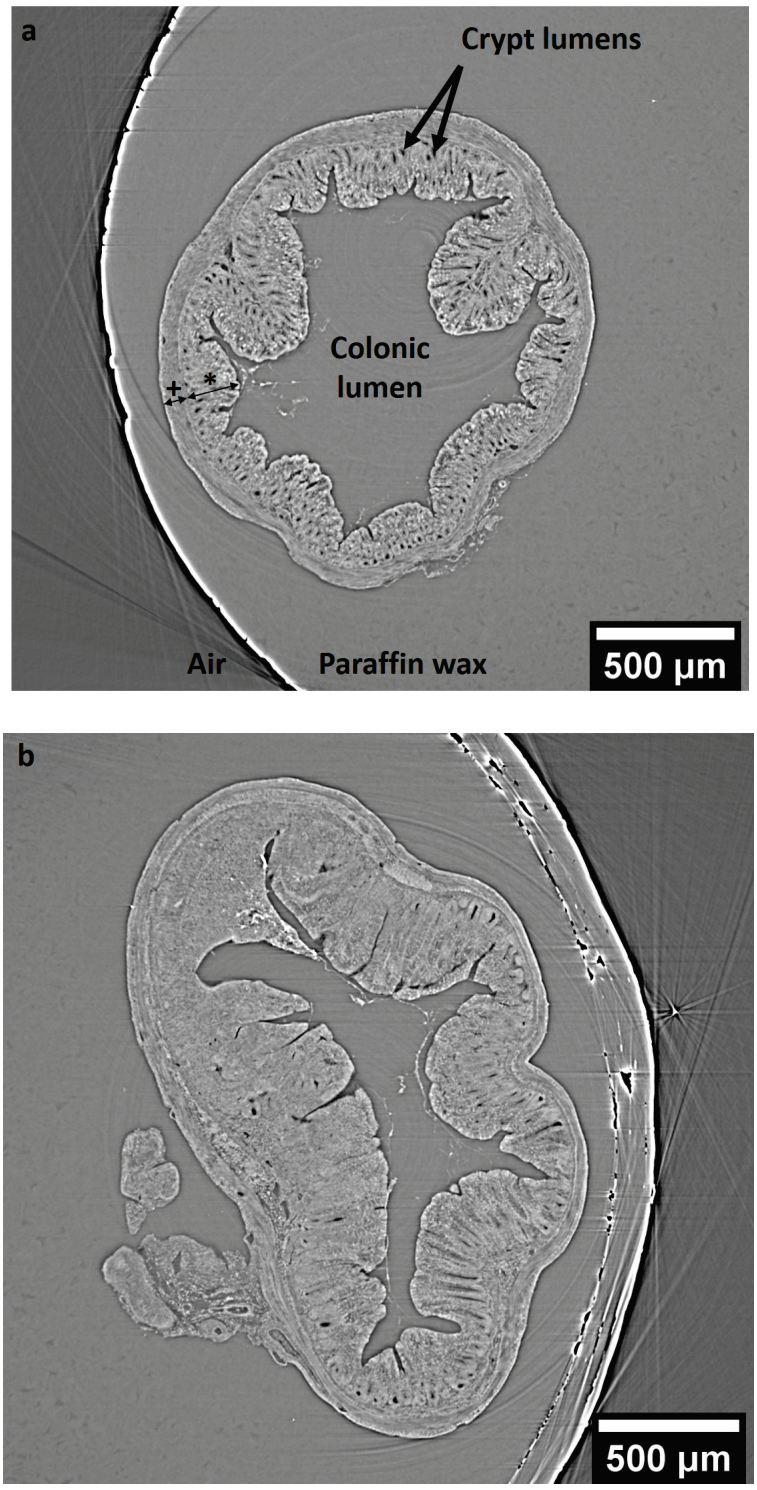


Figure 3.12: Low-resolution synchrotron phase-contrast CT of murine colons

A representative CT slice from an untreated control sample (a) and a sample threated with one cycle of AOM&DSS (b).The experimental settings were such that the entire colonic cross-section could fit in the field of view while the spatial resolution was enough to capture the gross colonic structure (muscle: '+' and mucosal: '*' layers) and microstructure (crypt lumens).

Chapter 4 Colonic 3D cyclorama: a method for digital volume unrolling

Chapter 2 deals with the production of FFPE murine colon samples with early-stage CRC with a focus on preserving the natural tubular shape of the colon. In Chapter 3, a method was developed to non-destructively image these intact colons in 3D using X-ray µCT, which enables subsequent traditional histology to be performed. The above-mentioned workflow is only conducive to quantitative morphometry when a method of identifying and extracting the 3D shape of crypts is available. However, the colonic crypts are not well aligned with the CT image plane (see section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons'), making their identification particularly difficult.

To this end, a novel method to digitally unroll the volume of deformed cylindrical tubes of non-uniform thickness was developed and implemented. In brief, the internal volume of the colon (deformed tube) is formulated as a set of onion-like 3D surfaces that follow its shape. These surfaces, which slice the colonic crypts perpendicularly, are subsequently mapped onto planar panoramic views. The resulting 3D image stack, which effectively unrolls the colon virtually, is termed here '3D cyclorama'. As the crypts extend radially through the colonic tissue, the 3D cyclorama method allows tracing individual crypts through the several tissue layers and identifying buddying and non-budding crypts.

Next, section '4.1 Introduction' examines existing techniques to digitally flatten curved surfaces. Section '4.2 Development' describes the mathematical formulation of the 3D cyclorama method, followed by section '4.3 Discussion'. Finally, section '4.4 Results' describes the specific use of 3D cyclorama particularly for digital unrolling of FFPE murine colons. In silico validation of the method is detailed in Appendix 'A.1 Validation using a digital phantom', and two case studies are provided in Appendix 'A.2 Interdisciplinary applicability: Case studies'.

4.1 Introduction

The problem of digitally unrolling surfaces has been extensively studied in the literature, where predominant applications include straightening wrinkled paper [235-239] or open-book documents [240], as well as virtual unrolling of ancient papyri [241-245]. These methods usually attempt to recover text that is written on deformed sheets by fitting 3D surfaces, and digitally mapping them onto planar images. A prime example of the need for volumetric digital unrolling in industry is for the study of initial defects due to delamination of the jelly-roll (internal design

made of a rolled sheet of an anode, a separator, and a cathode) of cylindrical batteries [246, 247]. Specifically, Ziesche and colleagues [247] have used a plug-in created by the Konrad-Zuse-Zentrum Berlin (ZIB) [241] for the commercial software package Amira (Amira 6.5, ZIB/Thermo Fisher Scientific, Germany/USA) to digitally flatten the jelly-roll, which is a scroll of uniform thickness. Unrolling volumetric deformed tubes or scrolls of non-uniform thickness though, adds another level of complexity.

In the biomedical field, volumes that are roughly cylindrical in shape are prominent, including certain bones and the colon. In bones, the vast bone cell networks formed by osteocytes, the most abundant bone cells, are receiving increasing interest as imaging and quantifying their 3D structure can enhance the understanding of bone development in health and diseased conditions, including osteoporosis and osteoarthritis [248, 249]. For the study of early-stage CRC in the murine colon, non-destructively imaging the whole sample in 3D rather than physically cutting it, enables visualising a large number of crypts in their natural shape. Digital unrolling of the volume becomes then necessary for investigating the tubular sample at different tissue depths.

4.2 Development

To unroll a deformed cylindrical tube (i.e. the murine colon sample imaged via µCT), the volume it occupies is initially defined using the inner and outer boundary of the tube in 3D. Given an image-stack depicting the deformed cylindrical tube, these boundaries are defined as a collection of inner and outer boundary contours drawn manually on 2D image slices. The sample boundaries are subsequently virtually charged with electrical charges, and the field lines of the resulting electrostatic field are simulated. Intermediate or internal contours that are non-intersecting and follow the shape of the sample can then be defined using the field lines. Each internal contour is drawn at a certain percentage of the local thickness of the deformed tube so that it corresponds to a relative local thickness. These are then used to create internal (‘onion-like’) re-slicing surfaces that are subsequently non-rigidly mapped onto rectangular 2D panoramic views. When stacked together, panoramic views at different depths of the cylindrical volume provide a flattened version of the deformed tube, termed here a ‘3D cyclorama’.

The 3D cyclorama method was implemented and tested as a Jython script in Fiji [229] (ImageJ version 2.0.0-rc-68/1.52g; Java 1.8.0_66), a scientific image processing software package written in Java that is distributed as an extended version of ImageJ [230, 231]. Jython is an implementation of the Python programming language that runs on Java platforms and is provided as one of the available scripting languages in Fiji.

4.2.1 Meshing: Generation of onion-like re-slicing surfaces

A sampling grid (set of onion-like surfaces) within the deformed cylindrical tube is generated as a set of intermediate or internal contours, corresponding to incremental relative local depth levels (0 % to 100 %) as follows: Given is a stack of images depicting 2D cross-sections of a vertically aligned deformed cylindrical tube, for instance provided as individual reconstructed CT slices at different heights of the sample. Thus, each slice represents a deformed 2D annulus that is defined by its inner and outer boundaries. The boundaries are subsequently used to interpolate the enclosed cross-sectional area and describe the internal contours at different relative local thickness levels.

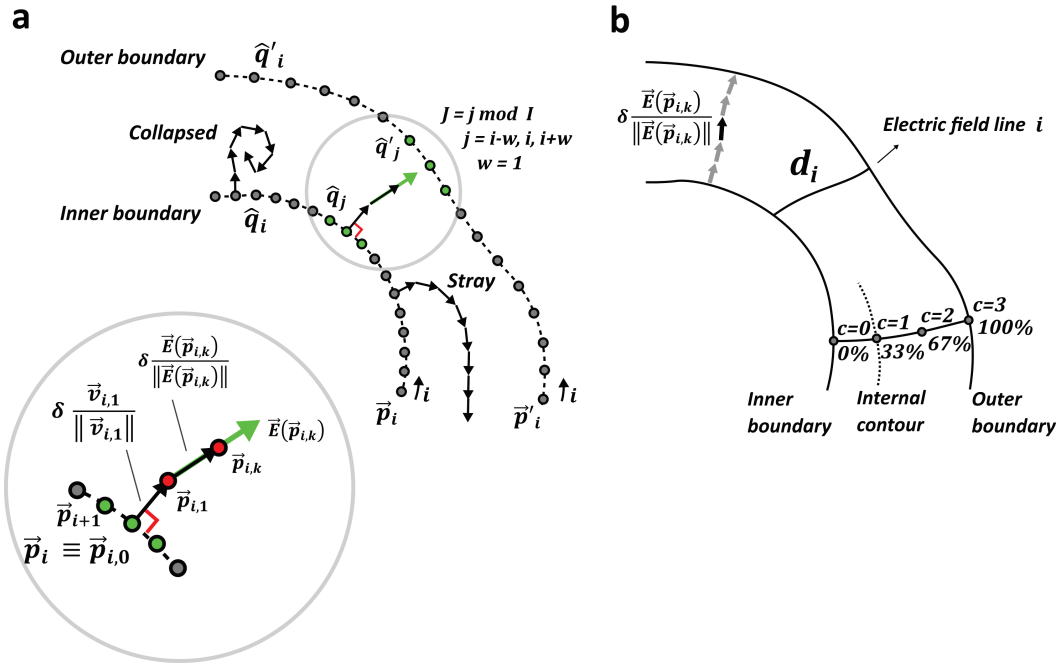


Figure 4.1: Definition of electric field lines and contours

(a) Boundary contours are defined on the deformed tube boundaries as sequences of charged points $\vec{p}_i$ and $\vec{p}'_i$. Each electric field line is built up incrementally starting at point $\vec{p}_i \equiv \vec{p}_{i,0}$. The first increment $\delta \cdot \frac{\vec{v}_{i,1}}{\|\vec{v}_{i,1}\|}$ is normal to the inner boundary. Each subsequent increment $\delta \cdot \frac{\vec{E}(\vec{p}_{i,k})}{\|\vec{E}(\vec{p}_{i,k})\|}$ is defined as the direction of the electric field scaled by δ (electric field line segment size), considering charges on both boundaries in a user-defined window (green points ●). This process might fail to connect point $\vec{p}_i$ to the outer boundary, resulting in 'collapsed' or 'stray' lines, which are identified and removed. (b) The local thickness is defined as the length d_i of field line i. Intermediate contours $cont_{h,c}$ are defined on a total of C relative depth levels $\frac{c}{C-1} \cdot 100\%$ ($c = 0,1, \dots, C-1$).

The local thickness is defined by employing notions from physics, namely electrostatics. Initially the inner and outer boundaries of the cross-sections are loaded by opposite electrical charges. The 2D electric field lines are then calculated within the enclosed space described by those boundaries as shown in Figure 4.1a. The theory governing electrostatic fields [250] ensures that electric field lines never cross, and that they never form closed loops (they always start at one boundary and cross the tube to reach the other boundary). Moreover, the electric field lines are always perpendicular to the boundaries. These properties make the electric field lines an ideal tool for the definition of the relative local 2D thickness, which is defined here as the length d_i of the respective electric field line i, spanning the inner and outer boundary of the tube's cross-section Figure 4.1b.

The electric field $\vec{E}$ at point $\vec{p}$ is defined as the sum of the contributions from points $\vec{p}_j$ in space, each charged with an electric charge q_j:

$$\vec{E}(\vec{p}) = \sum_{j=1}^{N} \frac{q_j}{4\pi\varepsilon_o} \frac{(\vec{p}-\vec{p}_j)}{\|\vec{p}-\vec{p}_j\|^3}, \qquad (4.1)$$

where ε_o is the electric constant, $j = 1 \dots N$ is the set of contributing charges, and $\|\cdot\|$ denotes the Euclidean distance. Due to the computational costs involved to calculate the exact electric field at all possible positions in space, equation (4.1) was not implemented directly here, but the following computational approach has been developed to approximate the electric field lines:

1. First, the boundary contours are defined as sequences of equidistant points $\vec{p}_i$ and $\vec{p}'_i$ ($i = 1 \dots I$) for the inner and outer boundary, respectively, with the total number of points I being a (positive) integer given by the user. These points define the positions of the electrical charges that are used here for the calculation of the field lines. Each point on the contour can be identified by an index i that is used to select the contributing charges as discussed next (Figure 4.1). Note that for tubes the boundary (and internal) contours are closed, and hence the points $\vec{p}_{i=1}$ and $\vec{p}'_{i=1}$ are arbitrary. However, it is important for the generation of cycloramas (see section '4.2.2 Mapping: Generation of 3D cycloramas') that the sequences of points on the internal and external boundary are paired, i.e. points with index i on the two opposing boundary contours are fairly close to each other. This will be ensured via appropriate selection of the starting point.

2. Each electric field line i is built up incrementally, by starting at the location of charge q_i (point $\vec{p}_i \equiv \vec{p}_{i,0}$ on the inner boundary contour) and computing the electric field at points $\vec{p}_{i,k}$. Due to the term $\|\vec{p} - \vec{p}_j\|$ in the denominator of equation (4.1) singularities may arise when point $\vec{p}$ is very close to the charged points $\vec{p}_j$, as this would result in division by zero. To avoid this problem, the first calculation of the electric field is bypassed and replaced by progressing along the

direction of the vector $\vec{v}_{i,1}$, at the location $\vec{p}_{i,0}$ of charge q_i and pointing inwards the annulus (Figure 4.1a). The vector $\vec{v}_{i,1}$ is calculated as the vector perpendicular to $\vec{u}_i = (u_{i,x}, u_{i,y}) \coloneqq \vec{p}_{i+1,0} - \vec{p}_{i,0}$ as follows:

$$\vec{v}_{i,1} \coloneqq \pm\left(u_{i,y}, -u_{i,x}\right), \tag{4.2}$$

The first step for each field line i is calculated by progressing in a straight line along the direction of the vector $\vec{v}_{i,1}$ for a user-defined length δ (electric field line segment size in number of pixels):

$$\vec{p}_{i,1} = \vec{p}_{i,0} + \delta \frac{\vec{v}_{i,1}}{\|\vec{v}_{i,1}\|}. \tag{4.3}$$

Note that equation (4.2) has two solutions ($\pm$), one points towards the annulus and the other points away from the annulus. The correct solution is chosen by selecting the resulting point $\vec{p}_{i,1}$ within the annulus. Since the inner and outer boundaries are closed lines, the annulus may be defined by the set difference of the areas they enclose (outer boundary minus inner boundary). In ImageJ the contours are defined as regions of interest (ROI) and the annulus is computed as their set difference (XOR function).

3. As the magnitude of the electric field drops with the second power of the distance from the contributing charges, the electric field at point $\vec{p}_{i,k}$ is approximated by only considering charges that are close to point $\vec{p}_{i,k}$. The contributing charges are selected as the sequences of points on the two boundary contours within a user-defined 'electric field search window' around index i (Figure 4.1a):

$$J \coloneqq j \bmod I,\ j = i - w, \dots, i, \dots, i + w,\ w \in \mathbb{N}, \tag{4.4}$$

where w is a user-defined integer that defines the window size. The modulo operator mod enables closing the contours by selecting points at either of the ends of the sequence $j = i - w, \dots, i, \dots, i + w, w \in \mathbb{N}$, when i is too small ($i < w$) or too large ($i + w > I$). The electric field line i is built upon the first computed point $\vec{p}_{i,1} = \vec{p}_{i,0} + \delta \frac{\vec{v}_{i,1}}{\|\vec{v}_{i,1}\|}$ as a sequence of straight electric field line segments of length δ:

$$\vec{p}_{i,k} = \vec{p}_{i,k-1} + \delta \frac{\vec{E}(\vec{p}_{i,k})}{\|\vec{E}(\vec{p}_{i,k})\|}, \tag{4.5}$$

where

$$\vec{E}(\vec{p}_{i,k}) = \sum_{j \in J} \hat{q}_j \frac{(\vec{p}_{i,k} - \vec{p}_{j,0})}{\|\vec{p}_{i,k} - \vec{p}_{j,0}\|^3} - \sum_{j \in J} \hat{q}'_j \frac{(\vec{p}_{i,k} - \vec{p}'_{j,0})}{\|\vec{p}_{i,k} - \vec{p}'_{j,0}\|^3}, \tag{4.6}$$

and integer $k = 2 \dots K$ is the index of the electric field line point $\vec{p}_{i,k}$, $\vec{p}_{j,0}$ are the locations of the contributing charges $\hat{q}_j$ on the inner boundary contour, $\vec{p}'_{j,0}$ the locations of the contributing charges $\hat{q}'_j$ on the outer boundary contour. Finally, $\hat{q}_j$ and $\hat{q}'_j$ are arbitrary positive electric charges. This process is repeated until either the electric field line reaches the opposite boundary or the user-defined maximum number of steps K_{max} has been reached. Whether the field line has reached the outer boundary, is checked by verifying if $\vec{p}_{i,k}$ is within the annulus or on/beyond the outer boundary contour. Note that Equation (4.4) defines a window of $2w + 1$ points on the inner (and outer) boundary contours that slides with the index i of the respective field line. This selection includes neighbouring points across the contours, which are not necessarily the closest points to $\vec{p}_{i,k}$ (when measured by their Euclidean distance within the plane). This might lead the algorithm to fail with 'stray' or 'collapsed' field lines due to inaccurate approximation of the electric field, as discussed next. To improve the estimation, one can increase w at the expense of computational time.

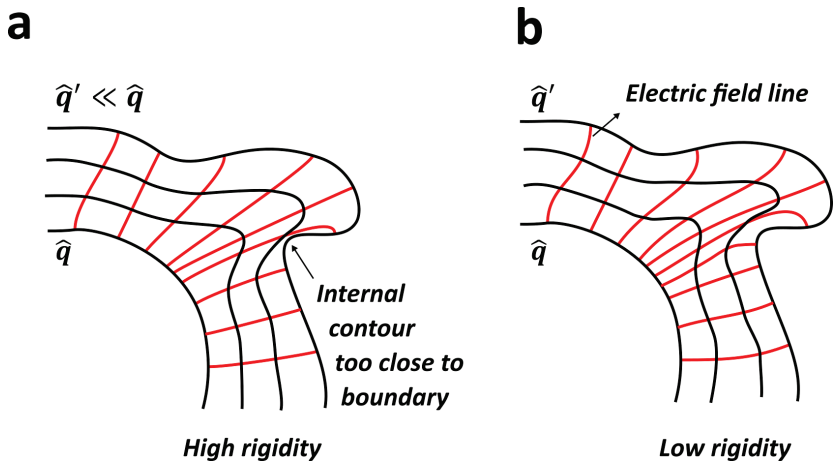


Figure 4.2: Electric field line rigidity

The rigidity of the electric field lines is defined as the ratio between the electrical charges on the outer and inner boundaries. Adjusting the rigidity yields stiffer or more flexible field lines that are useful to delineate contours of irregular cylindrical shapes that are heavily deformed. (a) High rigidity of the electric field lines results in stiff electric field lines that do not properly follow the curvature of the outer boundary contour. (b) Lower rigidity allows the electric field lines to be more flexible, pushing the intermediate contours away from the boundaries. This provides internal contours that better capture the irregular shape of the deformed tube.

A beneficial bi-product of using electric fields for the definition of the internal contours is the control it provides over the curvature of the electric field lines. This may be tuned via appropriate selection of the relative amplitude of the charges on the tube's boundaries, a parameter that may

be called rigidity (Figure 4.2). In practice this is achieved by defining $\hat{q}_j = 1$ for all charges on the inner boundary and $\hat{q}'_j = -rigidity$ (user-defined real value) for the charges on the outer boundary. This essentially weights the contributions to the electric field defined in equation (4.6), provided by the charges on the outer boundary ($\hat{q}'_j$) vs. the contributions of the charges on the inner boundary ($\hat{q}_j$). Tuning the rigidity results in straighter or more curved electric field lines, thus providing flexibility to capture irregular cylindrical shapes that are heavily deformed (see Figure 4.2b).

The computational approach to iteratively approximate the electric field lines described above, reduces the computational cost to calculate the exact value of the electric field at all possible positions in space, as per Equation (4.1). This approximation, involving (i) a discrete, user-defined electric field line segment size δ for progressing through the annulus, (ii) a user-defined electric field search window ($J = j \bmod I,\ j = i - w, \dots, i, \dots, i + w,\ w \in \mathbb{N}$) encompassing a limited number of contributing charges, and (iii) manual tuning of the relative electric charge magnitudes or rigidity might provide 'stray' or 'collapsed' electric field lines that do not reach the opposite boundary, as shown in Figure 4.1a. Such field lines are identified and removed as follows: 'Stray' lines are defined/identified as exceedingly long lines, and they are rejected by introducing a maximum number of iterations for the computation of each electric field line: $K \leq K_{max}$. 'Collapsed' field lines are defined as lines that do not progress far from the initial point $\vec{p}_{i,0}$, and are identified as the lines with variance V_i less than a user-defined minimum electric field line variance V_{min}:

$$V_i = \frac{1}{K}\sum_{k=0}^{K}\left(\vec{p}_{i,k} - \overrightarrow{\overline{p_i}}\right)^T\left(\vec{p}_{i,k} - \overrightarrow{\overline{p_i}}\right), \qquad (4.7)$$

where

$$\overrightarrow{\overline{p_i}} = \frac{1}{K}\sum_{k=0}^{K}\vec{p}_{i,k}\,. \qquad (4.8)$$

Each electric field line i is then represented by a cubic spline fitted through all points $\vec{p}_{i,k}$ ($k = 0 \dots K$). Note that once the electric field lines are transformed form a collection of points $\vec{p}_{i,k}$ to cubic splines, the notation $\vec{p}_{i,k}$ is hence abandoned and the cubic splines are resampled as follows: Firstly, the local thickness d_i at $\vec{p}_i$ is defined as the length in number of pixels of the electric field line i. Internal contours are then created at a total of C (user-defined) depth levels at incremental percentages (0% - 100%) of each individual electric field line's length d_i (see Figure 4.1b) as follows: Relative depth levels are defined as $\frac{c}{C-1} \cdot 100\%$ and absolute depth levels as $\frac{c}{C-1} \cdot d_i$ ($c = 0,1, \dots, C-1$) with c, C being integers. Explicitly, for $C = 3$: $\frac{0}{2} \cdot 100\% = 0\%$ (inner boundary), $\frac{1}{2} \cdot 100\% = 50\%$, $\frac{2}{2} \cdot 100\% = 100\%$ (outer boundary). See Figure 4.1b for an example of

four (C = 4) different relative depth levels (0%, 33%, 67%, 100%). The number of relative depth levels C determines the number of equidistant, non-overlapping contours (including the internal boundary at 0% and the external boundary at 100% relative depth level), which can be chosen arbitrarily. Each contour $cont_{h,c}$ is defined by fitting a cubic in-plane spline at height level h through all points on the electric field lines i corresponding to depth level c, or absolute depth levels $\frac{c}{C-1} \cdot d_i$. As some of the electric field lines may have been discarded ('stray' or 'collapsed' field lines) the resulting internal contours might consist of fewer points than the boundary contours. To obtain a uniform representation, each cubic spline representing one specific contour $cont_{h,c}$ is resampled with a total of I samples.

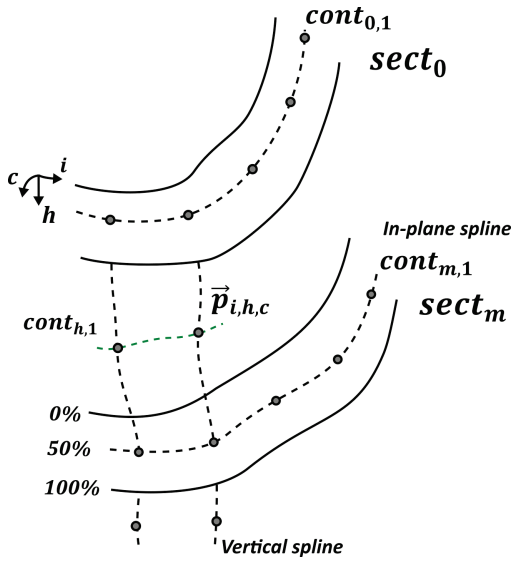


Figure 4.3: Interpolation of internal and boundary contours

The process shown in Figure 4.1 is performed on every m^{th} 2D section of the 3D image-stack depicting the deformed cylindrical tube, resulting in contours $cont_{h,c}$ on sections $sect_h$ ($h = 0, m, 2m, \ldots$) at depth level c. The remaining contours $cont_{h,c}$ on sections $sect_h$ ($h \neq 0, m, 2m, \ldots$) are generated by fitting vertical cubic splines (across all sections of the 3D image stack) through all field line points of each contour $cont_{h,c}$ at depth level c.

As the approximation of the electric field lines is the most computationally intensive part of this method to generate 3D cycloramas, it is only applied on every m^{th} slice or 2D section of the 3D image stack (m: interpolation interval) that depicts the deformed cylindrical tube, resulting in contours $cont_{h,c}$ on sections $sect_h$ ($h = 0, m, 2m, \ldots$ and $h \leq H - 1$ with H the total number of slices) at depth level c as shown in Figure 4.3. To generate the contours $cont_{h,c}$ on the remaining sections $sect_h$ ($h \neq 0, m, 2m, \ldots$), the contours of the selected 2D cross-sections are then interpolated across all sections of the 3D image stack. In detail, vertical cubic splines are fitted on

corresponding points, i.e. the i^{th} spline is created by selecting the i^{th} point on each contour for each depth level c (see Figure 4.3). Eventually $cont_{h,c}$ is defined as the set of all points at cross-section h and depth level c: $cont_{h,c} = \{\vec{p}_{i,h,c}\}$. Once contours have been defined at all depth levels c and for all 2D cross-sections h, 3D re-slicing or an onion-like surface at relative depth level $\frac{c}{C-1} \cdot 100\%$ is defined as the set of all 2D contours at depth level c. The basic components are then ready to generate internal 2D panoramic views of the deformed cylindrical sample and 3D cycloramas as explained in section '4.2.2 Mapping: Generation of 3D cycloramas'.

4.2.2 Mapping: Generation of 3D cycloramas

To generate 2D panoramic views at different relative depth levels $\frac{c}{C-1} \cdot 100$ % of the deformed cylindrical tube the 3D re-slicing surfaces need to be mapped onto 2D equivalents (2D cycloramas). This is done on a contour-by-contour basis, where the boundary and internal contours need to be mapped onto rectangular grids. When stacked together, these 2D panoramic views become a 3D cyclorama, which displays the boundaries and particularly the internal structure of the sample at different depth levels (see Figure 4.4a-c).

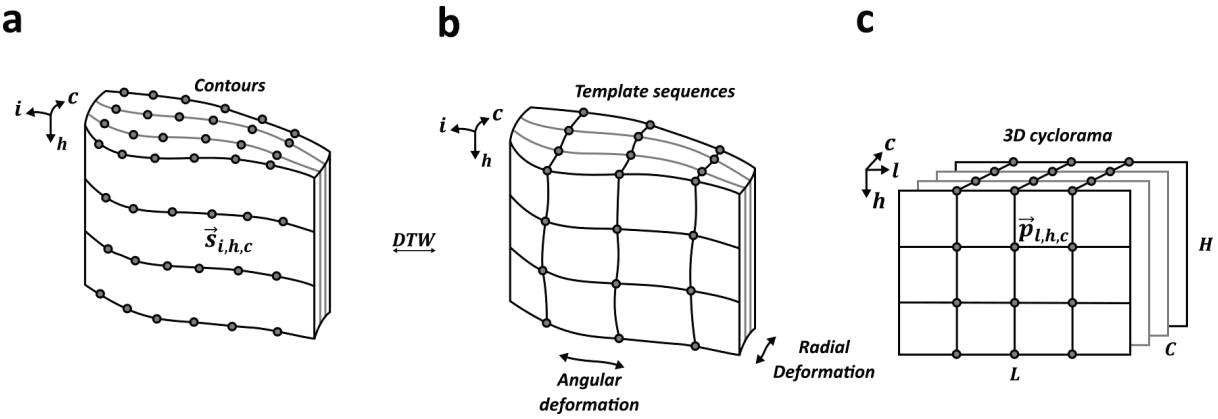


Figure 4.4: Mapping of internal and boundary contours onto a rectangular grid

3D cycloramas are created by selecting the appropriate points to probe the deformed cylindrical tube using a mapping computed by the dynamic time warping (DTW) algorithm. (a) Initial internal re-slicing surfaces are represented as unstructured internal contours that are sequences of points $\vec{s}_{i,h,c}$ (i: number of field line, h: section number of the 3D image-stack depicting the deformed cylindrical tube, c: depth level) at a certain depth level c, around the tube and across sections $sect_h$, among which the longest of all 2D contours $cont_{h,c}$ defines the length L of the 3D cyclorama. (b) DTW provides a mapping that selects the points in a. that correspond to the points in the structured grid shown in b. (c) The structured grid shown in b. is then used to probe the 3D image data and create the 3D cyclorama of length L. The

digital unrolling from a. to c. is achieved by introducing deformations in two directions: Angularly, by stretching and contracting the contours through DTW, and radially, through the definition of the absolute depth levels $\frac{c}{C-1} \cdot d_i$ using the length d_i of the electric field line i.

Initially, the length of all 2D panoramic views or 2D cycloramas is fixed at reference length L, defined by the longest contour in the volume: $L := \max_{h=0\ldots H-1} \left(\max_{c=0\ldots C-1} L_{h,c} \right)$ with integer $L_{h,c}$ being the length of contour $cont_{h,c}$ of section $sect_h$ at relative depth level c (see Figure 4.3) in number of pixels (see Figure 4.4c). The longest contour thus defines the length L of the 3D cyclorama, while all other contours ($L_{h,c} \leq L$) are dilated via one of two methods discussed later to match up with the longest contour. On this account, the resulting (non-uniform) pixel size of the 3D cyclorama will be equal (for the longest contour) or larger (all other contours) when compared to the pixel size of the original image data.

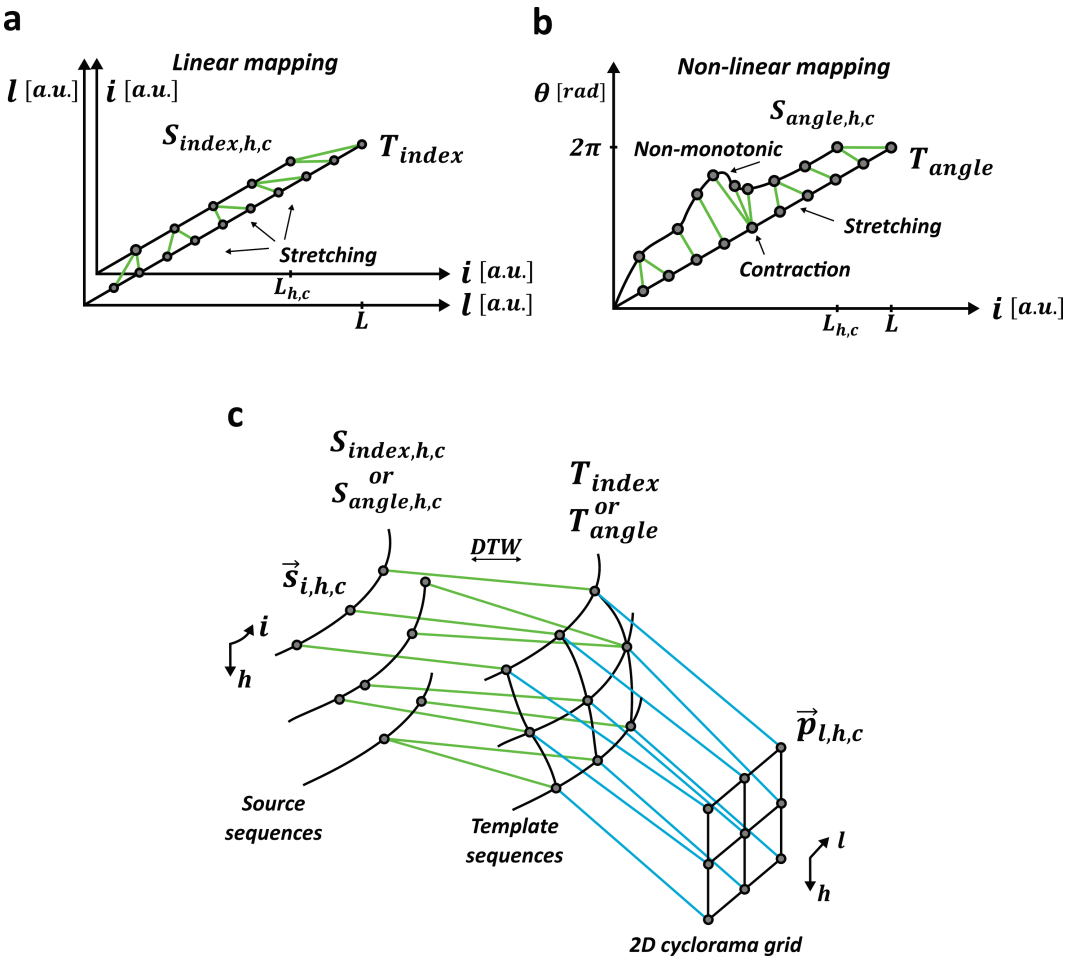


Figure 4.5: The dynamic time warping algorithm used to map individual contours

(a) Mapping the index sequence through the DTW algorithm is equivalent to linearly stretching the contour of length $L_{h,c}$ to fit the linear template sequence T_{index} of reference length L. (b) Mapping the angle sequence through DTW locally stretches or

contracts the angle sequence S_{angle} of length $L_{h,c}$ to fit the linear template sequence T_{angle} of reference length L. In sections where S_{angle} is not monotonic, several points of S_{angle} collapse to a singular point on the template, resulting in deformations of the cyclorama image. (c) 3D re-slicing surfaces are initially defined as sets of contours, which are sequences of points. After mapping on the template sequences, the contour points are restructured onto a rectangular grid representing the voxels of the 3D cyclorama. This grid is then used to create the 3D cyclorama by probing the 3D image stack at the corresponding points of the source sequences.

Mapping of the boundaries and internal contours onto 2D cycloramas or 2D panoramic views (with a fixed length L) is performed using the dynamic time warping (DTW) algorithm [251] (see Figure 4.5). The DTW algorithm locally stretches or contracts the time sequence of a time signal, a process called 'warping'. It matches each time point of a source time sequence S with a point onto a template time sequence T and *vice versa* so that a selected metric (Euclidean distance here) between the two sequences is minimised. One can imagine two runners that move across a path with different and variable speeds. The position of the two runners as a function of time may then be represented by time sequences. DTW may be used to warp the position of runner A so that they are next to runner B at any given time. This is done by forcing runner A to go faster or slower at different sections of their path, i.e. warping the time sequence of runner A onto the template sequence of runner B. Here, the FastDTW [251] algorithm has been applied, (specifically the Java implementation that is freely available at https://github.com/rmaestre/FastDTW), which approximates the original DTW algorithm to achieve high computational performance at reduced computational time. Note that although DTW was originally developed for time sequences, the sequences can actually represent any arbitrary set of elements upon which a metric or distance function can be defined. In this work, the time sequence is replaced with two alternative choices as discussed next, namely either an 'index' or an 'angle' sequence to yield two different mapping modes.

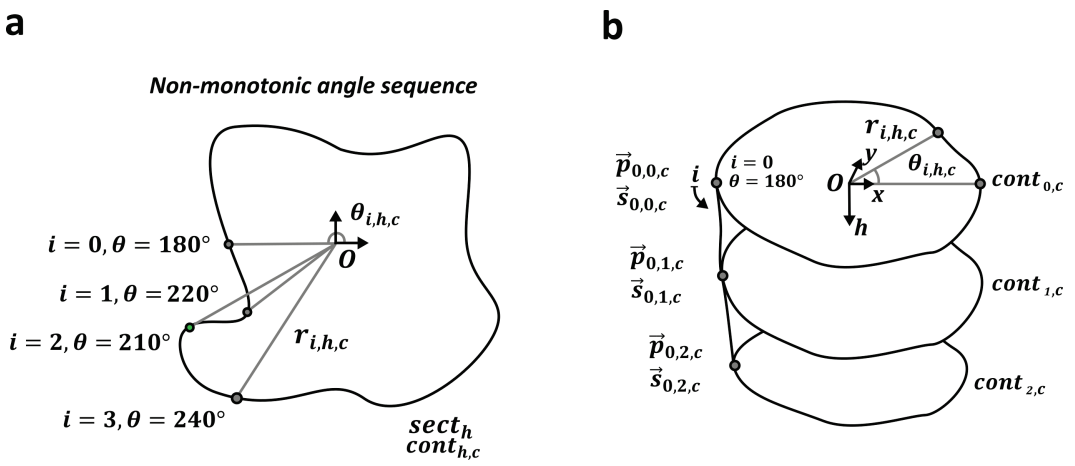


Figure 4.6: Selection of the re-slicing surface edge using an arbitrary axis of rotation

(a) An arbitrary axis of rotation O is selected inside the tube and all contour points are expressed in polar coordinates with respect to that axis. The resulting angle sequence depends on the position of the axis of rotation and the shape of the contour. In the specific case shown here the cavity in the contour results in a non-monotonic angle sequence: (180°, 220°, 210°, 240°). (b) The edge of the re-slicing surface at depth level c is then defined by selecting the first point $\vec{p}_{i,h,c} = \vec{p}_{0,h,c} = (r_{i,h,c}, 180°)$ of each of its constituting contours $cont_{h,c}$ with angle θ=180°. Re-slicing surfaces are then represented either in a Cartesian or polar coordinate system as $\vec{p}_{i,h,c}$ or $\vec{s}_{i,h,c}$, respectively.

Before applying the DTW algorithm to the contours, it is necessary to turn the closed contours into 'time' sequences. First, each of the contour splines is resampled so that contour $cont_{h,c}$ is (densely) represented by a total of $L_{h,c}$ points. Then, the first point $\vec{p}_{0,h,c}$ (point with index $i = 0$), is arbitrarily but consistently selected across all contours $cont_{h,c}$ as follows (see Figure 4.6): The points of each contour $cont_{h,c}$ that are initially defined in Cartesian coordinates $\vec{p}_{i,h,c} = (p_{i,h,c_x}, p_{i,h,c_y})$ can then be expressed in polar coordinates:

$$\vec{s}_{i,h,c} = \left(r_{i,h,c}, \theta_{i,h,c}\right), \quad i = 0 \dots L_{h,c} - 1\,, \tag{4.9}$$

with $O = (o_x, o_y)$ an arbitrary rotation axis origin inside the tube, $r_{i,h,c}$ the distance from the axis of rotation, and $\theta_{i,h,c} \in [0 \dots 2\pi)$ the azimuth angle in radians. The first point $\vec{p}_{0,h,c}$ or $\vec{s}_{0,h,c}$ is arbitrarily chosen as the point among all points on contour $cont_{h,c}$ with $\theta_{i,h,c}$ closest to π radians or 180 degrees (see Figure 4.6b). Two alternative source sequences are then defined that play the role of time as discussed beforehand. These are the 'index' sequence

$$S_{index,h,c} := (i),\ i = 0 \dots L_{h,c} - 1\,, \tag{4.10}$$

which is linear and monotonically increasing, and the 'angle' sequence

$$S_{angle,h,c} := \left(\theta_{i,h,c}\right),\ i = 0 \dots L_{h,c} - 1\ , \qquad (4.11)$$

which depends on the shape of the specific contour, and is not necessarily linear or monotonic (see Figure 4.6a). Using these sequences two different mapping modes were implemented, 'linear' and 'non-linear' mapping mode (see Figure 4.5a&b). For the linear mapping mode, the index sequence $S_{index,h,c}$ is mapped using the DTW algorithm onto a linear template sequence T_{index} with size L of the longest contour $cont_{h,c}$:

$$T_{index} := (l),\ l = 0 \dots L - 1\ . \qquad (4.12)$$

Note that in this linear mapping mode the use of the DTW algorithm is not essential as it is equivalent to uniformly sampling each boundary and internal contour with the same number of points as the longest/reference contour (see Figure 4.5a). In other words, all contours, except from the longest/reference contour, are stretched linearly to fit the fixed 3D cyclorama length L. For the 'non-linear' mode (see Figure 4.5b), the angle sequence of each contour is mapped (warped) through the DTW algorithm onto a linear template sequence with size L equal to the length of the longest contour:

$$T_{angle} := \left(\frac{l}{L} 2\pi\right), l = 0 \dots L - 1\ . \qquad (4.13)$$

As the non-monotonic and non-linear contour angle sequences $S_{angle,h,c}$ are mapped onto the linear template sequence T_{angle}, the DTW algorithm non-rigidly warps the angle sequences to minimise the Euclidean distance between the warped and the template sequences. The DTW algorithm assumes that time is monotonically increasing. Thus, when non-monotonic source sequences are mapped, the sections, where a source sequence (here $S_{angle,h,c}$) goes backwards, collapse to a singular point on the warped sequence, resulting in a contraction (see Figure 4.5b). This property of the DTW algorithm gives raise to the two different mapping modes employed here: The linear mapping mode maps the monotonically increasing index sequence on the monotonically increasing template sequence. This avoids collapsing of points and contraction, however it may result in 'wavy' cycloramas as each contour is stretched differently, depending on its length $L_{h,c}$ to fit the reference length L (see Figure 4.7a). The non-linear mapping mode, on the other hand, locally contracts the angle sequence at the sections where it goes backwards, resulting in local deformations in cycloramas and more uniform sampling across the rest of the image (see Figure 4.7b). Both mapping modes have a uniform pixel size across the vertical dimension (h-dimension in Figure 4.7) but their pixel size across the horizontal dimension (l-dimension in Figure 4.7) is different to the vertical one. The linear mapping mode results in a

uniform pixel size along individual lines (horizontal dimension) of the cyclorama image but not necessarily the same pixel size across all lines, as each contour is stretched individually. As the mapping is non-rigid for the non-linear mapping mode, the pixel size also varies in the horizontal dimension along individual lines (from pixel to pixel). This results in minimal deformation at sites away from singularities, at the expense of singularity patches as shown in Figure 4.7b. On the other hand, the linear mapping mode spreads the deformation across the entire image, resulting in wavy but continuous cycloramas as shown in Figure 4.7a.

At the end of the mapping process, the DTW algorithm yields a mapping of the indices in the source sequence S onto the template sequence T, by defining which index i in equation (4.10) or equation (4.11) corresponds to each index l in equation (4.12) or equation (4.13), respectively:

$$i \leftrightarrow l,\ i \in \{0 \dots L_{h,c} - 1\},\ l = 0 \dots L - 1\,, \qquad (4.14)$$

These source indices are then used to redefine the contours by selecting the points $\vec{s}_{i,h,c}$ that correspond to each point $\vec{p}_{l,h,c} = (l, h, c)$ in the 3D cyclorama (see Figure 4.5c). The mapping between the 3D surface, defined by all points with fixed depth level $\check{c}$, onto points on the 2D cyclorama plane is then defined as (see Figure 4.5c):

$$\vec{s}_{i,h,\check{c}} \leftrightarrow \vec{p}_{l,h,\check{c}}\,. \qquad (4.15)$$

One (planar) 2D cyclorama per depth level $\check{c}$ is created by probing the 3D image data volume at all points $\vec{s}_{i,h,\check{c}}$ (see Figure 4.5c). Note that all 2D cycloramas, which correspond to a specific relative depth level, are of equal size (equal length and height). This allows to create a stack of 2D cycloramas from 0 % (internal boundary) to 100 % (external boundary) of the tube's relative depth level (3D cyclorama), which is essentially an unrolled version of the deformed cylindrical tube (murine colon). This digital unrolling onto a 3D cyclorama has been achieved by introducing image deformations in two directions: Radially, through the definition of the relative depth levels using the electric field lines (see Figure 4.4), and angularly, by stretching and contracting the contours through the DTW algorithm (see Figure 4.5).

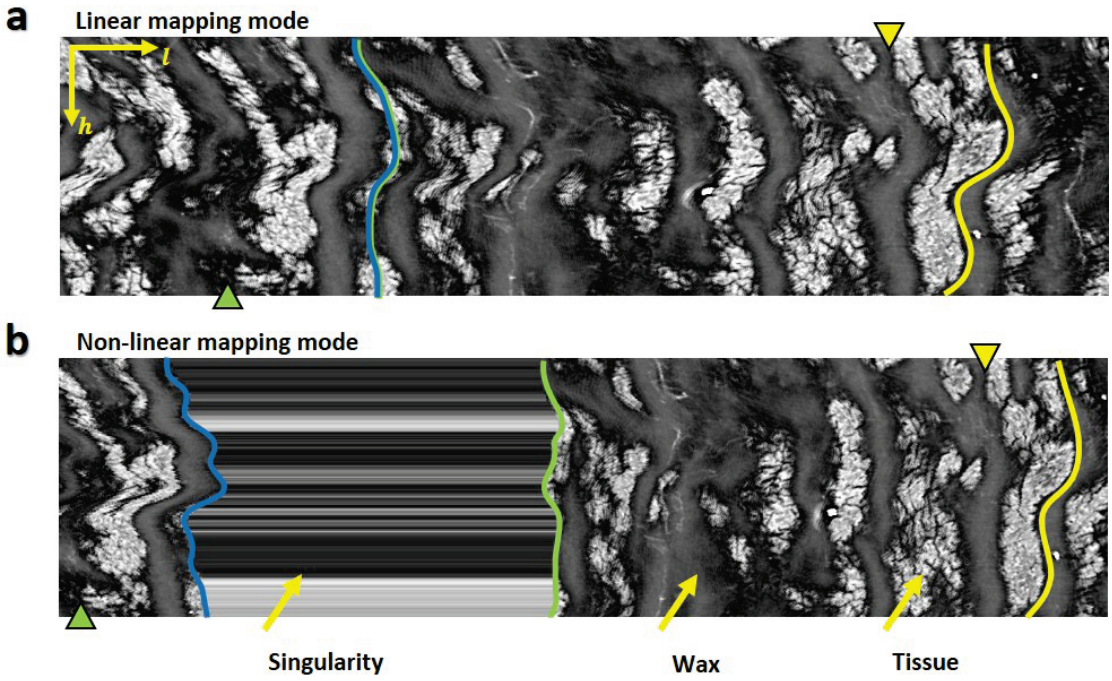


Figure 4.7: Linear and non-linear mapping modes providing different 2D cycloramas

Sections of 2D cycloramas at 0% relative depth level (inner boundary) of a healthy murine colon embedded in paraffin wax, which partially probe volumes from the background (wax) and the sample (tissue). (a-b) Both cycloramas are created mapping the same 3D re-slicing surface onto two different cycloramas using the a. linear and the b. non-linear mapping mode. The triangles mark corresponding points on the two cycloramas, revealing that they are deformed (stretched/contracted) differently with anisotropic pixel sizes. Linear deformation is apparent in both modes as vertically curved patterns, notably more pronounced in the a. linear than the b. non-linear mapping mode, as revealed by the yellow outline of the tissue. (b) A characteristic pattern ('Singularity') appears for the non-linear mapping mode. When a single point on the source sequence is mapped onto several points of the template sequence (stretching), the singular point is being repeatedly probed to create all corresponding pixels on the cyclorama ('Stretching' in Figure 4.5b), resulting in a line with a constant grey value. Notice that the blue (left) and green (middle) outlines coincide for the linear mapping mode, but not for the non-linear mapping mode. This shows that no information is lost during the mapping processes.

4.3 Discussion

The 3D cyclorama method described in section '4.2 Development' was purposefully designed to be generic so that it is not constrained to murine colon unrolling. This makes it applicable to digital unrolling of a wide range of samples imaged with different imaging modalities. Yet, developing a tool with such a broad use-case potential requires: i. thorough validation to verify that it works as expected in the general case (see Appendix 'A.1 Validation using a digital phantom') and ii. a means to make it available to the public. Thus, 3D cyclorama was adapted into an Imagej/Fiji plugin available through the Fiji update sites (see https://imagej.github.io/plugins/3d-cyclorama).

Two case studies irrelevant to unrolling of biological samples are presented in Appendix 'A.2 Interdisciplinary applicability: Case studies'. The first showcases the capability of 3D cyclorama for digital flattening of deformed sheets (Appendix 'A.2.1 Digital flattening of a pharmaceutical film'). The second showcases its capability to digitally unroll ancient scrolls (Appendix 'A.2.2 Virtual unrolling of an ancient lead amulet') through an ongoing international collaboration with the Australian Centre for Neutron Scattering, the MACQUARIE university museum of ancient cultures (MAC), and scholars at Rutgers University and the University of Cambridge. The project aims to combine neutron computed tomography and digital unrolling through the 3D cyclorama method to decipher, and translate ancient inscribed lead scrolls, unattainable by now.

Section '4.3.1 Limitations and outlook' discusses the limits and pitfalls of 3D cyclorama and section '4.3.2 Applicability of colonic cycloramas to the study of colorectal cancer' discusses issues specific to the study of colorectal cancer.

4.3.1 Limitations and outlook

The 3D cyclorama method can facilitate the visualisation of the internal structure of objects with a peculiar geometry, yet the projected volumes should be interpreted with care. The 3D cyclorama is deformed in both angular and radial directions in a non-uniform manner (angular and radial deformation, respectively), thus the scale is not preserved, resulting in anisotropic and non-uniform pixel sizes. This uneven scale implies that distances should not be measured on cycloramas. Nevertheless, the sample topology is always preserved, i.e. neighbouring features in the original volume will always remain neighbours in the projected space and *vice versa*. This originates in the construction of the re-slicing surfaces and the mapping: The re-slicing surfaces are created based on (virtual) electric field lines that never overlap. The dynamic time warping (DTW) algorithm is then used to warp each of the contours that constitute the re-slicing surfaces, which ensures that the warped contours are monotonic. This mapping approach offers two

benefits. First, each contour is mapped through the DTW algorithm, independent of each other, allowing implementation of the algorithm in a multi-threaded fashion to improve computational performance. Secondly, it creates cycloramas identical in size and shape, readily solving the subsequent registration problem, i.e. the problem of aligning 2D cycloramas from different depths of the tube, so that features extending over several layers match up and are represented as continuous entities within the 3D cyclorama. This unavoidably introduces image deformations: The non-linear mapping mode results in minimal deformation at sites away from singularities, at the expense of singularity patches as shown in Figure 4.7b. On the other hand, the linear mapping mode spreads the deformation across the entire image, resulting in wavy but continuous cycloramas as shown in Figure 4.7a. Thus, the user can choose if the linear or non-linear mapping mode is more appropriate for a specific application.

It is worth to investigate the possibility of employing alternative mapping strategies not addressed yet. Notably, the employed approach does not map a re-slicing surface at a time, but it maps each contour individually, building up each 2D cyclorama as a collection of mapped contours. Although this allows each contour mapping to be implemented in a separate computational thread, it confines the ability to impose certain properties on the 2D cycloramas that result from the mapping, such as minimal deformation, maximal smoothness, etc. Approaches such as manifold embedding could be used to impose such desired properties for the mapping of the 3D re-slicing surfaces. Such methods are studied in the field of differential geometry [252], where manifolds (topological spaces that are locally Euclidean) can be embedded in lower-dimensional equivalents. In the cyclorama case, the dimensionality of the re-slicing surface needs to be reduced from 3D to 2D. This mapping (or embedding) can be done in different ways, such as isometric mapping [253] for minimal deformation, or moving least squares [254] for maximal smoothness, depending on the desired properties of the target space. For instance, isometric mapping [253] computes the embedding that optimally preserves geodesic distances between the points on the manifold. This algorithm would preserve the shapes of features on the re-slicing surface, effectively minimising deformation of feature shapes on the mapped space. However, isometric mapping would not ensure consistency among the shape of all 2D equivalents since it does not necessarily yield rectangular images. This means that each 3D re-slicing surface would be mapped onto a 2D equivalent with different dimensions and shape. Hence, a subsequent step of registration among the 2D equivalents would be necessary to match up features in subsequent 2D surfaces, to combine them into a 3D stack depicting the unrolled volume. Instead, the adopted approach solves this problem by imposing identical shape and size for 2D cycloramas that are stacked to produce a 3D cyclorama, at the expense of the deformation types discussed in section ‘4.2.2 Mapping: Generation of 3D cycloramas’. Similarly to the isometric mapping algorithm, the moving

least squares approach [254] computes a mapping that is optimal in some respect, in that it imposes smoothness of the mapped surface. This method requires the user to provide a small set of control point pairs (a point on the source surface and its mapped position) and creates a smooth surface that respects the given set. This method, unlike isometric mapping, would be suitable to create rectangular 2D surfaces of identical size and shape. It is an elegant and mathematically sound method that is worth investigation and integration with the 3D cyclorama approach of reforming the deformed tube into onion-like re-slicing surfaces using electrostatics: The requirement of an input control point pairs set in the formulation of the moving least squares approach could be conveniently integrated with the 3D cyclorama method given the definition of re-slicing surfaces as a set of contours made up of sequences of points. Control point pairs could be created by matching up points $\vec{s}_{i,h,c}$ from the re-slicing surface (see Figure 4.4a) with points $\vec{p}_{i,h,c}$ on the 3D cyclorama grid (see Figure 4.4c). Selecting a small set of control point pairs would provide enough flexibility for image deformation in both directions h and l (see Figure 4.4a) in addition to the radial deformation (direction c), already introduced through the definition of the relative depth levels (see Figure 4.4). This flexibility of deformation in all directions would result in a smooth 3D deformation field (smoothness in directions h and l is imposed by the moving least squares method and in direction c is imposed by the electric field lines), avoiding singularities such as those shown in Figure 4.7b.

4.3.2 Applicability of colonic cycloramas to the study of colorectal cancer

The 3D cyclorama method developed here advances time-consuming traditional (2D) histology via the use of re-slicing onion-like surfaces, thus obtaining non-planar slices across the entire tissue block. This internal re-slicing, which follows the 3D shape of the deformed tube, is impossible in standard histology due to the nature of the involved (planar) physical sectioning. It enables all crypts within the tissue section to be visualised on a single image, providing an overview of the sample. The 3D cyclorama method facilitates studying the 3D morphology of the colon in its natural 3D shape. Since, there is no need of cutting the colon open and/or coiling it, tissue distortions through physical cutting typical for conventional histology [123] are absent.

Note that the 3D cyclorama method offers a great degree of adjustability when it comes to the contour length, the field lines rigidity, the mapping mode etc. Though, most of the available freedom is unnecessary when it comes to unrolling murine colons to identify budding crypts. It becomes apparent in section ‘4.4.1 Digital unrolling of murine colons’ and section ‘5.1 Workflow for identification and 3D segmentation of budding’ that in order to identify budding crypts it is enough that subsequent 2D cycloramas (image slices of the 3D cyclorama) are properly registered. That is, the cross-sections of each colonic crypt are matched-up between one 2D

cyclorama and the next. This is achieved by design in the 3D cyclorama method as discussed in section '4.3.1 Limitations and outlook'. Therefore, contours that capture a rough outline of the sample (not exactly following its outline) and the simpler, 'index' mapping mode are suitable for unrolling of murine colons even if this yields wavy cycloramas (see Figure 4.8 and Figure 4.9b).

4.4 Results

4.4.1 Digital unrolling of murine colons

From the CT stacks obtained in section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons', a subset of 301 reconstructed CT slices (height of 0.66 mm) were extracted, each 1841 × 1841 pixels^2 or 4.05 × 4.05 mm^2 large (see Figure 4.8a), and unrolled using the linear mapping mode with $C = 30$ depth levels, $I = 200$ point-long contours, an electric field search window of 31 points ($w = 15$), electric field line segment size $\delta = 2$ pixels, maximum number of steps $K_{max} = 75$, minimum electric field line variance $V_{min} = 50$, rigidity equal to 1, interpolation interval $m = 20$ and rotation axis origin $O = (1080, 780)$. The resulting 3D cyclorama's length was about $L \approx 2500$ pixels (varies between scans), the height $H = 301$ pixels, and the depth $C = 30$ slices.

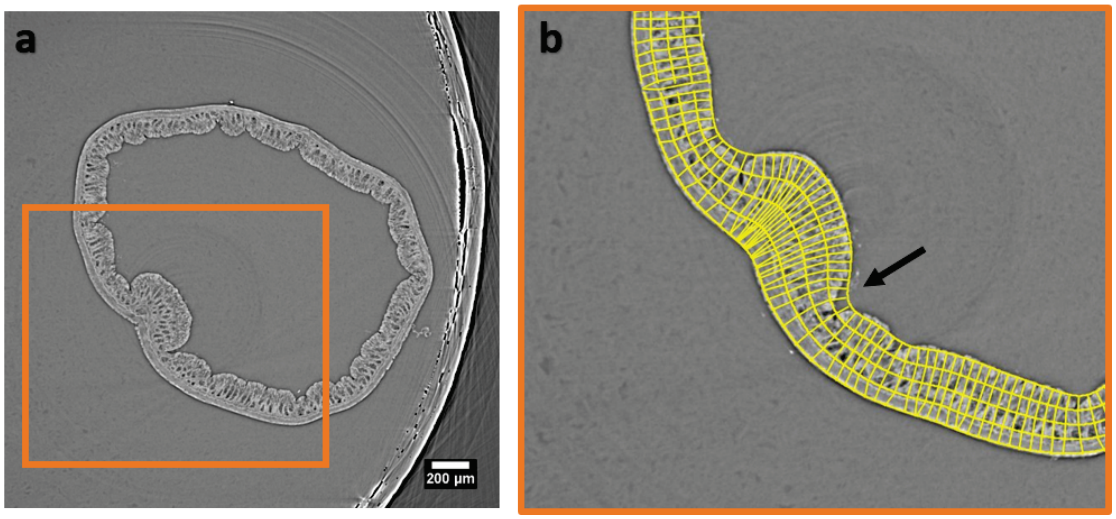


Figure 4.8: Contours on a murine colon

(a) A µCT section of a healthy murine colon embedded in paraffin wax. (b) Inset corresponding to the rectangle on a. shows how contours ($C = 5$ instead of the total $C = 30$ depth levels used are shown here for visual clarity) follow the natural shape of the sample. The electric field lines, used to introduce the concept of local thickness (length of electric field lines), are perpendicular to the contours. The arrow shows a

location where the boundary contour does not exactly follow the sample's boundary in order to minimise the resulting angular deformation of the 3D cyclorama.

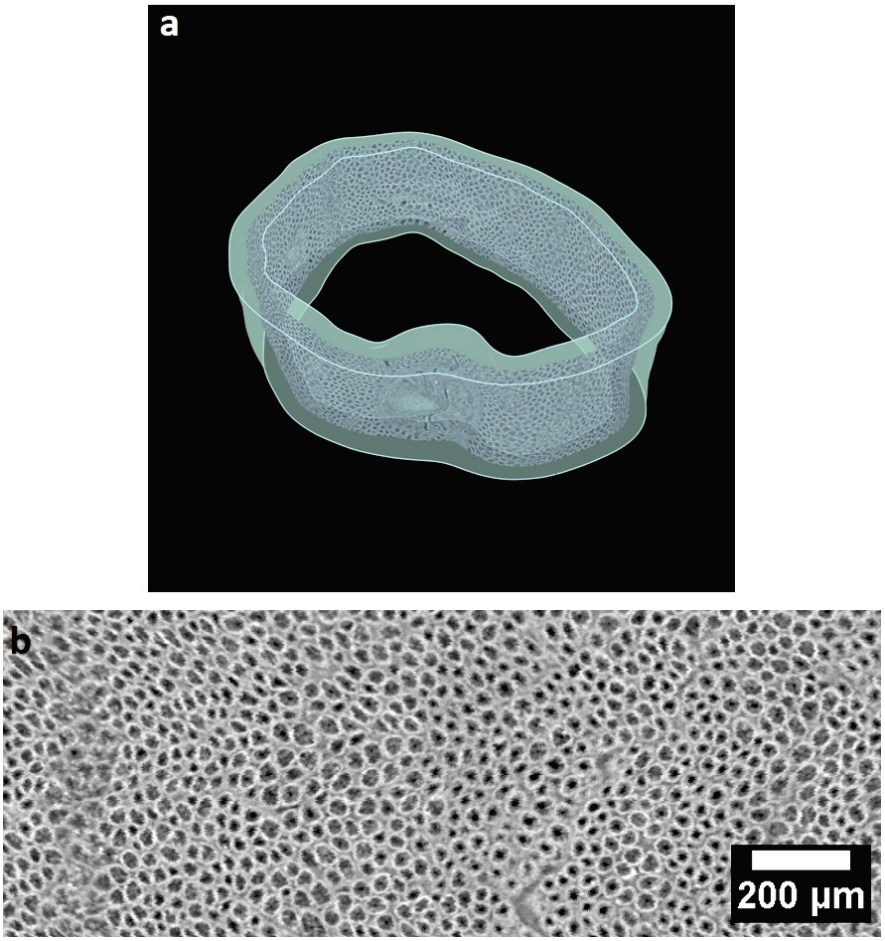


Figure 4.9: Re-slicing surfaces and 2D cyclorama of a healthy murine colon

(a) Boundary re-slicing surfaces illustrated as transparent grey facades and the onion-like re-slicing surface at 50% relative depth level with grey values probed from the 3D image stack passing through the µCT volume of the sample. (b) A cropped slice of the 3D cyclorama, i.e. a cropped 2D cyclorama, of a murine colon at 50% relative depth level renders the colonic crypt cross-sections as black circular profiles and the crypt epithelium appears grey. As this re-slicing surface is relatively deep into the tissue, it has a small curvature, which makes it 'easy' to unroll. This results in a 2D cyclorama that is not exceedingly 'wavy', in contrast to those closer to 0% relative depth level (see also Figure 4.7).

The inner and outer boundary contours were drawn manually using the ImageJ segmented-line ROI tool so that they roughly capture the shape of the colon. By avoiding to follow every fold of the colonic mucosa a smoother inner contour was created (see Figure 4.8b), and consequently a smoother inner re-slicing surface, which results in minimal deformation on the resulting

cyclorama. Figure 4.8b shows an example of the resulting electric field lines and contours (only $C = 5$ depth levels are shown here for visual clarity). Re-slicing surfaces created using these contours pass throughout the sample, slicing the colonic crypts perpendicularly. Figure 4.9 provides a 3D rendering that illustrates the sample's boundaries and the re-slicing surface at 50 % relative depth level, yielding the probed CT greyscale values. The re-slicing surfaces at each relative depth level were mapped onto planar 2D cycloramas that were combined into a 3D cyclorama. Figure 4.9b shows a segment of the 2D cyclorama at the same relative depth, which, being locally perpendicular to the colonic crypts, depicts the colonic crypts as circular profiles. The 3D cyclorama allows to trace each crypt through all tissue layers and hence, to identify crypt budding in 3D (see section '5.1 Workflow for identification and 3D segmentation of budding').

Chapter 5 Tools and workflows for 3D morphological quantification of budding structures

In Chapter 2 and Chapter 3, sample preparation and imaging protocols were presented, designed to preserve and depict the native shape and structure of the murine colons in 3D. This is an important prerequisite for accurate quantification of the microstructural characterisation of murine colonic crypts that follows in this chapter and Chapter 6. In Chapter 4, the method has been presented to digitally unroll colons onto 3D cycloramas, which facilitates the identification of budding crypts. The purpose of this chapter (Chapter 5) is to outline the process and tools developed for detailed 3D morphological quantification of the budding crypts that have been imaged by μCT and unrolled via the 3D cyclorama method.

In section 5.1, the workflow for identification and 3D segmentation of crypt budding using 3D cycloramas is outlined. In section 5.2, the morphological characteristics are defined, and the methodology for estimating them is described in detail. Finally, the procedure for statistical analysis of the estimated measurements is described in section 5.3. The computational methods described in section 5.1 were implemented in Fiji/ImageJ [229] using the Jython scripting application programming interface (API) [255]. The morphological quantification of crypts in section 5.2 and the statistical quantification in section 5.3 have been implemented separately in Python. All raw data and processing scripts are available online (see supplementary material S2 Dataset supporting the work presented in Chapters 2, 3, 4 & 6, and Appendices A & B DOI: 10.5258/SOTON/D1565).

5.1 Workflow for identification and 3D segmentation of budding crypts

To identify and segment the crypts, the workflow described next, and shown in Figure 5.1, was applied on the μCT stacks obtained in section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons'. Specifically a subset of 301 reconstructed CT slices (height of 0.66 mm) were extracted from each stack, each 1841×1841 pixels^2 or 4.1×4.1 mm^2 large (see Figure 5.2). The mucosal layer of the colons depicted in the μCT stacks were unrolled with the 3D cyclorama method (see Chapter 4). Budding crypts were identified within the 3D cycloramas, and the inverse mapping was used to identify and segment the crypts in 3D in the μCT data. The resulting segmented crypts were isolated and aligned into individual segmented stacks.

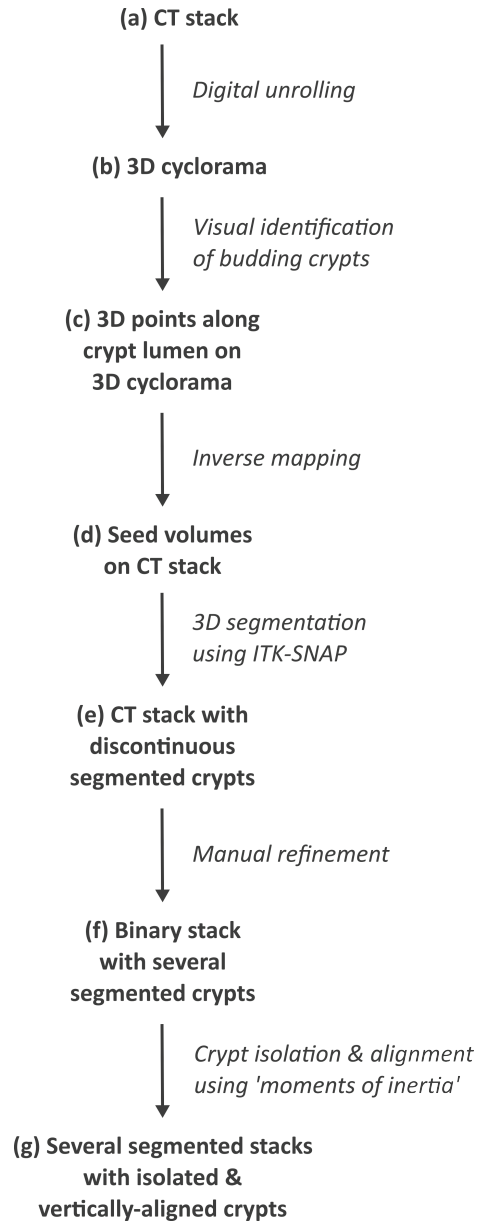


Figure 5.1: Crypt identification, 3D segmentation, and alignment workflow

µCT stacks of the intact murine colons (a) were unrolled using the 3D cyclorama method (b), and budding crypts were identified and marked with 3D points along their lumen (c). Those points were back-mapped using the inverse mapping between the 3D cyclorama and the µCT data to create seed volumes on the CT stack (d), which were used to initialise a region growing algorithm in ITK-SNAP to segment the crypts in 3D (e). The resulting discontinuous segmented crypts were manually refined, resulting in a binary stack with several segmented crypts (f). Those were isolated and aligned using the 'Moments of inertia' tool of BoneJ, resulting in several segmented stacks, each containing one isolated and vertically-aligned crypt (g).

5.1.1 Digital unrolling

The mucosal layer was defined in the µCT stacks by manually drawing the internal (a) and external (b) boundary contours for the 3D cyclorama method, as shown in Figure 5.2. Additionally, contour (c) was drawn to define the interface between the sample and wax, which was used to measure the colonic wall thickness and area as discussed in Appendix B: 'Global characterisation of the colonic structure'. The mucosal layer was then unrolled with $I = 200$ point-long contours, an electric field search window of 31 points ($w = 15$), electric field line segment size $\delta = 2$, maximum number of steps $K_{max} = 50$, minimum electric field line variance $V_{min} = 50$, rigidity equal to 1, interpolation interval $m = 20$, $C = 50$ depth levels and rotation axis origin $O = (1080, 780)$. The resulting 3D cycloramas had a length of about (differs per sample) $L \approx 2500$ pixels, height $H = 301$ pixels, and depth $C = 50$ slices.

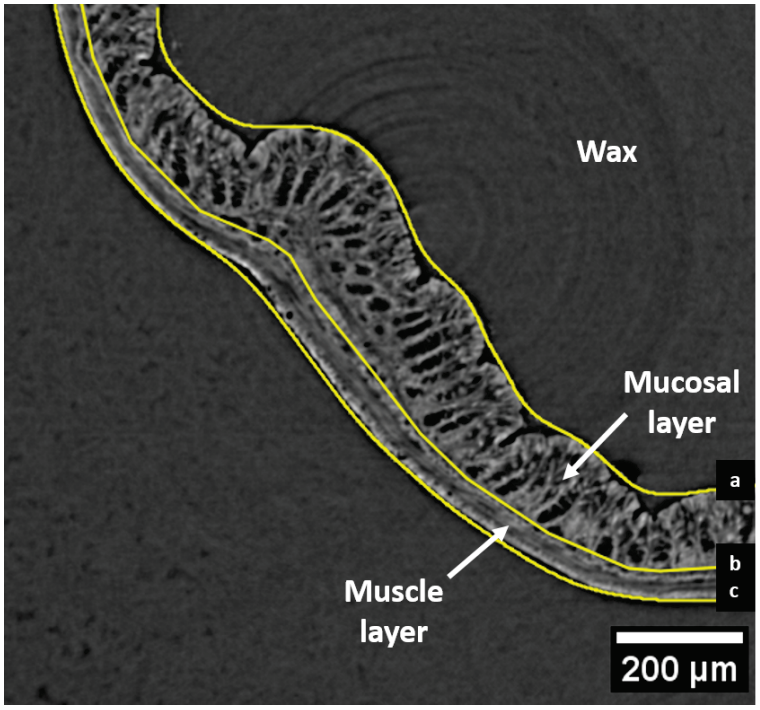


Figure 5.2: Mucosal and colonic wall definition

The three yellow lines (a, b & c) represent contours that were drawn by hand to define the mucosal layer (area between a and b) and the colonic wall (area between a and c). These served as boundary contours in the 3D cyclorama method used for digital unrolling of the colons.

5.1.2 Visual identification of budding crypts and inverse mapping

After unrolling, sequential slices of the 3D cyclorama depict cross-sections of the crypts at different depths of the colonic tissue (Figure 5.3 and Chapter 4), and are suitable to identify

budding crypts as follows[3]: Individual crypts were traced through the colonic tissue layers at different depths, i.e. traced through the slices of the 3D cyclorama (unrolled colon). Non-budding crypts were identified as crypts with a single cross-section that did not divide on any of the cyclorama slices. Budding crypts were identified by pinpointing those locations where a single cross-section of their lumen divided from one to several cross-sections. A 3D point ('Point tool' in Fiji) was manually created inside the lumen of individual crypts on each cyclorama slice. Hence, each crypt was represented by a set of 3D points defined on the 3D cyclorama (see Figure 5.1c and Figure 5.3). All crypts were manually counted on each 3D cyclorama by visual inspection.

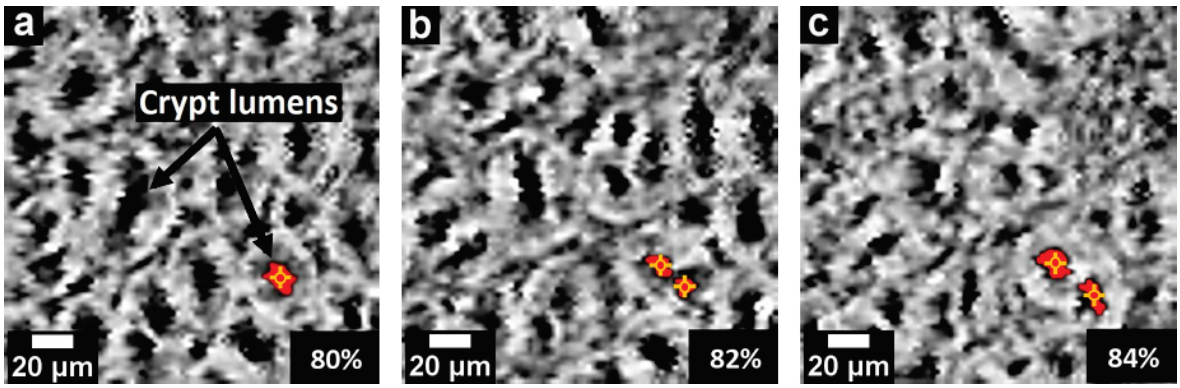


Figure 5.3: Crypt budding revealed in 3D cycloramas

From left to right (a – c) sequential slices of a 3D cyclorama with a total of $C = 50$ slices or depth levels. Cross-sections of colonic crypts are revealed at relative depth levels of 80%, 82% and 84%, corresponding to slices 39/49, 40/49 and 41/49, respectively (numbering of slices begins at 0 for the first slice). Crypt budding is revealed when a single cross-section, shown here in red, divides from one to several cross-sections when moving from one slice to the next at a different colonic tissue depth, i.e. moving from (a) to (b). Yellow crosses show points created inside the lumen of individual budding crypts on each cyclorama slice, used to represent each budding crypt on the 3D cyclorama. The jittering on the boundaries of the crypt lumens is a consequence of the fact that contours are mapped individually (see section '4.2.2 Mapping: Generation of 3D cycloramas'), which allows for inconsistencies on the vertical dimension of the 2D cyclorama.

[3] Thanks goes to Eileen Y. Li, 2nd year undergraduate student from the Chinese University of Hong Kong (CHUK), who meticulously counted the budding and non-budding crypts on the 3D cycloramas.

5.1.3 3D segmentation and manual refinement

ITK-SNAP [256] (www.itksnap.org), an open-source image processing application to segment 3D structures in biomedical image stacks, provides semi-automatic segmentation using active contour methods. ITK-SNAP, (v2.8.0-beta) was used here for 3D isolation of the crypt lumens. It requires a (set of) seed points that are expanded over the 3D image by applying a region growing algorithm. The iterative region growing algorithm maintains a curve or active contour that evolves under the influence of a region completion force, which tends to expand the curve and a smoothing factor, which tends to make the curve less rugged. On each iteration, ITK-SNAP reads the image grey values around the curve and decides how the curve needs to evolve in order to segment the structures of interest, i.e. the crypt lumens in this case.

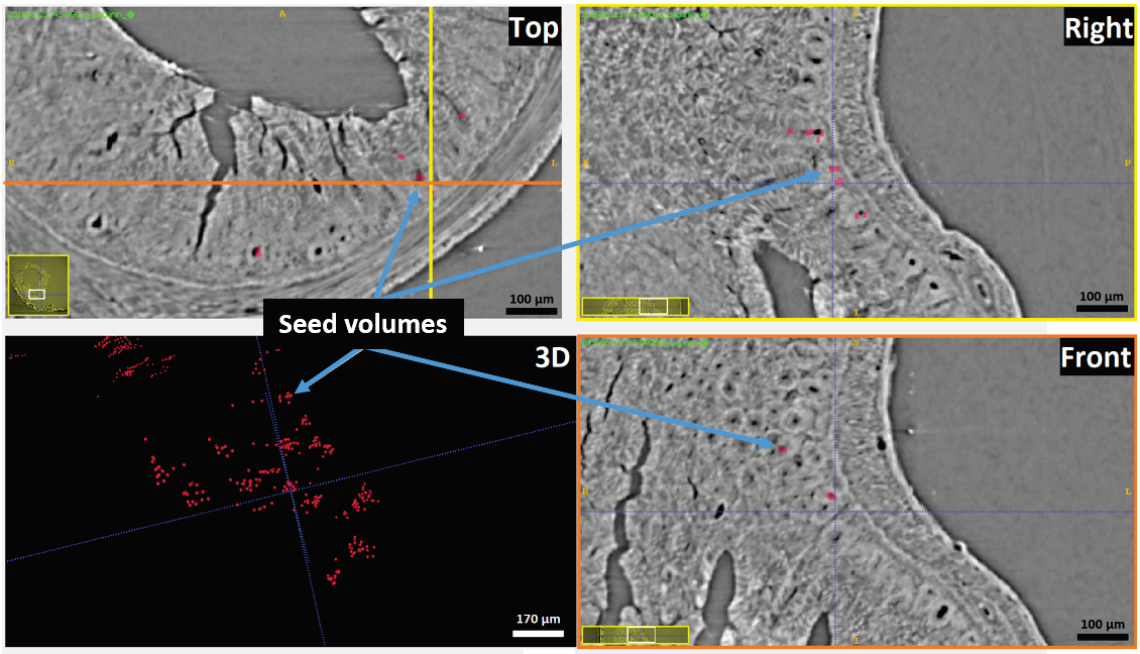


Figure 5.4: Budding crypt segmentation procedure

The viewport of ITK-SNAP is composed of four views (Top, Right, Front and 3D). The panels Top, Right and Front show orthogonal cross-sections of the µCT stack. The panels Right and Front show cross-sections across the yellow and orange lines on the panel named Top, respectively. (3D) The panel named 3D shows the µCT mask with the volumes of the back-mapped seed points (resulting from the inverse mapping between the 3D cyclorama and the µCT data). (Top, Right, 3D, Front) The 3D seed volumes shown in the panel 3D, which are overlaid on the µCT stack in the rest of the panels, are used to initialise the segmentation and expand it throughout the crypt lumens using region growing. (Top, Right, Front) The panels named Top, Right and Front only show a zoomed-in view of the respective orthogonal cross-sections. The

white rectangles within the yellow panes on the bottom-left corners of the Top, Right and Front panels show which part of the cross-sections is displayed in the panel.

As discussed in section '4.3.1 Limitations and outlook' of the 3D cyclorama, estimating the dimensions of features on cycloramas must be avoided due to non-linear deformations induced by the involved mapping. Thus, to perform image segmentation and quantification on the original (3D) µCT stack, and not on the reformed 3D cyclorama, the seed volumes for ITK-SNAP were first created as a binary mask with a zero value for the background, and cubes of 5 × 5 x 5 pixels^3 (or 11 × 11 × 11 µm^3) with a value equal to one, centred on each of the points on the cyclorama. This cyclorama mask was subsequently mapped back onto the µCT stack using the inverse of the computed mapping between the µCT stack and the 3D cyclorama, resulting in a (binary) µCT mask (see Figure 5.1d and the 3D view of Figure 5.4). Due to computer memory limitations (laptop computer with 16 GB of RAM), the bit-depth of the µCT stacks was reduced from 16 bit to 8 bit with grey-value clipping thresholds at 21,000 a.u. and 37,000 a.u. These values were derived using an automated threshold ('Auto-threshold tool' in Fiji, with the default algorithm: 'isodata' [257]) on a test µCT stack, and the rest were clipped with the same thresholds for consistency. Since all µCT stacks were produced using the same experimental settings to image similar samples, the grey value distributions throughout the stacks were comparable. Hence, the test µCT stack was randomly selected as a representative of the entire dataset. The back-mapped seed volumes resulting from the inverse mapping appear as red cubes in the 3D view of Figure 5.4. This 3D, binary, µCT mask had the same dimensions as the µCT stack of the colonic segment (1844 × 1844 x 301 pixels^3 or 4.1 × 4.1 x 0.7 mm^3), and was provided to ITK-SNAP as an initial image segmentation. The region growing algorithm was thus initialised with these seed volumes and expanded the 3D segmentation within the crypt lumens.

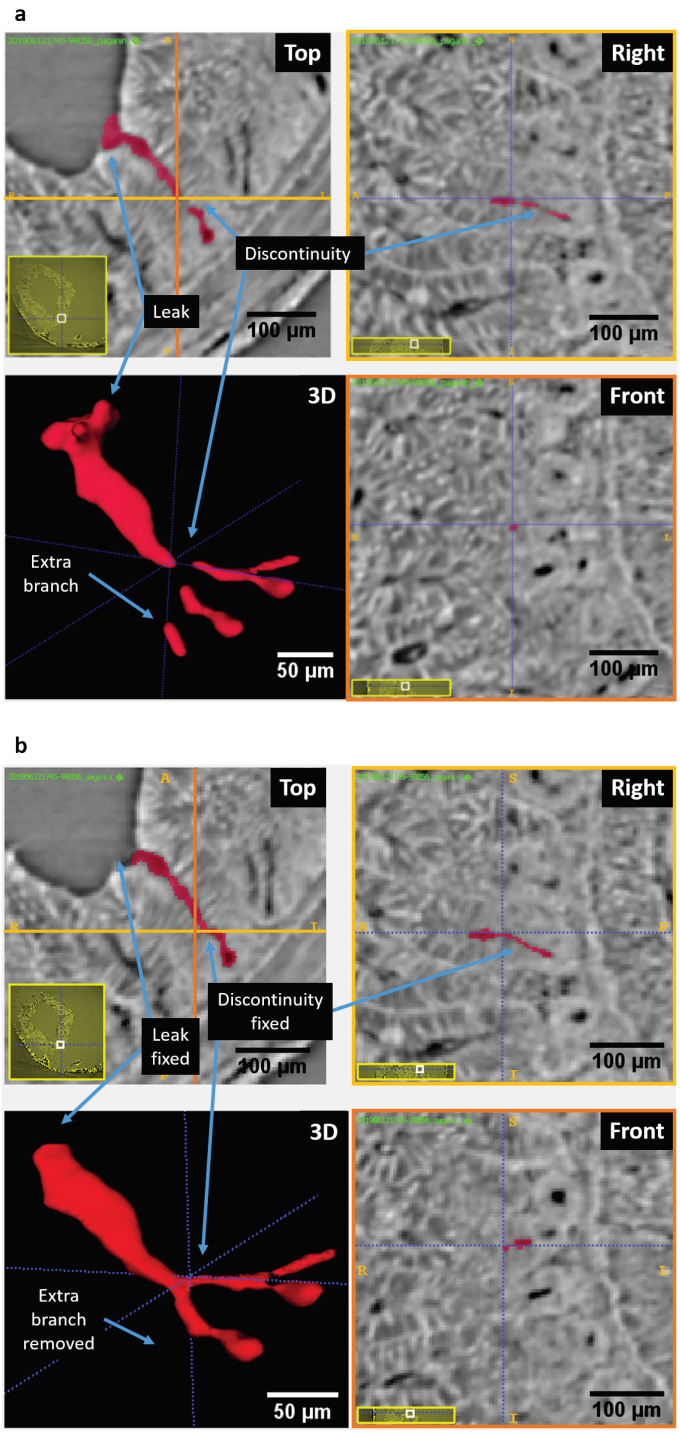


Figure 5.5: Problems with automated region growing algorithm

The viewport of ITK-SNAP is composed of the Top, Right and Front panels, which show zoomed-in views of orthogonal cross-sections of the µCT stack. The panels named Right and Front show cross-sections across the yellow and orange lines on the panel named Top, respectively. The white rectangles within the yellow panes on the bottom-left corners of the panels named Top, Right and Front show which part of the cross-sections is displayed in the panel. The red overlay represents the segmented crypt resulting after 200 iterations of the region growing algorithm. The 3D view

panel shows a 3D rendering of the segmentation. (a) Two problems arise with the region growing algorithm: First, the segmented crypts are often discontinuous where the crypt lumen becomes too narrow, due to limited spatial resolution. Second, the segmentation might 'leak out' to neighbouring structures, either creating extra branches that do not correspond to the crypt or extending through the colonic lumen. (b) These issues were manually fixed by completing missing parts or deleting areas where the segmentation exceeded the crypt lumen's boundary.

The ITK-SNAP region growing algorithm was run for 200 iterations with a low-pass filter threshold of 117 a.u., a region completion force equal to 1 and a smoothing factor of 0.2. When finished, the lumens of the budding crypts were roughly segmented. Two problems arise with the automated region growing algorithm, shown in Figure 5.5: First, the segmented crypts are often discontinuous where the crypt lumen becomes too narrow, due to limited spatial resolution. Second, the segmentation might 'leak out' to neighbouring structures, either creating extra branches that do not correspond to the crypt, or extending through the colonic lumen. These issues were visually identified and manually fixed using the segmentation tools provided by ITK-SNAP. Missing parts were completed and areas where the segmentation exceeded the crypt lumen's boundary were deleted by concurrently working on all views shown in Figure 5.5 (issues were identified on the 3D panel and completing/deleting was carried out on the Top, Right and Front panels). Although ITK-SNAP quickly provides a rough initial segmentation, the manual refinement process is particularly laborious and slow. Hence, since time constraints did not allow for complete segmentation of all the crypts, a random subset of 15 to 20 budding crypts in each of the colonic segments was considered for manual refinement and inclusion in the dataset for analysis discussed later on. For the purposes of a comparison between the untreated control group and the AOM&DSS-treated cancer model (see Appendix B.2: 'Colonic structure morphometry in health and disease') a random sub-set of 15 to 20 crypts were also segmented in each of the colonic segments of the untreated control group. As crypt budding was almost absent in this group (one or two budding crypts were found in each sample), the subset of segmented crypts in the untreated control group contained mainly non-budding (normal-looking/healthy) and the few identified budding (possibly fissioning/healthy) crypts.

5.1.4 Crypt isolation and alignment

The output of the segmentation process outlined above (see Figure 5.1f) is a binary image for each sample with the voxels corresponding to the background having the value of zero (black). The pixels corresponding to the crypt lumens have the value of one, which is shown as red in Figure 5.5 or white in Figure 5.6. Segmentation only portrays the shape of the crypt lumens

(empty space), discarding the excess information related to the surrounding tissue. The next step was to isolate individual crypts and align them vertically (longitudinal axis of the crypt being perpendicular to the image plane) in separate image stacks, so that the morphological quantification described in section '5.2 Definition of budding and non-budding crypt characteristics' can be performed automatically. This alignment was implemented in Fiji using functionality provided by the MorphoLibJ (v1.4.1) collection of mathematical morphology methods [258] and the BoneJ plugin (v6.1.1) for bone image analysis [259]. The connected components labelling tool of MorphoLibJ was used to identify individual 3D entities (segmented crypt lumens) in the binary image. This resulted in an 8-bit grayscale image stack with each voxel value being equal to a unique label number for each crypt lumen (zero for the background and one to twenty for the voxels belonging to crypt lumens). Next, the 3D bounding box tool of MorphoLibJ was used to compute the 3D boundaries of each crypt lumen, which was then used to isolate them ('Duplicate tool' in Fiji) in individual image stacks with an added 5 pixels or 11 µm clearance on each side. These were subsequently aligned vertically across their long axis using the 'Moments of inertia tool' in BoneJ (see Figure 5.1g). The latter calculates the three orthogonal principal moment of inertia axes, and creates a new image stack (segmented stack) with the crypt lumen centred and rotated so that the plane described by the two minor moment of inertia axes coincides with the image plane of the segmented stack (Figure 5.6c). Employing the moments of inertia provides an automated method to consistently define the orientation of the segmented crypt lumens so that cross-sections appear as uniform as possible after alignment.

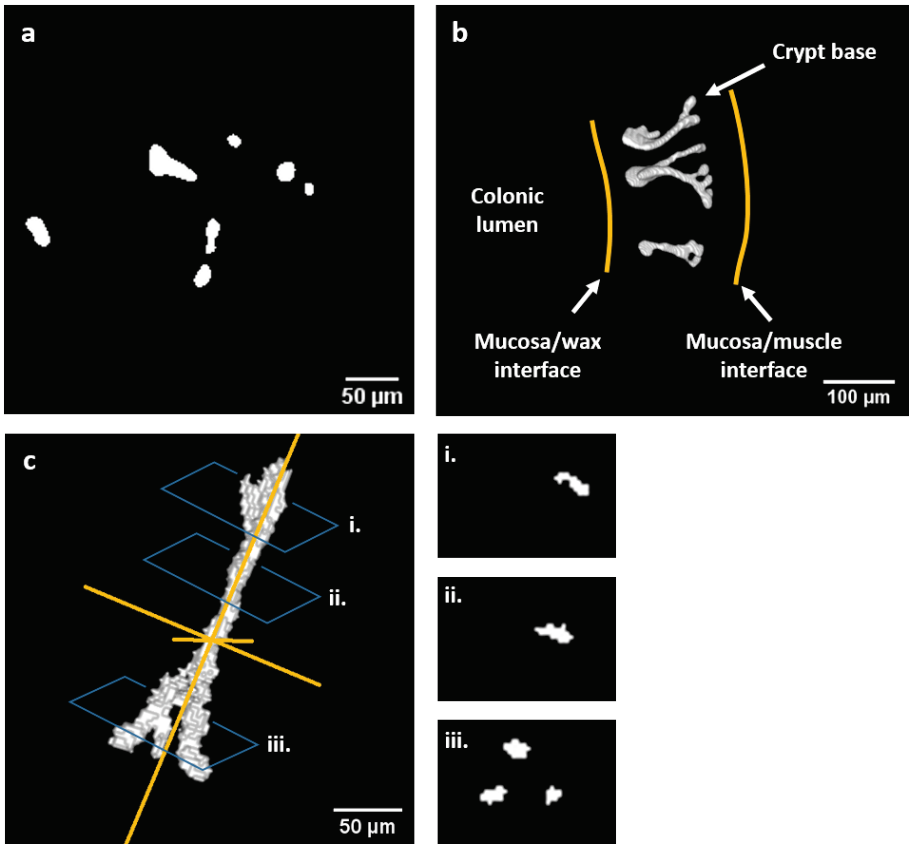


Figure 5.6: Isolation and alignment of segmented crypt lumens

(a) Each slice of the image stack produced by the segmentation procedure is a binary image, with pixels corresponding to the crypt lumens having the value of one (white) and zero (black) elsewhere. (b) A 3D rendering of the binary image stack reveals the 3D structures (budding crypts) and their position in space within the tissue. (c) A 3D rendering of a budding crypt. The yellow lines represent the computed principal axes of inertia. The crypt is then aligned so that the plane described by the two minor moment of inertia axes coincides with the image plane of the segmented stack. The slices of the segmented stack after alignment (e.g. slices i, ii and iii corresponding to the respective planes in (c)) depict the lumen's cross-sections of a single, isolated crypt, which are used to identify and quantify its branches. Note that the crypt lumen is perpendicular to planes i. and ii., whereas the three branches at its lower part are not perpendicular to plane iii.

In practice, this approach resembles the case where serial planar slices, (locally) perpendicular to the colonic mucosa, were taken across a colonic crypt. Note however that for the principal moment of inertia axis to run through the centre of the crypt's lumen, and the alignment to produce perpendicular cross-sections, the crypts would need to be perfectly straight. This is never the case because non-budding crypts are generally curved, and budding ones have more than one

branches, as shown in Figure 5.6c and discussed in detail in section '6.2.1 Global phenotyping of budding crypts'. Regardless, this approach was employed here because it is not prone to human error, and automated, since no manual input is required. The fact that not all cross-sections in the aligned image are locally perpendicular to the crypt lumen is carefully considered in the analysis of section 6.2.1. Eventually, each aligned crypt lumen was saved in an individual image stack, called segmented stack here. Since individual, aligned stacks only contain cross-sections of a single crypt's lumen, they were binarised again to remove the grey value from their pixels, which previously corresponded to the crypt label number. After this latter binarisation, a zero grey value is reassigned for the background and the value of one is assigned to the crypt lumen.

5.2 Definition of budding and non-budding crypt characteristics

The designed and implemented methods presented here aim to fully characterise the 3D shape of budding crypt lumens, with key motivation being the support of the development of 3D computational models (see section '1.3.4 Computational models of the intestine') for the study of the early-stages of colorectal cancer. To avoid introducing physical deformations of the colons, resulting in inaccurate morphological quantification, the murine colons have been fixed in their native shape following a minimally invasive approach (see section '2.1.2 Tissue sample acquisition and preparation'), and imaged in 3D using a non-destructive imaging modality (see section '3.2.1 Materials and methods'). They were subsequently digitally unrolled using the 3D cyclorama method developed for this purpose, and the colonic cycloramas were used to identify crypt budding and locate the corresponding crypts in the µCT stack. The crypt lumens of individual crypts have been subsequently segmented from the µCT image stack and isolated into individual image stacks containing one crypt each (see section '5.1 Workflow for identification and 3D segmentation of budding').

The description of the morphological quantification that follows uses these image stacks of isolated crypts as input. As the shape of budding crypt lumens is fundamentally different (they consist of different number of branches) comparison of their structure is not trivial. To enable comparison between the budding crypt lumens, a graph was developed to represent their budding structure, called 'topology graph' here. A graph is a mathematical tool widely used in computer science to represent relationships between entities. To construct the topology graph, concepts from the graph theory [260], have been employed to represent individual crypt branches (entities) and how they are interconnected (relationships). This enabled the separation of the morphological characteristics of the crypt's lumen into two groups (see Figure 5.7): The (i) branch characteristics, which refer to individual branches of a budding crypt lumen. The (ii) global characteristics, which refer to the entire crypt lumen that includes all branches. Further sub-

divisions and detailed definition of the branch and global characteristics follow in sections 5.2.2 and 5.2.3, respectively. Following this approach, a non-budding crypt lumen is structurally equivalent to a budding crypt lumen with a single branch and no subdivision, as will become evident in section '5.2.1 Topology graph'. Therefore, the morphological quantification method described here was designed for the more general case of the budding crypt lumens, while non-budding lumens can be considered as a special case.

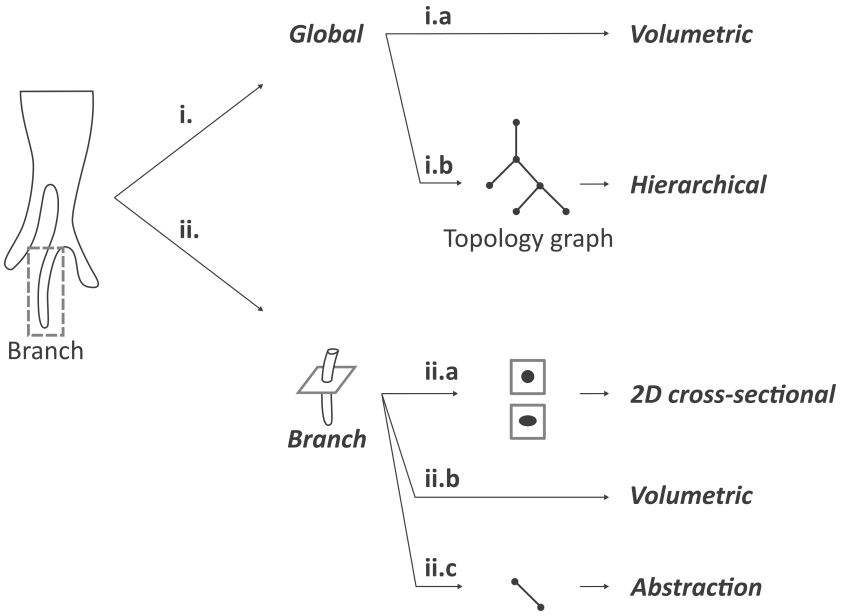


Figure 5.7: Groups of morphological characteristics

Morphological characteristics are divided in two groups: (i) global characteristics are derived from the entire crypt, and (ii) branch characteristics are derived from individual branches. The global characteristics are further divided to the (i.a) volumetric, which characterise the crypt as a whole, and the (i.b) hierarchical, which represent the topology of the crypt through a topology graph. Branch characteristics are measured on each of the branches (identified using the topology graph). Those characteristics are either (ii.a) 2D cross-sectional, which are derived from the cross-sections that make up each branch, (ii.b) volumetric, which are derived from the branch as a whole or (ii.c) abstraction, which reduce each branch to a line segment.

5.2.1 Topology graph

The hierarchical structure of the crypt lumen is initially captured by a topology graph, which depicts the individual branches and how they are related to each other. The topology graph then enables the attribution of morphological characteristics to individual branches that may be localised within the crypt structure.

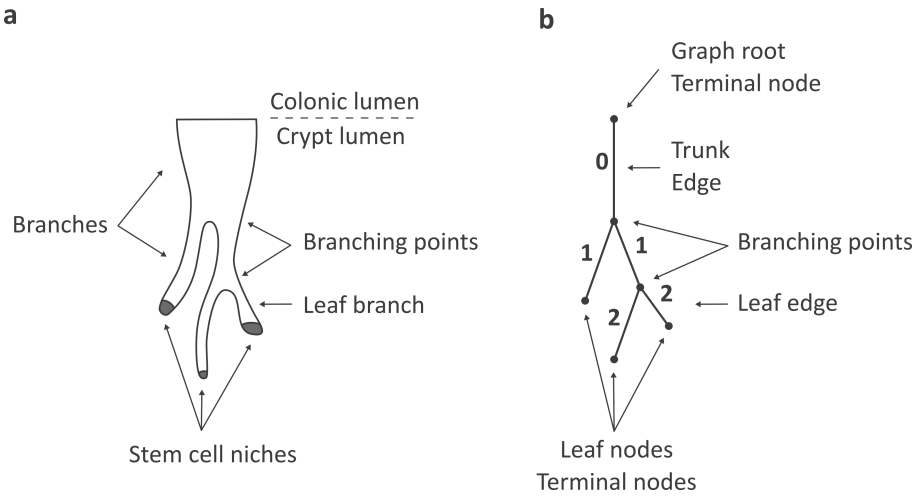


Figure 5.8: Terminology of the topology graph of a budding crypt

(a) The top boundary of the vertically aligned crypt lumen represents the opening of the crypt into the colonic lumen. Segments of the crypt that are called branches here, divide at branching points. Branches towards the bottom of the crypt that do not divide further are assumed in the literature to have a stem cell niche at their bases (see sections 1.2 & 1.3), and are called leaf branches here. (b) The corresponding topology graph is composed of nodes shown as full disks (•) and edges shown as lines (|). Each branch of the crypt lumen is represented by an edge with a parent node at its top and a child node at its bottom. The edge attached to the root node is called trunk, and the edges attached to the rest of the terminal nodes are called leaf edges. The crypt extremities and branching points are represented by nodes. Branch generations are indicated on the side of each edge.

As discussed in section '1.2 Colorectal cancer and crypt budding', the current understanding of the budding process in the literature deems a crypt that bifurcates (divides into two new crypts). This bifurcation process begins at a stem cell niche located at the crypt's base and results in two new crypts that might not undergo full division, resulting in two branches of the budding crypt. This process might be repeated several times, resulting in a budding crypt with several branches that have a stem cell niche at their base (Figure 5.8b). The terminology introduced here to describe the topology of a budding or non-budding crypt, however, does not assume that there is a stem cell niche at the bottom of any of the branches, nor that the branches are parts of a crypt. Instead, the characterisation is merely topological, hence the terms refer to the morphology of the segmented structure. Figure 5.8b shows the terminology introduced here to refer to several entities of the topology graph. This terminology is introduced to describe the shape of a budding or non-budding crypt at a fixed point in time. The wording here does not imply growth over time, nor it tries to explain the sequence of biological events that drive such a growth.

Figure 5.8a shows a schematic representation of a typical budding crypt. The top boundary of the vertically aligned crypt lumen represents the interface or opening of the crypt into the colonic lumen. Segments of the crypt that are called branches here, divide at branching points. Branches towards the bottom of the crypt that do not divide further are assumed in the literature to have a stem cell niche at their bases and are called leaf branches here. Figure 5.8b shows the corresponding topology graph. Each branch of the crypt lumen is represented by an edge shown as a line (|), and each branching point (location where segments are connected together) is represented by a node shown as full a disk (•). Nodes that have a single edge attached to them are called terminal nodes, and they are located at the extremities of the topology graph. The root of the graph is represented by a terminal node located at the top of the budding crypt lumen, i.e. at the interface between the crypt and the colonic lumen. The edge attached to the root node is called the trunk, and the edges attached to the rest of the terminal nodes are called leaf edges. Beginning from the root, a newly attached node is called a child, i.e. a node directly connected (through an edge) to another node when moving away from the root. Similarly, a parent node is the node to which a given node was attached through an edge upon creation. Finally, generation (or hierarchical level) is the number of nodes between a given edge and the root node (excluding the root node), i.e. the trunk has generation equal to zero, and the generation increases for the children of each subsequent non-terminal node (branching point).

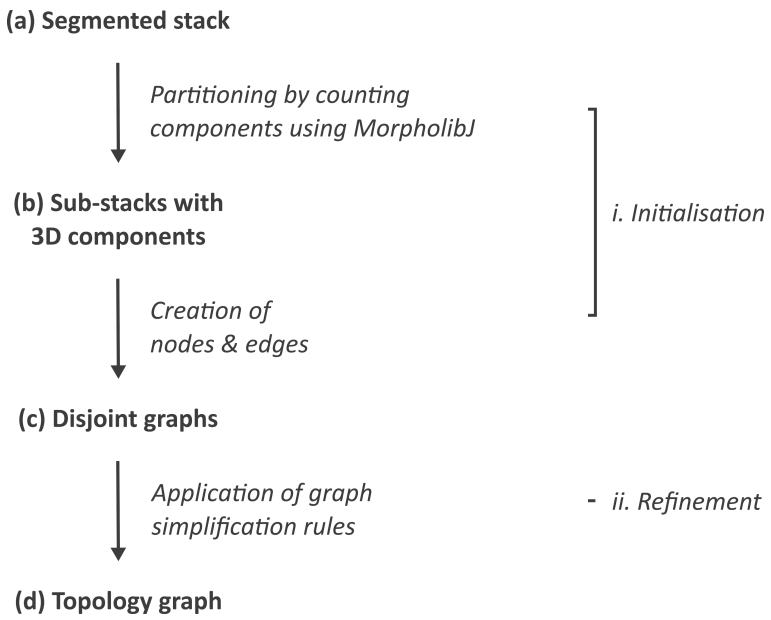


Figure 5.9: Workflow for the derivation of the topology graph

The topology graph is created in two steps, through (i) initialisation and (ii) refinement. Initially, the segmented stacks with isolated and aligned crypt lumens (a) are partitioned between the locations where the number of components changes. This results in sub-stacks containing 3D components (b) that are parts of the crypt

lumen branches (see Figure 5.10a&b). Next, a parent and child node is created at the top and bottom slices of each 3D component (see Figure 5.10b) and connected via an edge, resulting in several disjoint graphs (c). The disjoint graphs are subsequently connected together by applying a set of simplification rules (see Figure 5.11), resulting in the final topology graph (d).

The topology graph is created in two steps, through (i) initialisation and (ii) refinement, following the workflow shown in Figure 5.9. Slices of the image stacks prepared as described in section '5.1 Workflow for identification and 3D segmentation of budding' portray cross-sections of the branches of the crypt lumens. Each slice might contain one or more cross-sections, depending on the number of branches it slices through (Figure 5.10a). During the (i) initialisation step, these are counted on each slice using the 'connected components labelling tool' of MorphoLibJ, which yields the number of non-connected components in each slice of the segmented image stack (image stacks of isolated crypts discussed earlier). The number of components reduces between consecutive slices (starting from the top of the crypt at the colonic lumen and proceeding towards the crypt base) at the end of a branch and increases where a new branch is spawned. Thus, the image stack is partitioned in several sub-stacks between the locations where the number of components changes (Figure 5.10b). In each of the sub-stacks, 3D entities (parts of the crypt lumen branches) are identified using the '3D connected components labelling tool' of MorphoLibJ that identifies 3D components extending through the image stack. Graph nodes are then created at the two ends of each component, a parent node n_p at the level of the top slice and child node n_c at the bottom, and linked through an edge $e: n_p \rightarrow n_c$ (see Figure 5.10a&b). The nodes are defined at the centroid of the components' cross-sections in the top and bottom slices of each sub-stack. These are computed using the 'Analyse regions tool' of MorphoLibJ, which computes several 2D morphological characteristics of individual components, including their centroid. This process results in disjoint graphs representing parts of the crypt lumen's branches (Figure 5.10b).

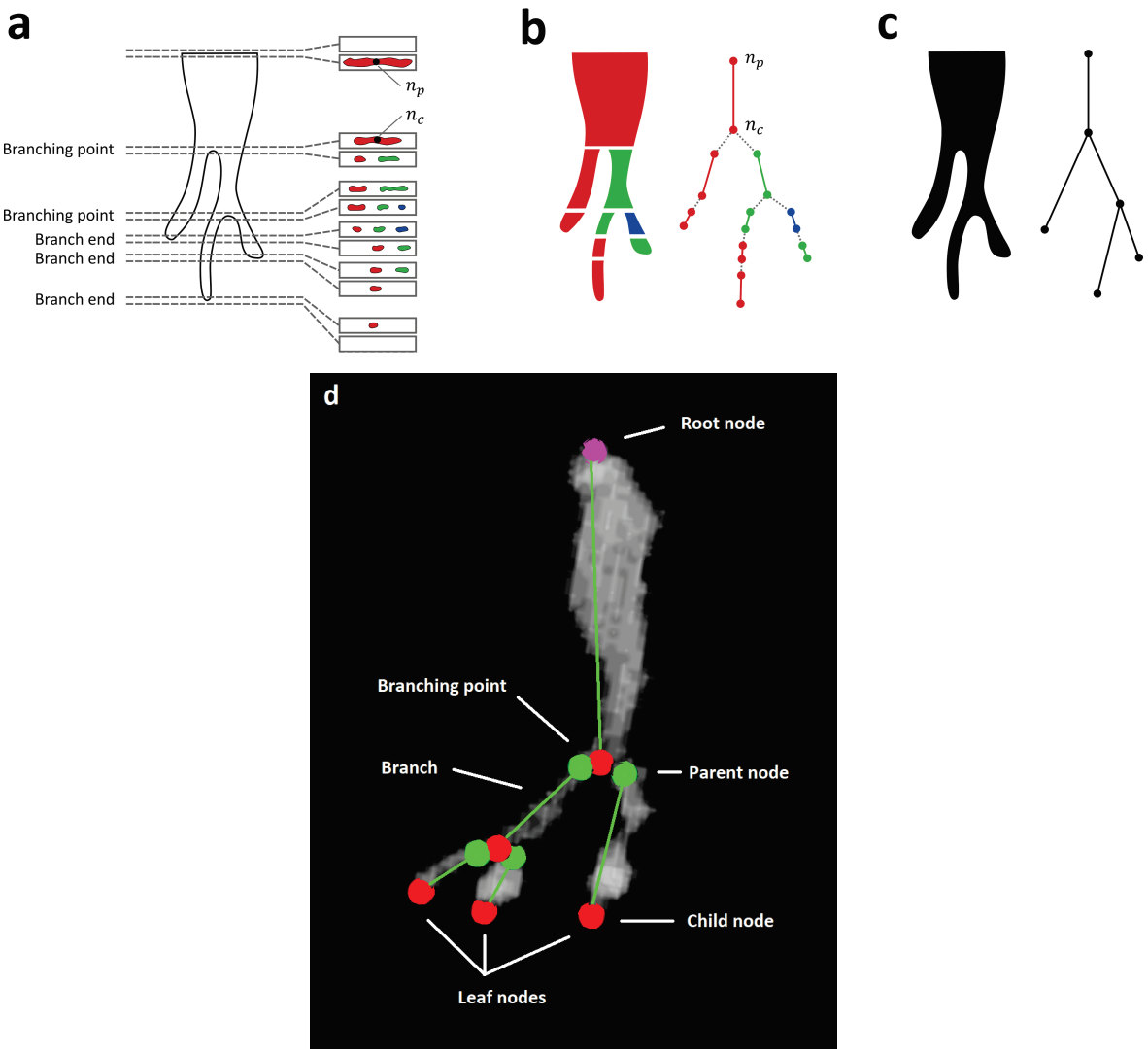


Figure 5.10: Computation of the topology graph for budding crypts

(a) Components within the cross-sections of the crypt's branches in the segmented stack are labelled with a unique grey value on each slice, here represented by different colours (1: red, 2: green, 3: blue). (b) Several sub-stacks are created between the locations where the number of components changes, and 3D components are labelled within each sub-stack. A parent node n_p and a child node n_c are created at the centroid of each component at the top and bottom slices of each sub-stack. This results in disjoint graphs, collectively representing a given branch with more than one edge. Disjoint graphs are connected together by applying the rules shown in Figure 5.11. (c) Applying these rules results in a connected topology graph that represents the entire crypt structure. (d) The topology graph overlaid on top of a 3D rendering of a segmented crypt lumen. Spheres represent topology graph nodes, with the root node in purple, parent nodes in green and child

nodes in red. Branches are shown as green lines connecting their parent to their child node. For further details on the terminology see also Figure 5.8b.

During the (ii) refinement step, a connected graph is created by applying a set of simplification rules shown in Figure 5.11. Firstly, edges of two disjoint graphs with the child of one being identical to the parent of the other (located at the same position in 3D space) are replaced with a single edge as follows: Consider a graph $g_1 = \{e_1 : n_{p,1} \rightarrow n_{c,1}\}$ close to the root, and a second graph $g_2 = \{e_2 : n_{p,2} \rightarrow n_{c,2}\}$ further way, where $n_{c,1}$ and $n_{p,2}$ are located in the same position in 3D space and $\{\}$ denotes the set operator. Then the two graphs can be merged by replacing e_1 and e_2 with a new edge $e : n_{p,1} \rightarrow n_{c,2}$, thus defining a connected graph $G = \{e : n_{p,1} \rightarrow n_{c,2}\}$ as shown in Figure 5.11a. A second simplification is applied when two graphs have the same parent, i.e. $n_p \coloneqq n_{p,1} = n_{p,2}$ are located at the same position in 3D space. Then the two graphs are merged into a new graph creating the connected graph $G = \{e_1 : n_p \rightarrow n_{c,1}, e_2 : n_p \rightarrow n_{c,2}\}$ as shown in Figure 5.11b. Finally, the two child nodes can be merged together when $n_c \coloneqq n_{c,1} = n_{c,2}$ are located at the same position in 3D space, defining the connected graph $G = \{e_1 : n_{p,1} \rightarrow n_c, e_2 : n_{p,2} \rightarrow n_c\}$ as shown in Figure 5.11c. Applying these simplifications to all disjoint graphs repeatedly, until the connected graph G cannot be simplified anymore, results in the final topology graph representing the crypt lumen's hierarchical structure (see Figure 5.10c).

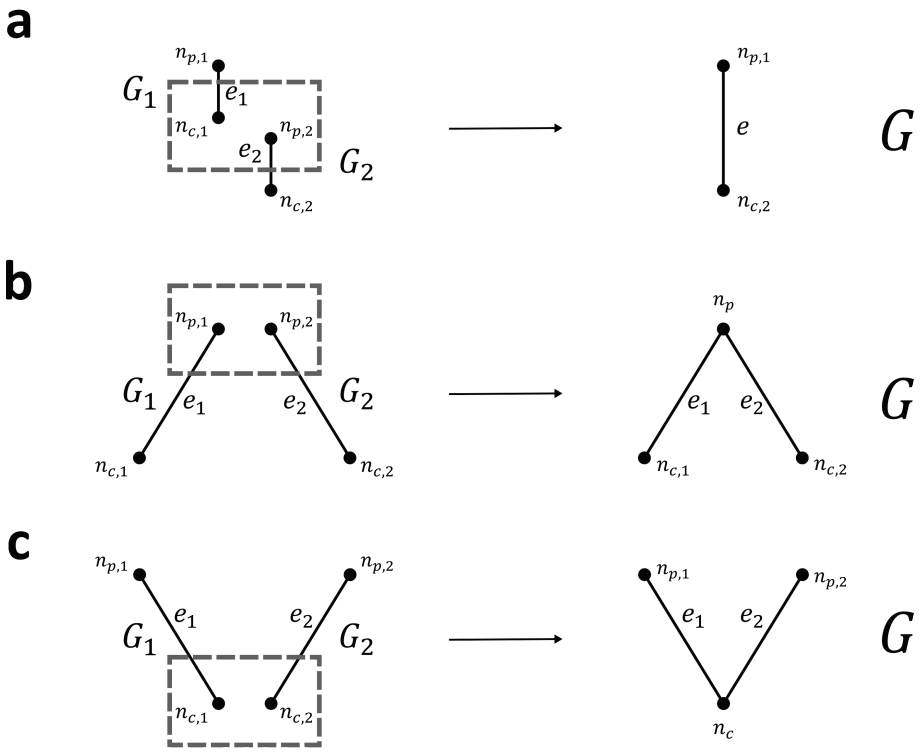


Figure 5.11: Topology graph simplification rules

(a) Two disjoint graphs $g_1 = \{e_1 : n_{p,1} \rightarrow n_{c,1}\}$ and $g_2 = \{e_2 : n_{p,2} \rightarrow n_{c,2}\}$, with the child of one being identical to the parent of the other, are merged by discarding their

common node and replacing their edges with a single edge $e: n_{p,1} \rightarrow n_{c,2}$, thus defining a connected graph $G = \{e: n_{p,1} \rightarrow n_{c,2}\}$. (b) When two disjoint graphs have the same parent, i.e. $n_p \coloneqq n_{p,1} = n_{p,2}$ are located at the same position in 3D space, they are merged into a connected graph $G = \{e_1: n_p \rightarrow n_{c,1}, e_2: n_p \rightarrow n_{c,2}\}$. Similarly, when the two child nodes are identical, i.e. $n_c \coloneqq n_{c,1} = n_{c,2}$ are located at the same position in 3D space, they are merged together into a connected graph $G = \{e_1: n_{p,1} \rightarrow n_c, e_2: n_{p,2} \rightarrow n_c\}$.

The resulting topology graph portrays the individual branches in a crypt and their interconnections in a hierarchical structure that can be used to attribute measures to their corresponding locations in the crypt. In the case of the non-budding crypts, the connected graph G is composed of a single edge that is both a trunk and a leaf. This edge connects the only two terminal nodes in the graph, one on the top of the crypt and the other on the base. Contrarily, budding crypts are composed of several branches that are arranged in several generations, and represented by different topology graphs. Therefore, to facilitate comparison between dissimilar budding structures, morphological characteristics were divided in two categories: Branch morphological characteristics (see section 5.2.2) that quantify the structure of individual branches of the crypt lumens, and global morphological characteristics (see section 5.2.3) that quantify the structure of the crypt's lumen as a whole.

5.2.2 Branch morphological characteristics

Each crypt is represented by a topology graph. Each branch is represented with an edge in the graph (see Figure 5.10b) and a segmented image stack. The topology graph is used to identify the branch as an interconnected part of the crypt, and most importantly to determine the branch's generation. Morphological characteristics of each branch can then be compared across several crypts by aggregating them by generation (see section 6.2.1). Figure 5.7 and Table 5.1 show how the several morphological characteristics have been defined and derived from individual branches. The branch morphological characteristics are separated into three groups (see Figure 5.7): (ii.a) The 2D cross-sectional correspond to measures computed for each slice of the image stack and averaged for each branch. (ii.b) the volumetric measures describe the branch's shape as a whole, and (ii.c) abstraction measures simplify the branch into a 3D line segment.

For the first group (ii.a), 2D quantification of cross-sections of the crypt lumens has been performed on individual slices of the segmented stack of the isolated crypt using the '2D analyse regions tool' of MorphoLibJ. It readily provides, among others (a complete list of all available measurements can be found in the MorphoLibJ user manual [270]), the perimeter, the

cross-sectional area and the circularity of the 2D cross-sections (see Figure 5.12). As the boundary of the segmented crypt lumen coincides with the crypt epithelium, the cross-sectional perimeter is a measure proportional to the number of epithelial cells in the cross-section of the crypt. The cross-sectional area, although it corresponds to the empty space inside the crypt lumen, is a measure of the crypt lumen's size at a given location of the branch. Both the cross-section perimeter and area are estimated in MorphoLibJ by employing the Crofton formula, which estimates the true (continuous) perimeter or area of a discretised shape. It considers sets of isotropic test lines and counts the number of intersections with the discrete cross-section boundaries to derive the true perimeter of the cross-section. This method provides a better estimation of the cross-section perimeter or area than simply counting the number of pixels, because it avoids systematic errors of the perimeter estimation due to the discretised nature of the actual, continuous shape of the cross-section [258, 261]. The circularity, also referred to as shape factor or isoperimetric deficit index, is the ratio of the area A over the square of the perimeter P^2, normalised such that the value for a disk equals to one ($Circularity = 4\pi\frac{A}{P^2}$). It is a measure of how close the cross-sectional shape resembles a circle. The solidity or convexity ratio is defined as the ratio of the area of the crypt lumen to the area of its convex hull [129], and it measures the extent to which a shape is convex. Convex cross-sections have larger solidity than concave ones. The area of the convex hull was computed using the 'convex hull tool' in Fiji, and the solidity was derived by dividing the crypt lumen area by the convex hull area. MorphoLibJ also fits the inertia ellipse, which is an ellipsoid centred at the cross-sectional centre of mass, with its axes being inversely proportional to the square root of the corresponding moment of inertia (see Figure 5.12). Thus, a complex cross-sectional shape is modelled with a simplified one, i.e. an ellipse. The fitted inertia ellipse is useful to compute the straightness discussed next, and the inertia ellipse elongation. The latter is defined as the ratio between the length of the principal axes of inertia and it is a measure of how non-circular the cross-section is. The maximum Feret diameter is the maximum distance computed over all the pairs of points belonging to the cross-section, and is a measure of how far it extends throughout the cross-sectional plane (see Figure 5.12). The last measure in the 2D cross-sectional group is the maximum inscribed circle radius (see Figure 5.12), i.e. the radius of the largest circle that can be inscribed within the 2D cross-section. The maximum inscribed circle radius is a measure of how large the largest part of the cross-section of the crypt lumen is. All of the above measures were derived for each of the 2D slices of each branch, and the mean value was computed per branch. Note that when the branches are not aligned with the vertical axis, the sectioning planes (images of the label image stack) are not perpendicular to the branches (see Figure 5.13a). Therefore, measures on 2D cross-sections are inherently biased, and must be interpreted with care. To this end, the leave angle (a measure of the misalignment) is defined later on.

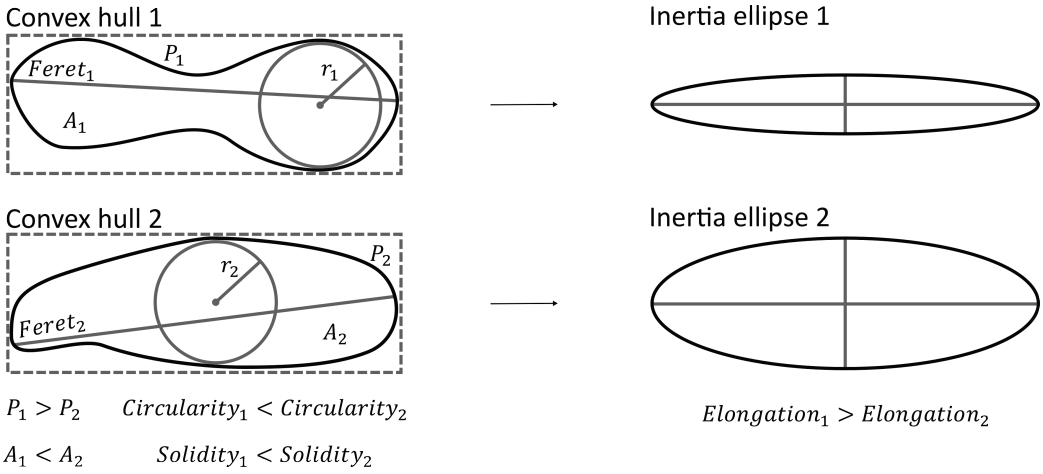


Figure 5.12: 2D cross-sectional characteristics

Two cross-sections with different perimeters P_1 and P_2 and areas A_1 and A_2, respectively, could possibly be enclosed in convex hulls of the same size. The relative values of the perimeters, cross-sectional areas and convex hull areas result in dissimilar circularities (measure of similarity to a circle) and solidities (measure of the cross-section convexity). The radii r_1 and r_2 of the largest inscribed circles are shown here, along with the maximum Feret diameters (maximum distance computed over all the pairs of points belonging to the cross-section). Each cross-section may be represented by an ellipse of inertia or inertia ellipse. The ratio of the inertia ellipse principal axes (elongation) is a measure of how non-circular the cross-section is.

On the other hand, (ii.b) volumetric measures shown in Figure 5.13a quantify the branch as a whole and are not dependent on the slicing orientation. These have been estimated using the '3D analyse regions tool' of MorphoLibJ for each branch separately. The computed volumetric measures are: the volume, the surface area, the mean breadth and the maximum inscribed ball radius. The volume refers to the empty space corresponding to the segmented crypt lumen, and it is a measure of the branch size. The surface area refers to the surface of the segmented crypt lumen and can be considered as a measure of the surface of the crypt lumen's epithelium. Note, however, that the surface area of individual branches slightly overestimates the surface of the branch lumen's epithelium. This is because the branch's surface area includes the surface of the cross-sections at the top and bottom of the branch, shown in Figure 5.13b as A_1 and A_2, respectively. These are typically adjacent to the branches above and below, thus they are not part of the branch's epithelium. Similarly to the cross-sectional perimeter and area, the volume and surface area are computed using the improved estimation implemented in MorphoLibJ that employs the Crofton formula [261]. The mean breadth is the average of the calliper diameter over all directions, and it is a measure of the branch's extent throughout the mucosa. Finally, the maximum inscribed ball radius is a measure of the size of the largest part of the branch lumen.

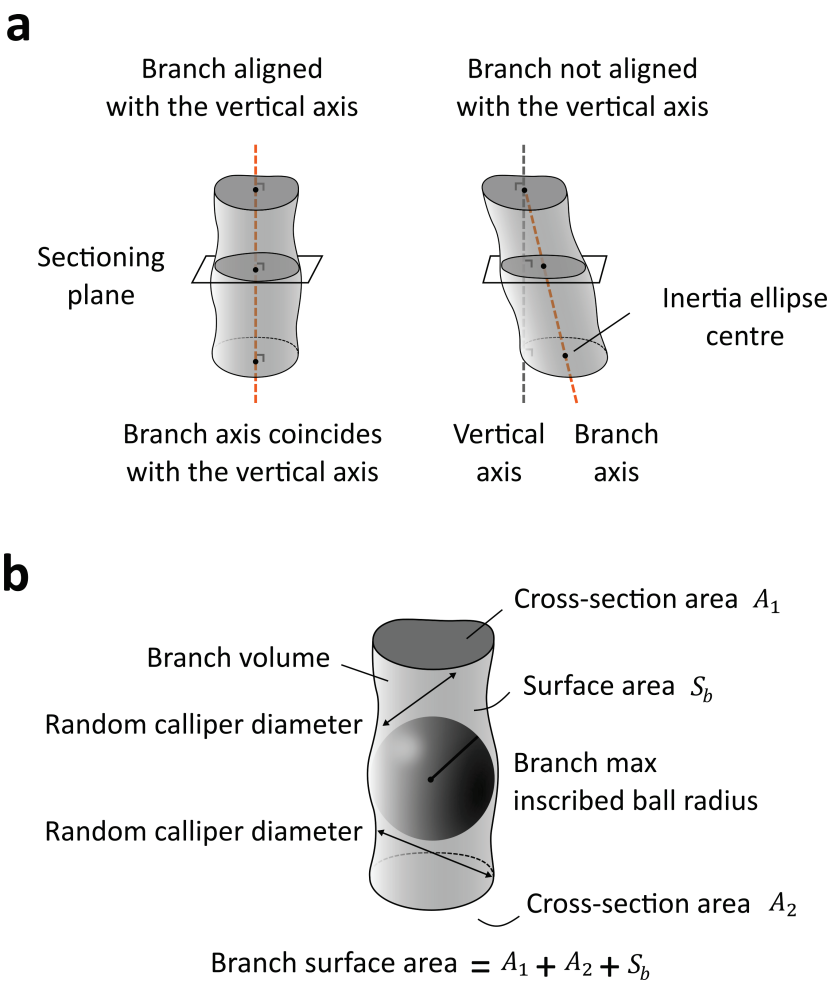


Figure 5.13: Branch volumetric characteristics

Schematics of the 3D structure of individual branches, used to represent the several volumetric characteristics. (a) Each branch is represented by a set of cross-sections in the sectioning planes or slices in the segmented stack. When the branch axis is not aligned with the vertical axis, cross-sections at similar locations of a given branch have different shape. Hence, 2D cross-sectional characteristics are inherently biased. (b) Each branch is quantified by its volume and surface area. The branch surface area includes the cross-section areas of the top and bottom slices A_1 and A_2, respectively, and surface area S_b, which corresponds to the crypt epithelium. The two-sided arrows show the calliper diameter at two random directions, used to derive the mean breadth (average of the calliper diameter over all directions). The branch maximum inscribed ball radius, is the radius of the largest ball that can be enclosed within the branch.

The group of abstraction measures (ii.c) simplifies the branch representation, allowing to reduce it to a 3D line segment. The segment's length is quantified as the Euclidean distance between the centroids of the parent and child nodes used to define it (see section 5.2.1 and Figure 5.14a). This

measure, although simple to compute, underestimates the actual (geodesic) length, which is larger due to the inherent curvature of the branch (see Figure 5.14a&b). The geodesic distance to the closest leaf (see Figure 5.14c and Figure 5.15a) is defined as the geodesic distance between the parent node of a given branch to the closest leaf child node of the crypt. Assuming that a stem cell niche is located at each leaf node, the geodesic distance to the closest leaf is the distance that a differentiated stem cell needs to travel to reach the top of a given branch. This measure facilitates the quantification of two other important measures: First, the minimum geodesic distance to root (see Figure 5.14d). That is the geodesic distance between the root and the closest leaf edge, which corresponds to the distance that an epithelial cell needs to travel to reach the colonic lumen. Second, the geodesic length of the leaf edges (see Figure 5.14b), which corresponds to the actual length of the branches that are typically assumed to have a stem cell niche at their base.

The geodesic distance to the closest leaf is computed using the '3D geodesic distance map tool' of MorphoLibJ. The algorithm is initialised with a set of seed points and propagates throughout the branch (constrained inside the segmented image stack) assigning each voxel an estimation of its distance from the closest seed point. For computational purposes incremental distance values are computed as 'chess knight' distances (an approximation of the real Euclidean distance that is simpler to compute). To compute the geodesic distance to the closest leaf, the 3D geodesic distance map tool is initialised with a seed point on each of the leaf nodes (see Figure 5.15a). The geodesic distance to the closest leaf is then computed for each branch as the mean of the pixel values in the top cross-section of the branch. The leaf geodesic length is derived as the geodesic distance to the closest leaf computed on the leaf branches. The leaf geodesic distance to root corresponds to the distance from a given leaf node to the root node. Note that there is one leaf geodesic distance to root for each of the leaf branches. In contrast, there is one geodesic distance to closest leaf for each branch, including leaf and non-leaf branches (see Figure 5.14c&d). It is computed following a similar approach, with the difference that the 3D geodesic distance map is initialised on the root node in this case (Figure 5.15b). Then, the geodesic distance from the root to a given leaf node is computed as the mean of the pixel values of the bottom cross-section of each of the leaf branches.

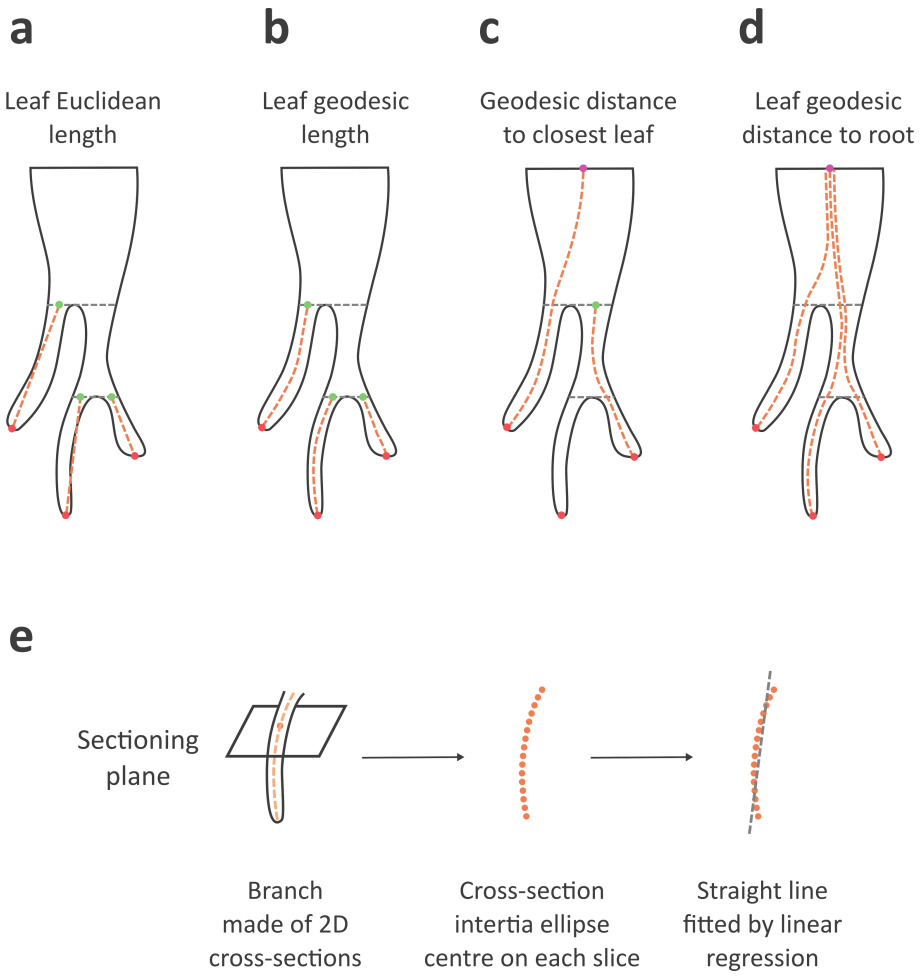


Figure 5.14: Leaf lengths, distances and straightness

(a - d) Individual branches are separated by grey dashed lines. Where relevant, parent nodes and leaf child nodes are shown as green and red full disks, respectively. The root node is shown as a purple full disk. (a) The Euclidean length is a straight line connecting the leaf branch's parent and child nodes, which underestimates the branch's actual length. (b) In contrast to the Euclidean length, the geodesic length follows the branch curvature which yields a better estimation of the branch's actual length. (c) The geodesic distance to closest leaf of a given branch connects its parent node with the closest leaf child node. (d) The leaf geodesic distance to root or leaf-to-root geodesic distance connects a given leaf's child node to the root node. (e) The inertia ellipse centres (3D points) of the branch's cross-section on each sectioning plane are fitted with a 2D straight line by linear regression. The resulting coefficient of determination R^2 is called straightness here.

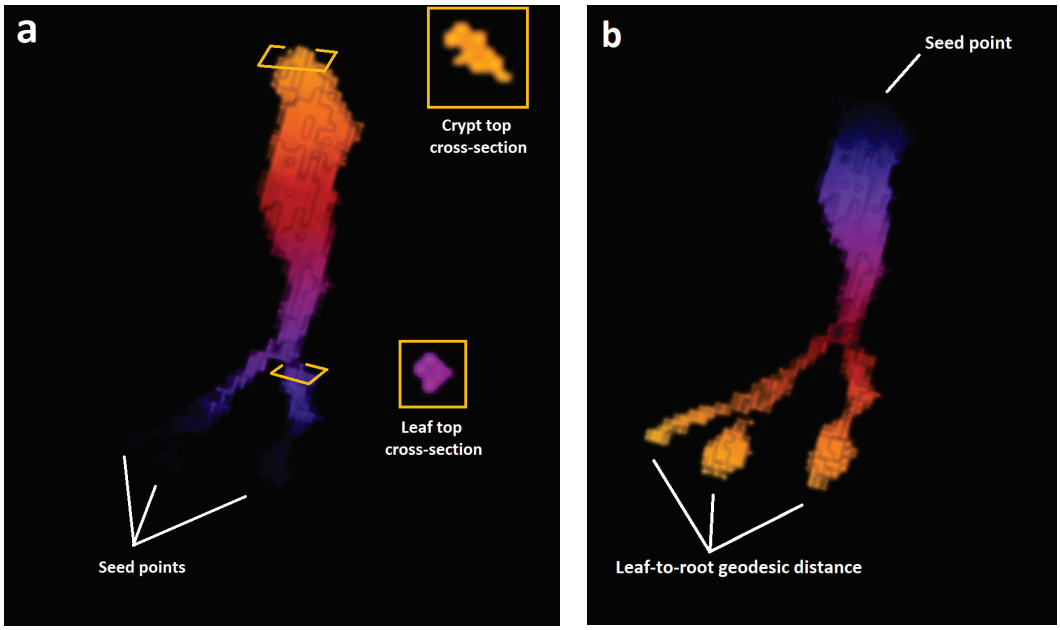


Figure 5.15: Geodesic distance to closest leaf, leaf geodesic length and leaf-to-root geodesic distance

Renderings of a budding crypt illustrating geodesic distance maps. The value of each voxel is equal to the geodesic distance between the given voxel and the closest seed point defined by the user. Brighter colours represent larger values. (a) The algorithm is initialised with seed points at the bases of the leaf branches (stem cell niches), in order to compute the geodesic distance to the closest leaf and the leaf geodesic length. The geodesic distance to the closest leaf is computed as the mean value of the voxels of the crypt top cross-section, while the leaf geodesic length is computed as the mean value of the voxels on the leaf top cross-section. (b) When the algorithm is initialised with a single seed point on the top of the crypt (root), each voxel value corresponds to its distance from the root. Thus, the leaf-to-root geodesic distance is computed for each leaf branch as the mean value of the leaf bottom cross-section.

The relative position of the branch is computed as the vertical Euclidian distance of the parent node (distance from the root node measured along the axis perpendicular to the image stack) as a fraction of the total crypt height (distance between root node and the furthest leaf node, measured along the axis perpendicular to the image stack), which is shown in Figure 5.16a and Figure 5.17. This relative quantification of the branch position allows comparison across several crypts as discussed in section '6.2.1 Global phenotyping of budding crypts', and it enables the quantification of the highest mitotic figure discussed in section '5.2.3 Global morphological characteristics'. The radial offset is defined as the distance of a given branch child node from the vertical axis (perpendicular to the image plane) passing through the parent node of a given branch (see Figure 5.16d). The leave angle is defined as the angle between the vertical axis and the

straight line segment representing the branch (see Figure 5.16d). It characterises the spatial relationship between a given branch and the branch upon which it is attached to.

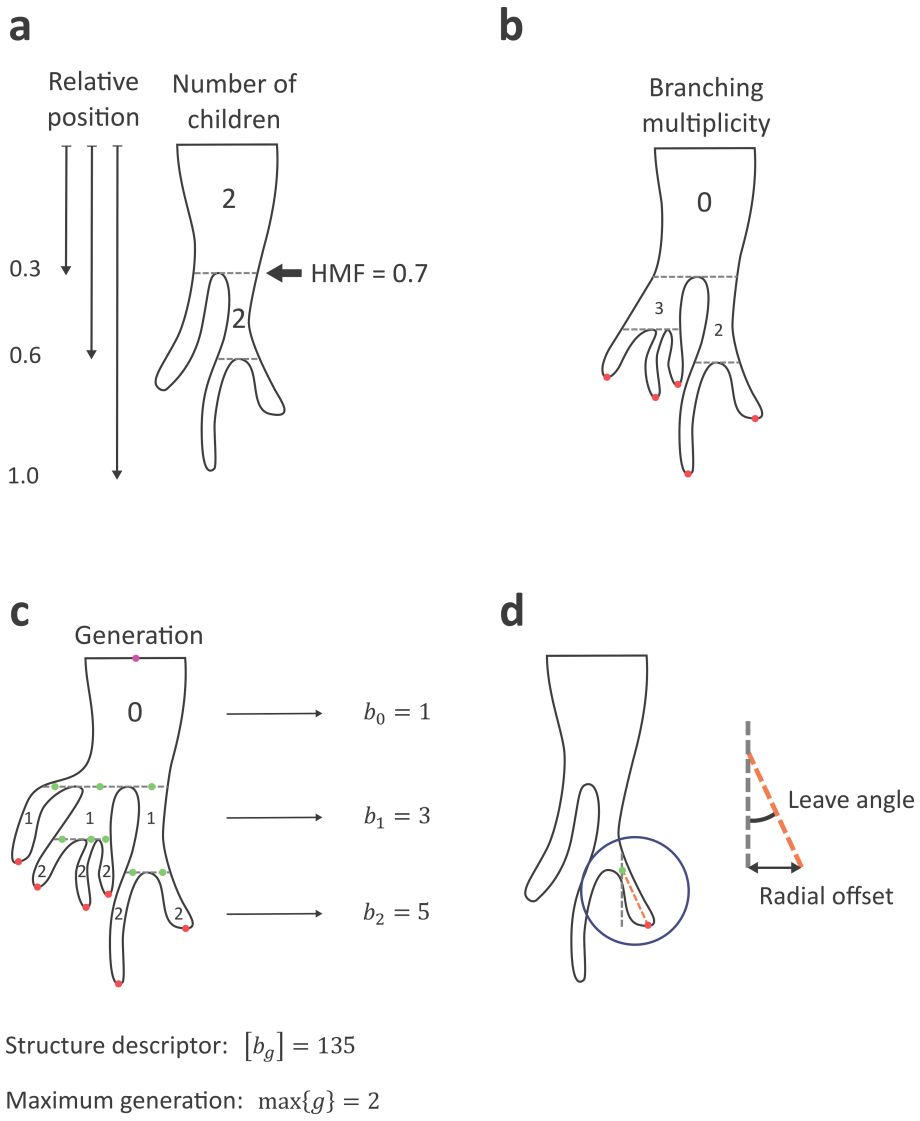


Figure 5.16: Branch number of children, multiplicity, generation and leave angle

(a - d) Where relevant, individual branches are separated by grey dashed lines, parent nodes and leaf child nodes are shown as green and red full disks, respectively, and the root node is shown as a purple full disk. (a) The number within non-leaf branches shows the number of children of the respective branch. Branching points are indicated by grey dashed lines along with their relative position (relative with respect to the crypt height). The highest mitotic figure (HMF) is the reciprocal of the relative position of the uppermost branching point HMF=1-min{relative position}. (b) The number within non-leaf branches shows the branching multiplicity of the respective branch, i.e. the number children when at least one of the children is a leaf branch. (c) The number within each branch shows its generation g, i.e. the number of budding events that occurred above it. The number of branches b_g in generation g is

used in a sequence $[b_g] \coloneqq b_0 b_1 b_2$ to define the structure descriptor of the crypt. (d) The leave angle of a branch is the angle between the vertical axis passing through its parent node and the branch axis or the line segment representing the branch. The radial offset of a branch is the distance between its child node and the vertical axis passing through its parent node.

The crypt straightness was originally introduced by Qi *et al.* [129] as the fraction of the variation in the skeleton of the crypt that is explained by a straight line. They defined and quantified straightness as the coefficient of determination R^2 resulting from fitting a straight line through the crypt skeleton. This was implemented here using the linear regression module from the scikit-learn library [262] to fit a 2D straight line through the inertia ellipse centres on all cross-sections of each branch computed beforehand (see Figure 5.14e). The module yields the coefficient of determination R^2. The branch generation corresponds to the number of budding events that occurred above it, between a given branch and the trunk (see Figure 5.16c). It is computed as the number of nodes between the edge on the topology graph that represents the given branch and the root node (see section '5.2.1 Topology graph'). The number of children of a given branch is the number of branches attached to it when moving away from the root, i.e. the number of branches attached to its child node (see Figure 5.16a) and it allows quantification of the branching multiplicity as explained next.

Table 5.1: Branch morphological characteristics

Morphological characteristics of individual branches. These are separated in three different groups: (i) The 2D cross-sectional correspond to measures computed on each slice of the image stack and averaged for each branch, (ii) the volumetric measures describe the branch's shape as a whole, and (iii) abstraction measures simplify the branch into a 3D line segment.

Measures	Description
2D cross-sectional	**Measures computed on each image slice.** * This group is biased, depending on the slicing orientation.
Perimeter (P)	Estimate of the cross-section perimeter. A measure proportional to the number of epithelial cells.
Area (A)	Estimate of the cross-sectional area. A measure of the crypt lumen's size at a given location of the branch.
Circularity ($4\pi \frac{A}{P^2}$)	'Shape factor' or 'isoperimetric deficit index'. The ratio of the area over the square of the perimeter, normalised such that the value for a disk equals to one. A measure of the crypt lumen's cross-sectional shape's similarity to a circle.

Measures	Description
Solidity	The ratio of the cross-section area over the convex hull area. A measure of the extent to which the cross-section is convex. Convex cross-sections have larger solidity than concave ones.
Inertia ellipse	'Centroid' or 'centre of gravity' around which the pixels of the cross-section are distributed, along with a fitted ellipse. A minimal representation of the cross-section, which facilitates modelling of the shape and calculation of the straightness.
Inertial ellipse elongation	The ratio of the cross-sectional inertial ellipse's principal axes. A measure of how non-circular the cross-section is.
Maximum Feret diameter	Maximum distance computed over all the pairs of points belonging to the cross-section. A measure of the cross-section extent throughout the cross-sectional plane.
Maximum inscribed circle radius	The radius of the largest circle that can be enclosed within the cross-section. A measure of how large the largest part of the cross-section of the crypt lumen is.
Volumetric	**Unbiased measures describing the entire branch.**
Volume	Estimate of the empty space occupied by the crypt lumen. A measure of the branch size.
Surface area	Estimate of the crypt lumen's surface. A measure of the surface of the crypt epithelium, which is proportional to the number of epithelial cells. Slightly overestimates the actual epithelial surface since it includes the interface between neighbouring branches (cross-sectional areas A of the top and bottom slices).
Mean breadth	The average of the calliper diameter over all directions. A measure of the branch's extent throughout the mucosa.
Max inscribed ball radius	The radius of the largest ball that can be enclosed within the branch. A measure of how large the largest part of the branch is.
Abstraction	**Simplifies the branch into a 3D line segment.**
Euclidean length	Straight-line distance between the two ends of a branch. A measure of the branch length. The actual branch length is underestimated by the Euclidean length when the branch is curved.
Geodesic distance to closest leaf	Length of the path from the top of the branch to the closest leaf child node. A measure of the distance that an epithelial cell (differentiated stem cell) needs to travel starting from the stem cell niche at the leaf child node, to reach the top of a given branch.

Measures	Description
Leaf geodesic length	The length of the curve connecting the two ends of a branch, estimated by tracing its 3D shape. Quantified as the geodesic distance to the closest leaf when the studied branch is a leaf.
Leaf geodesic distance to root	The geodesic distance from the child node of a given leaf edge to the root node. Corresponds to the distance epithelial cells (differentiated stem cells) need to travel to reach the top of the budding crypt.
Relative position	Position of the branch's parent node relative to the crypt's height. Allows comparison across branches of different crypts.
Radial offset	Distance of child node from the vertical axis (perpendicular to the image plane) passing through the parent node of the branch. Allows quantification of the leave angle.
Leaf leave angle	The angle between the vertical axis (perpendicular to the image plane) passing through the parent node and the line segment representing the branch. Quantifies the spatial relationship between a given branch and the branch upon which it is attached to.
Straightness	The coefficient of determination R^2 resulting from linear regression fitting of a straight line though the inertia ellipse centres of the branch's cross-sections. A measure of how straight the branch is. $R^2 \leq 1$ with 1 denoting a perfectly straight branch.
Generation	The branch's generation corresponds to the number of budding events that occurred above it (between a given branch and the trunk).
Number of children	Number of branches attached to the child node of a given branch.

5.2.3 Global morphological characteristics

The branch morphological characteristics defined in the previous section quantify the structure of individual branches of the crypt lumens. In conjunction with the topology graph defined in section '5.2.1 Topology graph', individual branches compose the entire crypt and the crypt's morphological characteristics are quantified in a piecewise approach. In this section, the morphological characteristics of the entire crypt's lumen as a whole are defined. The global morphological characteristics reduce the representation of the entire crypt down to a few descriptive characteristics that enable comparison between non-budding and budding crypts. These were divided in two groups (see Figure 5.7): (i.a) volumetric characteristics that portray the

dimensions of the crypt lumen and (i.b) hierarchical characteristics that describe how the crypt lumen is distributed in space throughout the mucosa.

Table 5.2 summarises the global morphological characteristics employed here. The 3D bounding box that contains the entire crypt lumen is computed using the '3D bounding box tool' of MorphoLibJ. As the crypt is already aligned within the segmented image stack (see section '5.1 Workflow for identification and 3D segmentation of budding') the height of the bounding box runs across the axis perpendicular to the image plane, and it is used to define what is hereby termed the crypt height. This does not take into account the curvature of the crypt lumen, but it represents the height of the smallest box that can contain the segmented crypt lumen. The bounding box width and length are the dimensions of the bounding box within the image plane.

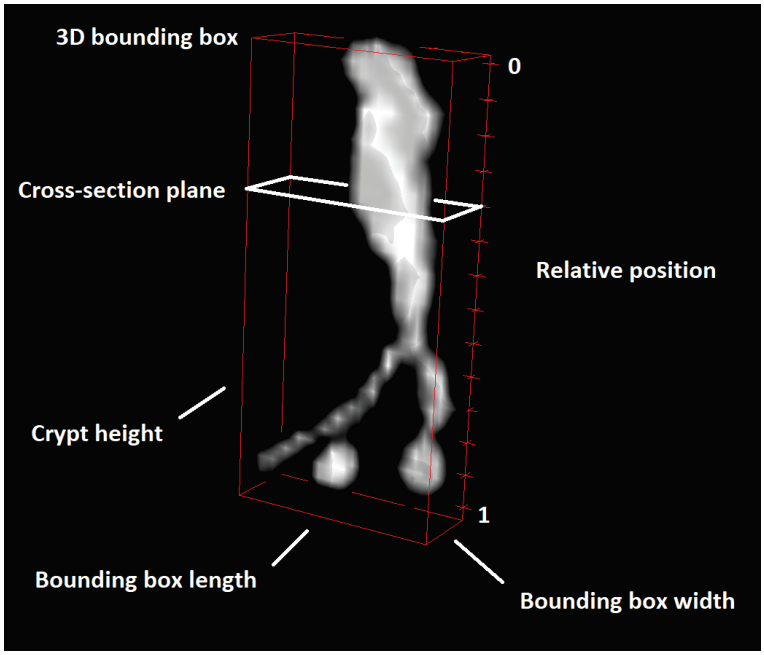


Figure 5.17: 3D bounding box, crypt height and relative position

The vertically aligned segmented crypt lumen is enclosed in a 3D bounding box shown in red. Its length and width are defined within the cross-section plane (slice of the segmented image stack). The third dimension of the 3D bounding box runs across the axis perpendicular to the cross-section plane and is used to define the crypt height. The relative position of each cross-sectional plane with respect to the crypt height is measured across the same axis, starting from 0 at the top of the crypt and ending at 1 on the bottom.

The total volume, the total surface area, the maximum inscribed ball radius and the mean breadth, shown in Figure 5.18, were measured using the '3D analyse regions tool' of MorphoLibJ for the entire segmented crypt lumen. The total volume is the total empty space occupied by the

crypt's lumen, and it is a measure of the crypt's size. The total surface area is a measure of the entire crypt epithelial area. Note that this includes the cross-sectional area of the top slice, shown as cross-section area A_1 in Figure 5.18, which is the interface with the colonic lumen and does not portray epithelial tissue. Hence, the total surface area slightly over-estimates the crypt epithelial area. The radius of the largest ball that can be inscribed within the segmented lumen (max inscribed ball radius) is a measure of how large the largest part of the crypt lumen is. The mean breadth is computed as the average of the calliper diameter of the crypt lumen over all directions, and it is a measure of the crypt's extent throughout the mucosa.

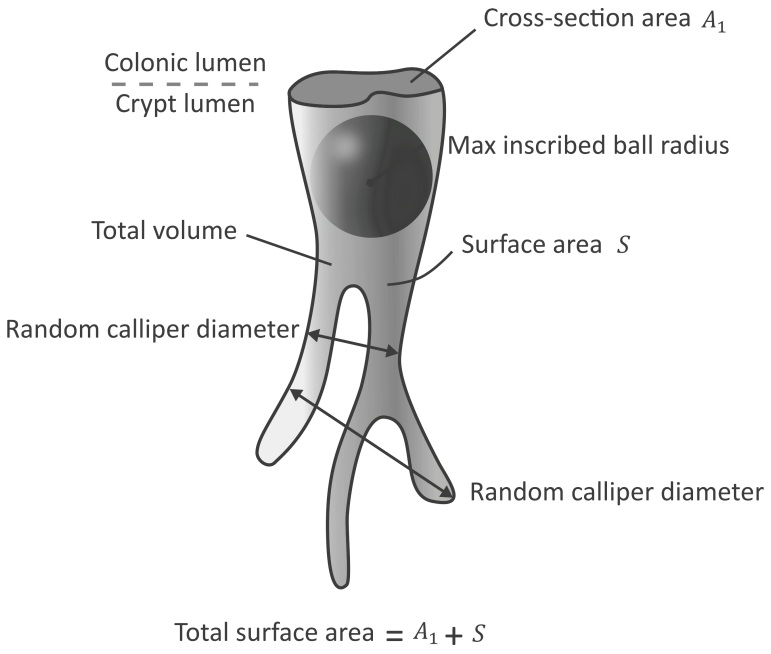


Figure 5.18: Global morphological characteristics

The crypt shape is quantified as a whole by its total volume and the total surface area, which is equal to the sum of the surface area S and the top cross-section area A_1. The maximum inscribed ball radius is the radius of the largest ball that can be enclosed within the crypt's lumen. The two-headed arrows show the caliper diameter across two random directions. These are used to derive the mean breadth, which is the average of the calliper diameter over all directions.

The leaf branches are of particular interest for the study of crypt budding because of the underlying assumption that stem cell niches exist at the base of each of these branches (see sections '1.2 Colorectal cancer and crypt budding', '1.3 Standard approaches for crypt budding research' and '6.2.1 Global phenotyping of budding crypts'). Therefore, the leaf branches were treated separately, and the following measures were evaluated for the leaf branches (see section '5.2.2 Branch morphological characteristics' and Figure 5.14): Mean leaf geodesic length is the geodesic length computed and averaged among all leaf branches of a given crypt. Similarly, the

mean leaf-to-root geodesic distance is the leaf geodesic distance to the root averaged over all leaf branches, and the mean leaf relative position is the branch relative position averaged over all leaf branches. The highest mitotic figure (HMF), which was introduced Wasan *et al.* [9] as the uppermost relative position where budding was observed, was evaluated here using the relative position measure (see Figure 5.16a). The HMF ranges from 0 to 1 starting from the bottom to the top of the crypt. In contrast, the relative position ranges from 0 to 1 from top to bottom. Hence, the HMF was computed as HMF = 1 – min{relative position}.

The hierarchical characteristics describe the topology graph of the crypt. The structure descriptor shown in Figure 5.16c provides a symbolic representation of the crypt's structure as a sequence of numbers $[b_g] \coloneqq b_0 b_1 b_2 \cdots$, where b_g is the total number of branches with generation g. This symbolic representation allows comparison across crypts to identify the predominant budding structure within a given group. The maximum generation is the largest generation among all branches (or the largest generation g in the structure descriptor) and quantifies the maximum number of repeated branching events in a path from the root to a leaf edge. The number of leaf branches is computed as the total number of branches that have no children, and corresponds to the number of daughter crypts that constitute the crypt. The mean branching multiplicity quantifies the mean number of new branches created at each budding event. It is computed as the mean number of branches spawned by budding events that yield at least one leaf branch (see Figure 5.16b). The definition of the asymmetry index defined by Tan *et al.* [6] was extended for multiple crypts. Initially, what was called 'mean crypt length C' was quantified here as the mean leaf-to-root geodesic distance $\overline{L_\iota}$, $i = 1..n$ where n is the total number of leaf branches (see Figure 5.14d). The normalised differences ΔL_i of the individual leaf-to-root geodesic distances L_i from the mean crypt-to-root geodesic distance $\overline{L_\iota}$ were then calculated as $\Delta L_i = \frac{|L_i - \overline{L_\iota}|}{\overline{L_\iota}}$ and the asymmetry index AI was computed as the mean of the normalised differences $AI = \overline{\Delta L_\iota}$, $i = 1..n$. For easier comparison between the asymmetry index and the symmetry ratio presented next, the asymmetry index is presented as a percentage, i.e. AI × 100 %. The asymmetry index quantifies the amount of asymmetry in the budding crypt (the larger AI, the more asymmetric a crypt is). The symmetry ratio (SR) by Langlands *et al.* [22] (division of the length of the shorter daughter crypt by the length of the longer daughter crypt) was quantified as the ratio of the geodesic length of the shortest of the leaf edges over the geodesic length of the longest of the leaf edges (see Figure 5.14b). This estimates the amount of symmetry in the crypt (the larger the SR, the more symmetric the crypt is), and it is also presented as a percentage, i.e. SR × 100 %. In contrast to the AI, the SR only takes into account the extreme cases of the leaf geodesic lengths. Thus, the SR punishes the most extreme of the asymmetries, in contrast to the AI, which averages out all of the asymmetries in the crypt.

Table 5.2: Global morphological characteristics

Morphological characteristics that reduce the representation of the entire crypt down to a few descriptive characteristics. These are divided into two groups: (i) volumetric characteristics that portray the dimensions of the crypt lumen and (ii) hierarchical characteristics that describe how the crypt lumen is distributed in space.

Measure	Description
Volumetric	**Describe the dimensions of the crypt lumen.**
Bounding box width	The width of the smallest box containing the crypt's lumen.
Bounding box length	The length of the smallest box containing the crypt's lumen.
Crypt height	The height (dimension across the axis perpendicular to the image plane) of the smallest box containing the crypt's lumen.
Total volume	An estimate of the negative space occupied by the entire crypt lumen. A measure of the crypt's size.
Total surface area	An estimate of the entire crypt lumen's surface area. A measure of the epithelial layer's surface area.
Max inscribed ball radius	The radius of the largest ball that can be enclosed within the crypt's lumen. A measure of the thickest part of the lumen.
Mean breadth	The average of the calliper diameter over all directions. A measure of the crypt's extent throughout the mucosa.
Mean leaf geodesic length	The leaf geodesic length averaged over all leaf branches.
Mean leaf-to-root geodesic distance	The leaf geodesic distance to root averaged over all leaf branches.
Mean leaf relative position	The branch relative position averaged over all leaf branches.
Highest mitotic figure	The uppermost relative position where budding is observed (following the definition by Wasan et al. [9], where the relative position is zero at the bottom and one at the top of the crypt: HMF = 1 – min{relative position}).
Hierarchical	**Describe the distribution of the crypt lumen in space.**
Structure descriptor	A symbolic representation of the crypt's structure as a sequence of numbers $[b_g] \coloneqq b_0 b_1 b_2 \cdots$ where b_g is the number of branches with generation g.
Maximum generation	The largest generation among all branches (or the largest generation g in the structure descriptor). It quantifies the maximum number of branching events within a budding crypt.
Number of leaf branches	The total number of branches that have no children.

Measure	Description
Mean branching multiplicity	The mean number of branches spawned by budding events yielding at least one leaf branch. It quantifies the mean number of new branches created in every branching event.
Asymmetry index	The amount of asymmetry in the crypt. Definition of Tan *et al.* [6] extended to multiple branches. Ranges from 0 to 100%.
Symmetry ratio	The amount of symmetry in the crypt. Definition of Langlands *et al.* [22]. Ranges from 0 to 100%.

5.3 Statistical analysis

To reduce the several measurements for a given treatment group down to a small set of descriptive characteristics, the measurements from individual samples were analysed as follows (see Figure 5.19): Six µCT image stacks (Sample 1 to 6), one per animal, were analysed per group (biological replications). The measures discussed in section 5.2 were quantified multiple times, one per segmented crypt or branch, on each sample (technical repetitions).

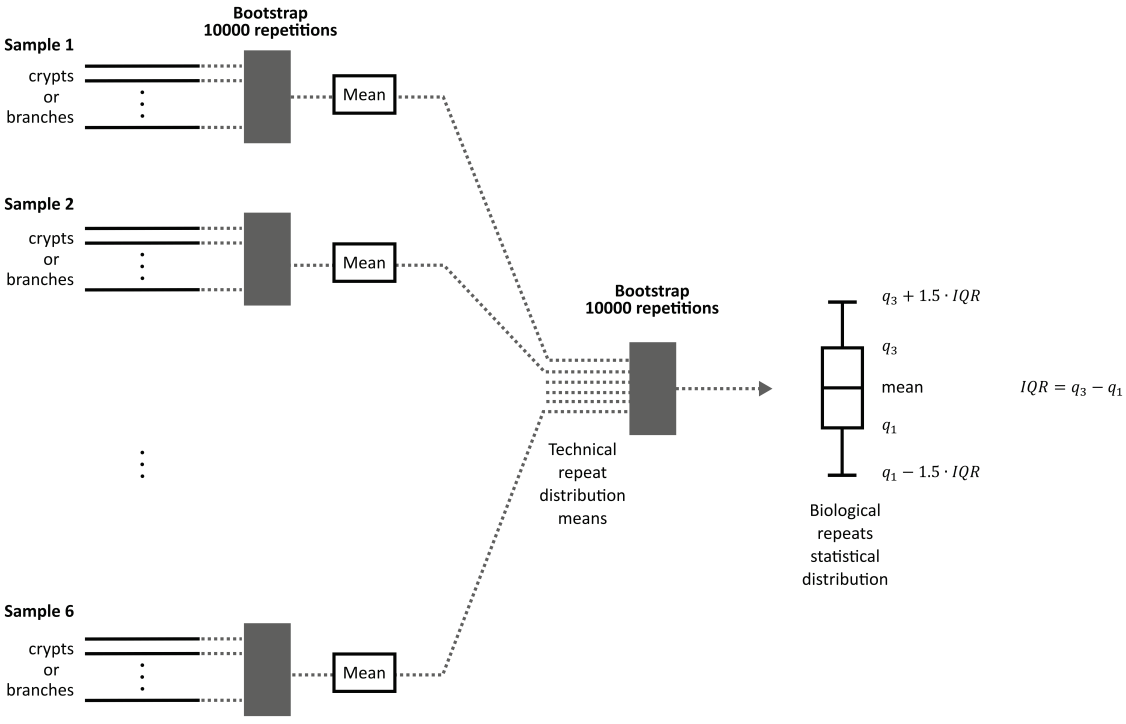


Figure 5.19: Statistical analysis scheme

A statistical distribution, also referred to as sampling distribution, was obtained per treatment group for each characteristic of interest in two steps: (i.) The Bootstrap method was applied on all measurements (technical repetitions) of a given characteristic in each biological sample (biological replication), yielding six

distributions. (ii.) The Bootstrap method was applied once more on the distribution means of the technical repetitions to derive the sampling distribution of each treatment group. These are represented by box-and-whiskers plots, where the middle horizontal line shows the median value of the sampling distribution and the bottom and top vertical lines show the first (q_1) and third (q_3) quartiles, respectively. The inter-quartile range (IQR) is the difference between the two quartiles ($q_3 - q_1$), and the length of the whiskers is equal to 1.5 times the IQR.

Combination of the technical repetitions and biological replications into a single descriptive characteristic per measurement was achieved in two steps: (i) one mean value representing the sample distribution of technical repetitions was estimated for each biological replication, where (ii) these means were combined to estimate the statistical mean representing the treatment group for each measurement. The number of technical repetitions is often marginally small for the sample distribution to be approximately normal for certain measures shown in Table 5.1. For instance, the total number of technical repetitions when the measures are aggregated by generation, e.g. considering only branches with generation equal to 2, might be as low as five or less. The same holds for the number of biological replications (×6). Therefore, the Bootstrap method [263] was applied to estimate both the technical repetition means (i) and the sampling distribution (distribution from which the six samples are assumed to be selected) of each treatment group (ii). Bootstrap is a statistical resampling method with replacement, i.e. it randomly selects one of the samples within a population (10,000 times here) to create a larger sample population. The statistical distribution of this larger sample population is proven [264] to converge to a normal distribution, with means and confidence intervals that are asymptotically more accurate than the standard intervals that would be obtained using the sample mean. After bootstrapping, once at step (i) followed by another at step (ii), one asymptotically normal statistical distribution is obtained per treatment group for each morphological characteristic, also referred to as sampling distribution. Therefore, each treatment group is characterised by a set of characteristics discussed in section '5.2 Definition of budding and non-budding crypt characteristics', which is described by their sampling distributions, and presented using box-and-whiskers plots, as shown in Figure 5.19. Note that as Bootstrapping ensures that the resulting distribution is asymptotically normal, its median value coincides with its mean for exceedingly high number of repetitions, as it is the case here (10,000 times). Therefore, the median values shown in the box-and-whiskers plots are approximately equal to the mean values of the sampling distributions.

A resampling method [265] has been used to overcome the limitations imposed by the small number of samples when performing statistical comparisons. Specifically, a (nonparametric)

permutation test [266, 267] was used to test the null hypothesis that two measures under examination come from the same distribution. The test is based on the premise that any permutation of the experimental measurements is equally likely to be observed, i.e. the order of observation is irrelevant. It also makes no assumption related either to the number of samples or their distributions. Initially, the difference d_{ab} between the mean values of the sampling distributions of measurements a and b is computed, and its distance from unity $d_0 = |1 - d_{ab}|$ is calculated. Next, the two sampling distributions are combined into one pool of 6 + 6 = 12 measurements and all the possible permutations (12! = 479,001,600) are computed. For each permutation, the pool is divided into two sets i and j of size = 6 each and the difference d_{ij} of their means is computed. It can be shown [268] that the *p*-value (the probability that the two groups come from the same distribution) is equal to the number of times the distance of recorded differences from unity $d_k = \left|1 - d_{ij}\right|$ is larger than d_0, divided by the total number of permutations used: $p = \frac{1}{12!}\sum_{k=1}^{12!} I(d_k > d_0)$, where $I(\cdot)$ is the indicator function.

All statistical comparisons between morphological characteristics were performed using the permutation test described here. Statistical significance was inferred after application of the Bonferroni correction [269, 270] for multiple comparisons. This ensures that the increased probability of false-negative results (deciding that there is a statistically significant difference between the mean values of two distributions where there is not), due to multiple comparisons, is counter-balanced by selecting stricter boundaries for the levels of significance. Specifically, the critical levels for significance were divided by the number of comparisons n that were performed on each distribution. Statistical significance is presented as plot annotations following a 'star' notation that is summarised as follows: no significant difference or 'ns' for $\frac{0.05}{n} < p \leq 1$, one star '*' for $\frac{0.01}{n} < p \leq \frac{0.05}{n}$, two stars '**' for $\frac{0.001}{n} < p \leq \frac{0.01}{n}$, three stars '***' for $\frac{0.0001}{n} < p \leq \frac{0.001}{n}$, and four stars '****' for $p \leq \frac{0.0001}{n}$.

Discrete (numeric categorical) measures, such as the number of branches etc., were treated as continuous variables as described above in order to compute the sampling distribution, and subsequently discretised by rounding to the closest integer.

The strength of statistical association between distributions of measures was calculated using the Pearson's correlation coefficient. The correlation coefficient between measures a and b ranges from -1 to +1, with -1 representing a strong anti-correlation (increase of a is associated with a reduction in b and *vice versa*), +1 representing a strong correlation (increase of a is associated with an increase in b and *vice versa*), and 0 representing a weak correlation (change of a is not associated with a change of b and *vice versa*).

Chapter 6 Quantitative morphometry of crypt budding in a mouse model

Chapter 5 establishes the workflows and tools needed for objective morphological quantification of the colonic structure and budding crypts using µCT imaging (Chapter 3) and digital volume unrolling (Chapter 4). In this chapter these workflows and tools are applied on samples from the exploratory animal study of section '2.1 Animal model materials and methods' as a proof of concept to showcase their applicability for the study of the early-stages of colorectal cancer.

Specifically, morphological quantification was applied to the 1st cycle AOM&DSS-treated cancer model (6×) samples. This yield a detailed phenotype that includes the morphological characteristics detailed in section '5.2.2 Branch morphological characteristics' and section '5.2.3 Global morphological characteristics'. As briefly mentioned in Chapter 5, budding is (almost) absent in the untreated control group. Therefore, this chapter focuses on detailed phenotyping of crypt budding in the AOM&DSS-treated cancer model. For a comparison between the untreated control and the AOM&DSS-treated cancer model, the reader is referred to Appendix B: 'Global characterisation of the colonic structure'.

Detailed analysis of the global morphological characteristics is presented in section 6.2.1, followed by detailed analysis of the branch-wise morphological characteristics in section 6.2.2. Section 6.3 discusses the key findings, while certain limitations and outlook are discussed in section 6.4.

6.1 Materials and methods

FFPE samples of intact murine colons were produced as detailed in section '2.1.1 Azoxymethane and dextran sodium sulphate mouse model' and subsequently imaged by synchrotron X-ray phase-contrast imaging as detailed in section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons' (see Table B 1 for a summary of the used colon samples). One sub-stack of 301 µCT slices or 662 µm – long colonic section was randomly selected for unrolling for each animal in the AOM&DSS-treated group. These sub-stacks were digitally unrolled, and the colonic crypt lumens were isolated following the workflow of section '5.1 Workflow for identification and 3D segmentation of budding'. The crypt lumen morphology was then quantified by applying the methods of section '5.2 Definition of budding and non-budding crypt characteristics'.

Subsequently, the statistical analysis described in section 5.3 was performed for each morphological characteristic. In brief, each measure's statistical distribution (technical repetitions)

was estimated via Bootstrapping, and one mean value for each of the 6 biological replications (animal tissue samples) was then computed. The bootstrapping method was subsequently applied once more on the 6 means to compute the sampling distribution of the biological replications for each treatment group. In all plots presented in section 6.2, the computed means corresponding to the 6 samples are presented as full disks (●) and the sampling distributions are presented as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles, respectively, and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The bootstrapping method computes a statistical distribution that approaches the normal distribution. As the median value of the normal distribution is identical to its mean, the median value of the box-plots (midline) approaches the mean value of the statistical distribution. Note that, although the advanced statistical methods applied here provide improved estimates, the results should be interpreted with care concerning their biological implications. This is a consequence of the limitations imposed by the design of the exploratory animal study, which are discussed in section '6.4 Limitations and outlook'.

6.2 Results and discussion

Section '6.2.1 Global phenotyping of budding crypts' deals with the phenotyping of the AOM/DSS budding crypts as a whole, assigning a representative value on each of the measures of section '5.2.3 Global morphological characteristics'. This, in conjunction with section '6.2.2 Branch-wise phenotyping of budding crypts' form a complete picture of the representative budding crypt for the AOM/DSS group. The global crypt measurements are summarised in Table 6.1, and the branch measurements are summarised in Table 6.2.

6.2.1 Global phenotyping of budding crypts

The statistical distribution of the sample mean measurements for the crypt height in Figure 6.1a shows that budding crypts in the AOM/DSS group are segregated in two distinct groups. The budding crypts were thus assigned with a crypt height class: 'short' for crypt height < 211 µm and 'tall' otherwise, resulting in a short group of 160 ± 9 µm (mean ± standard deviation) and a tall group of 255 ± 7 µm (Figure 6.1b).

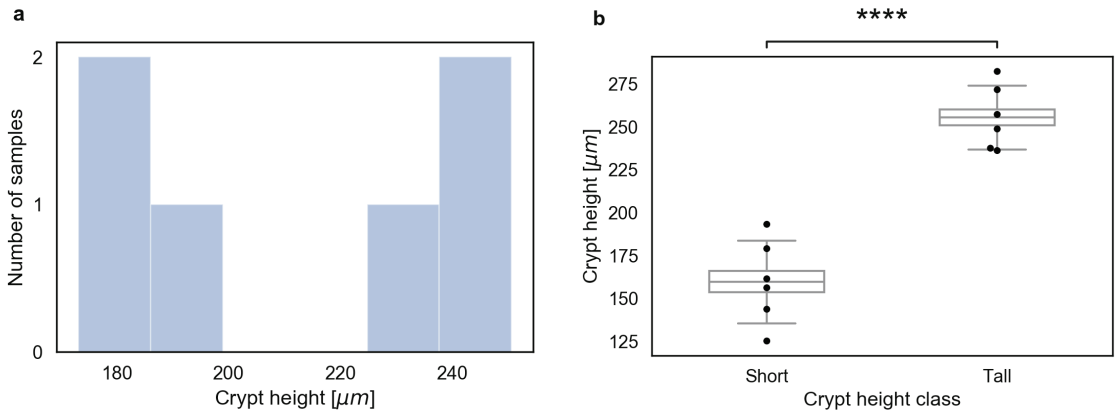


Figure 6.1: Crypt height

(a) The statistical distribution of the crypt height suggests that crypts can be classified as 'short' for crypt height < 211 µm and 'tall' otherwise. (b) This classification yields a short group with a crypt height of 160 ± 9 µm and a tall group with a crypt height of 255 ± 7 µm. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

This classification is then carried-on to the crypt's bounding box width and length as well as the mean breadth, the crypt total volume and total surface area (Figure 6.2). All of these measurements yield a smaller statistical distribution mean value for the short class compared to the tall one (bounding box width 39 ± 4 µm and 49 ± 3 µm, bounding box length 63 ± 7 µm and 75 ± 4 µm, mean breadth 190 ± 20 µm and 290 ± 20 µm, crypt total volume 34000 ± 4000 µm^3 and 62000 ± 6000 µm^3, crypt total surface area 13000 ± 2000 µm^2 and 24000 ± 2000 µm^2 for the short and tall class, respectively). The difference in the mean values of these measurements is statistically significant for the mean breadth, the total volume and total surface area (Figure 6.2c-e), and not statistically significant for the bounding box width and length (Figure 6.2a, b). This suggests that crypts in the tall class are generally larger than those in the short class in terms of total volume, total surface area and mean breadth. They are, however, not wider in terms of bounding box width and length, i.e. the branches do not extend further away from the crypt central axis in the tall class compared to the short class. Though the latter should be interpreted

with caution because one of the 'short' samples is exceedingly larger (outlier) in bounding box width and length, affecting the statistical comparison.

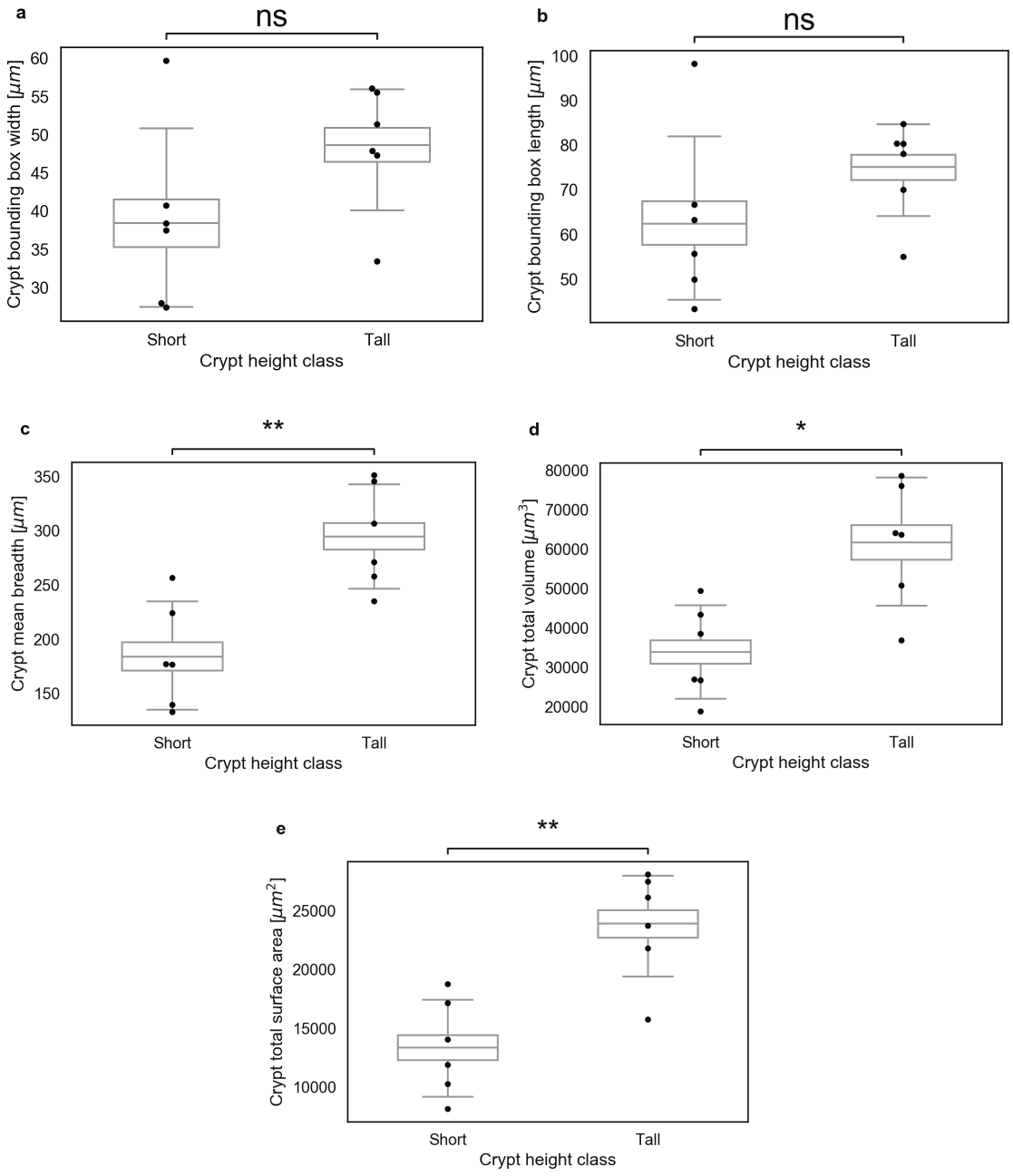


Figure 6.2: Bounding box width and length, total volume and total surface area

(a) The mean and standard deviation of the bounding box width is 39 ± 4 µm and 49 ± 3 µm for the short and tall classes, respectively. Similarly, the bounding box length (b) is 63 ± 7 µm and 75 ± 4 µm, the mean breadth (c) is 190 ± 20 µm and 290 ± 20 µm, the crypt total volume (d) is 34000 ± 4000 μm^3 and 62000 ± 6000 μm^3, the crypt total surface area (e) is 13000 ± 2000 μm^2 and 24000 ± 2000 μm^2 for the short and tall class, respectively. The difference in the mean values of these measurements is statistically significant for the mean breadth, the total volume and total surface area

(c-e), and not statistically significant for the bounding box width and length (a, b). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

The classification between short and tall crypts can be explained by taking a closer look at the locations of individual crypts within the sample. Short crypts tend to be found at locations of mucosal folds, whereas tall ones are found at sites where the mucosa is thicker. As the crypts always extend throughout the entire mucosal tissue (from the interface between the mucosal and muscle layers all the way to the colonic lumen), the crypt height is directly related to the mucosal thickness. It is worth to underline here the crucial role of *in situ*, µCT imaging employed here, which allowed to identify the locations of individual crypts within the tissue. Realising that the crypt height, and thus, the crypt size as measured by the total volume, the total surface area and the mean breadth, is only a consequence of the crypt location, has direct implications in the study of crypt budding. This finding, which is also supported by the analysis of the correlations between the several measures later on (see Figure 6.7), further supports the findings of Pin *et al.* [90, 154], who challenged the traditional belief that crypt fission is initiated after the crypt has reached a critical size (see section '1.3.4 Computational models of the intestine').

Focusing at the global morphology of the branches shows that the thickest part of the crypt is 14 µm (maximum inscribed-ball radius of 7 ± 0 µm) in both the short and tall classes. Note that the pixel size of the µCT scan was 2.2 µm, which makes all measurements in Figure 6.3a virtually identical. The mean geodesic length of the leaf branches is equal to 60 ± 10 µm for the short class and 70 ± 10 µm for the tall one. Although, the difference between the two is not statistically significant, this is again due to an outlier point in the short class (Figure 6.3b). When the outlier sample is disregarded (data not shown here) the two means become 48 ± 4 µm and 79 ± 7 µm for the short and tall classes, respectively, with a statistically significant difference (p=0.00794). The mean leaf relative position (Figure 6.3c) shows at which point the branching occurs relative to the crypt height (starting from the colonic lumen side). The highest mitotic figure (Figure 6.3d) shows at which point the highest of the branches in a crypt is located (starting from the side of the

mucosal and serosal interface[4]). The mean leaf relative position is equal to 0.65 ± 0.04 and 0.73 ± 0.03 for the short and tall classes (difference not statistically significant) and the highest mitotic figure is equal to 0.47 ± 0.06 and 0.40 ± 0.04 (difference not statistically significant). Taken together, these measurements show that the crypts tend to branch on the lower half of the crypt (towards the base) with the highest branching point being located close to the middle. The relative positions of branches are examined in more detail further below, where branches of individual generations are grouped together.

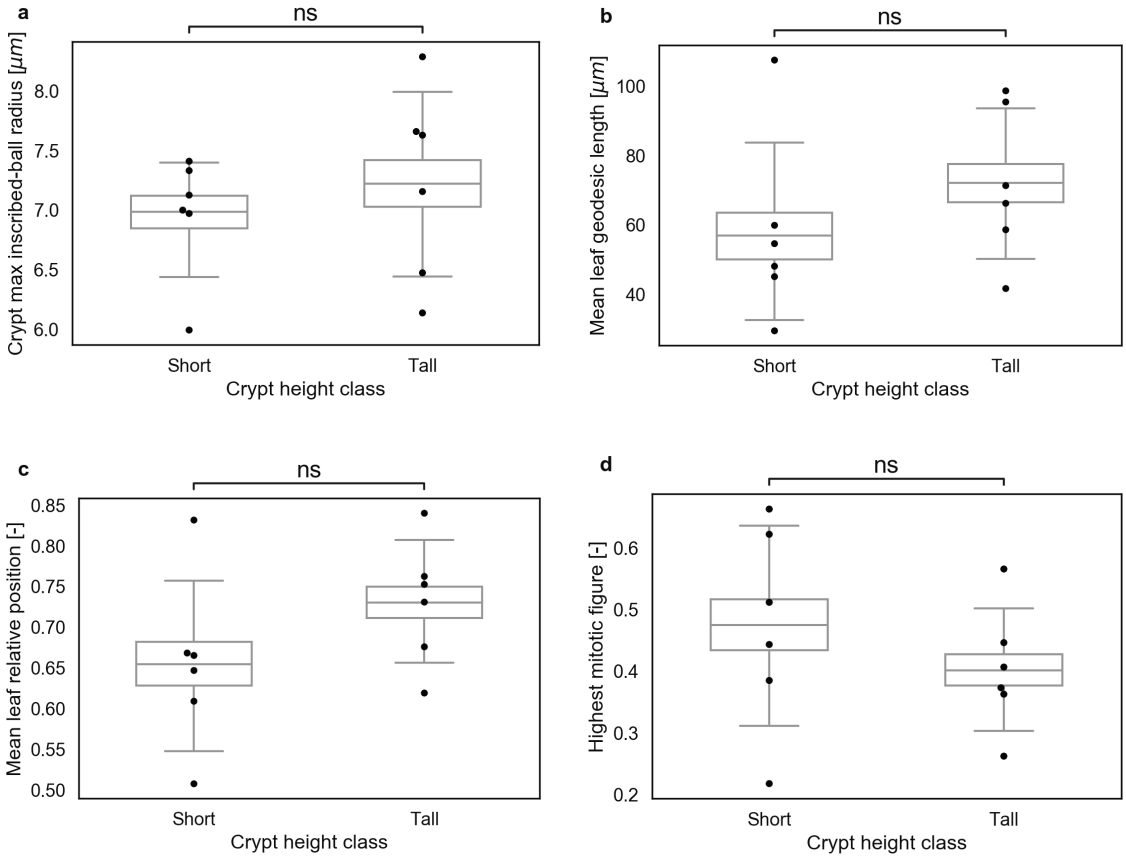


Figure 6.3: Crypt max inscribed-ball radius, mean leaf geodesic length and mean leaf relative position

(a) The maximum inscribed-ball radius is equal to 7 ± 0 μm in both the short and tall classes (the pixel size of the μCT scan was 2.2 μm, making the measurements of the two classes virtually identical). (b) The mean geodesic length of the leaf branches is equal to 60 ± 10 μm for the short class and 70 ± 10 μm for the tall one. (c) The mean leaf relative position (starting from the colonic lumen side) is equal to 0.65 ± 0.04 and 0.73 ± 0.03 for the short and tall classes (difference not statistically significant). (d)

[4] The highest mitotic figure implementation respects the original definition of Wasan *et al.* [6] which is complementary to the mean leaf relative position since it starts measuring from the opposite side.

The highest mitotic figure (starting from the crypt base) is equal to 0.47 ± 0.06 and 0.40 ± 0.04 (difference not statistically significant). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

The total number of successive budding events (maximum generation) is slightly different for the two classes. The majority of crypts in the short class branched once or twice (Figure 6.4a) while crypts in the tall class typically branch two or three times (Figure 6.4b). Based on this, one might be inclined to infer that the number of budding events might be proportional to the crypt height. However, the correlation coefficient between the crypt height and the maximum generation is low (0.35), suggesting that there is no statistical reason to believe that the two are related. Regardless of the height class, the vast majority of budding events yields two branches (Figure 6.4c), which supports the current understanding of crypt budding as a bifurcation process. The result of two successive budding events yielding two branches each is a total number of 3 leaf branches (Figure 6.4d).

These measurements can be better understood by looking at the structure descriptor, which specifies the number of branches in each generation. Figure 6.5a shows the total number of crypts in the AOM/DSS group for each structure descriptor. Interestingly, there is a large range of different budding structures, revealing the great complexity of crypt budding morphology. Structure descriptors that begin with '2' correspond to budding crypts with two trunks, and structure descriptors that are decreasing from one generation to the next (e.g. '2124' decreases from 2 branches in generation zero to 1 in generation one) correspond to branches that merge together. In some cases, branches might split and subsequently merge together (e.g. '12122' starts with a single trunk in generation zero that bifurcates in generation 1 and re-merges to a single branch in generation two) in a structure where the hollow crypt lumen surrounds an 'isle' of tissue.

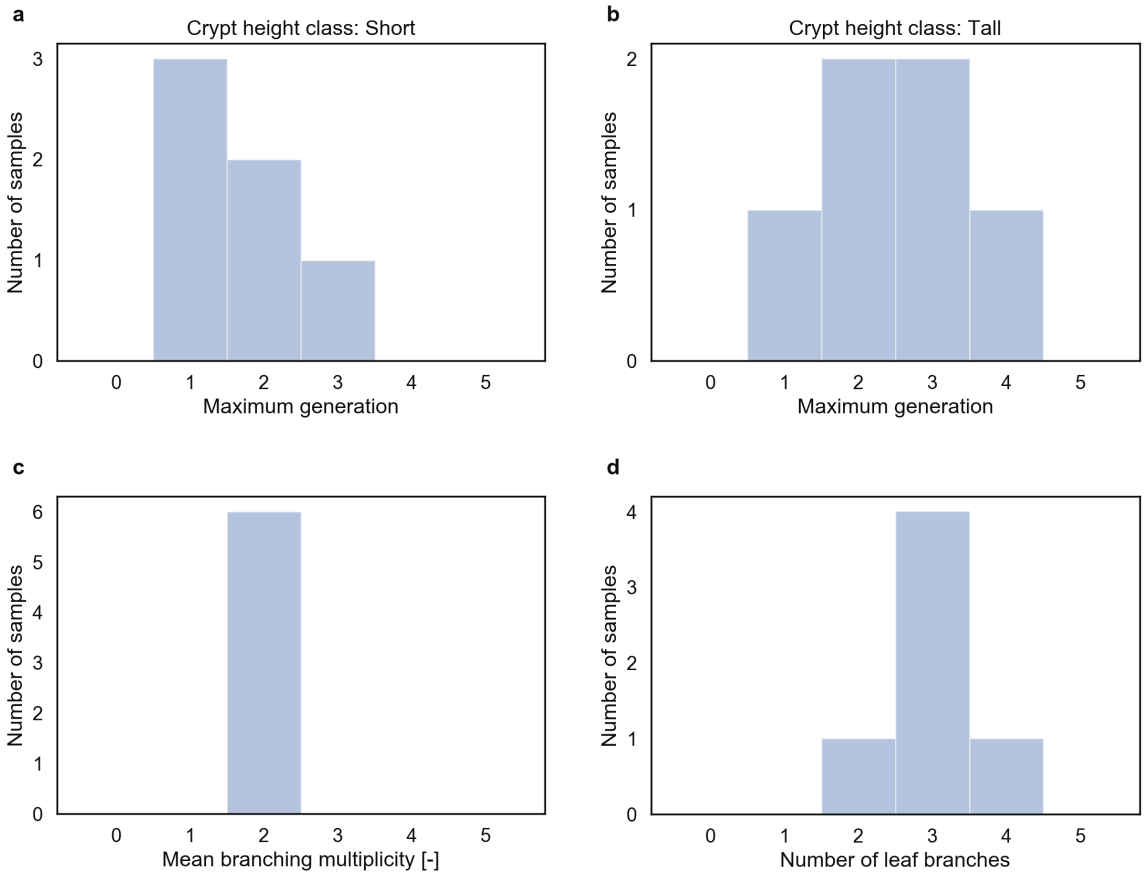


Figure 6.4: Maximum generation, mean branching multiplicity and number of leaves

(a) The maximum generation for the short class ranges from 1 to 3, while the maximum generation in the tall class (b) ranges from 1 to 4. (c) The mean branching multiplicity is independent of the height class, and it is equal to 2, which means that the vast majority of budding events yields two branches (bifurcation). (d) The number of leaf branches ranges from 2 to 4, and it is centred around 3. This is a consequence of typically two successive budding events (maximum generation of 2) yielding two branches each (branching multiplicity of 2).

It is, however, readily apparent that two topologies are dominant, a single bifurcation resulting in two leaf branches and two bifurcations resulting in three leaf branches (‘12’ and ‘122’, also schematically represented by the topology tree of Figure 6.5c&d). Focusing on these two dominant topologies, Figure 6.5b shows the structure descriptors for each sample when the rest of the structure descriptors (all excluding ‘12’ and ‘122’) are collected together in a single group (‘other’). It shows that an assortment of topologies may exist in a single biological sample. Given that ‘12’ represents a single bifurcation, possibly the healthy case of crypt fission, ‘122’ represents an added and abnormal budding event, and ‘other’ represents a whole range of complex topologies, the assortment of structure descriptors in each sample (Figure 6.5b) could be a useful

indicator of the progression of the microstructural deformation in the colonic mucosa. For instance, the fact that AD0 is mainly composed of '12' topologies while AD5 is mainly composed by 'other', suggests that the deformation in AD5 is more severe compared to AD0[5]. Taken together, the results presented in Figure 6.4, Figure 6.5a and Figure 6.5b can be illustrated in the graphic representations of Figure 6.5c and Figure 6.5d for the structure descriptors '12' and '122', respectively. The labels A to C represent the leaf branches, and the numbers above each branch represent the branch's generation.

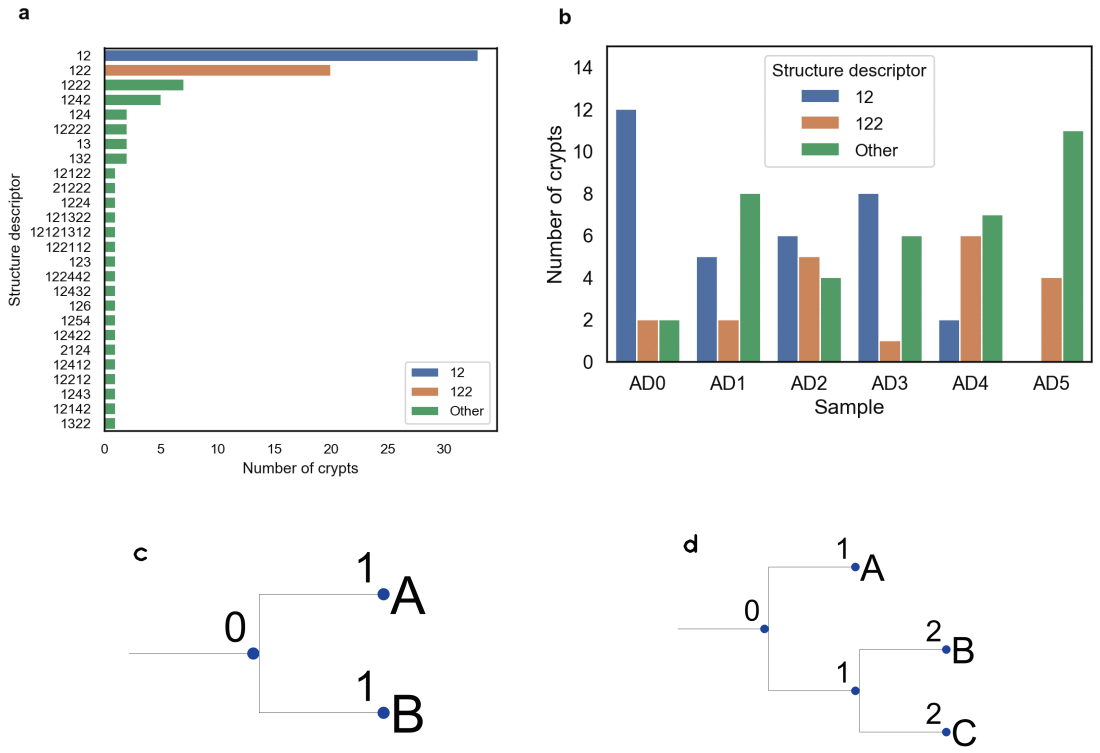


Figure 6.5: Crypt structure descriptor and topology trees

(a) The total number of crypts in the AOM/DSS group for each structure descriptor. Two topologies are dominant: a single bifurcation resulting in two leaf branches '12', and two bifurcations resulting in three leaf branches '122'. In (b) the structure descriptors excluding '12' and '122' for each sample are collected together in a single group ('other'). The topology trees in (c) and (d) correspond to the structure descriptors '12' and '122', respectively. The labels A to C represent the leaf branches, and the numbers above each branch represent the branch's generation.

[5] Note: In order to arrive to robust conclusions following this line of thought, one has to make sure that the quantified crypts have been selected consistently and uniformly for each of the samples. In the case under examination, the crypts were randomly and manually selected without matching the sampling sites or any other factor that might be relevant.

The budding crypts' asymmetry was quantified using the asymmetry index (AI) and the symmetry ratio (SR), shown in Figure 6.6. Crypts in both the short and tall classes have an AI greater than 1 %, which was defined by Tan *et al.* [6] as the lowest threshold to consider a budding crypt as asymmetric. The asymmetry index (Figure 6.6a) was found here much greater, at 9 ± 3 % and 6 ± 1 % for the short and tall classes, respectively (difference not statistically significant). The symmetry ratio (Figure 6.6b) was found equal to 50 ± 9 % and 40 ± 6 % for the short and tall classes, respectively (difference not statistically significant), which is lower than the values (> 55 %) found in organoids by Langlands *et al.* [22].

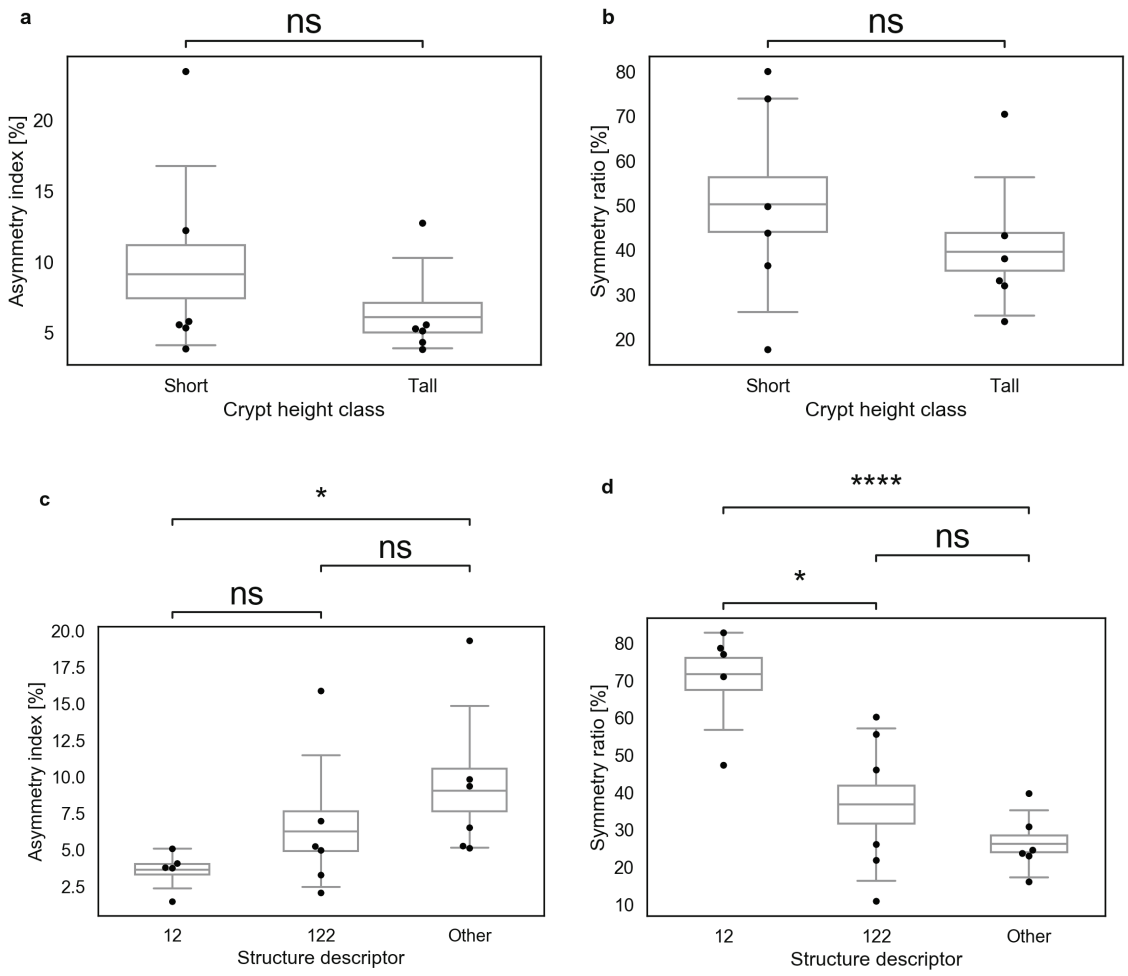


Figure 6.6: Asymmetry index (AI) and symmetry ratio (SR)

(a) The asymmetry index is equal to 9 ± 3 % and 6 ± 1 % for the short and tall classes, respectively (difference not statistically significant). (b) The symmetry ratio is equal to 50 ± 9 % and 40 ± 6 % for the short and tall classes, respectively (difference not statistically significant) (c) Crypts in with the '12' structure descriptor have a tendency towards a lower asymmetry index (4 ± 1 %) compared to the '122' crypts (6 ± 2 %) and the 'other' (9 ± 2 %). Statistical comparison reveals no statistically significant differences between '12' and '122' (p = 0.277), nor between '122' and

'other' (p = 0.381), but a marginally significant difference between '12' and 'other' (p = 0.017). In (d), the symmetry ratio similarly reveals that crypts in the '12' group are more symmetric (71 ± 6 %) than those in the '122' group (35 ± 7 %), which tend to be more symmetric than those in the 'other' group (26 ± 3 %). The differences of the asymmetry ratio distributions between the '12' vs 'other' and '12' vs '122' groups are statistically significant (p = 0.000 and p = 0.015 respectively) while the difference between '122' vs 'other' is not (p = 0.253). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). In (a) and (b) the 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$. In (c) and (d) the critical levels for statistical significance have been adjusted with Bonferroni correction for repeated comparisons (2×), thus the 'star' notation reads: 'ns' for $0.025 < p \leq 0.5$, '*' for $0.005 < p \leq 0.025$.

Taken together, all measurements of the short and tall classes presented so far suggest that both classes represent abnormally budding crypts that are similar in every aspect apart from the crypt height. As discussed earlier, the crypt height is directly related to, or determined by, the crypt's location within the mucosal tissue. Therefore, the crypt height classes, and consequently the crypt locations within the mucosa, do not appear to relate to structurally different crypts. Collected together (disregarding the height class) the budding crypts may also be grouped according to their structure descriptor (Figure 6.5b), which elucidates the relationship between the budding crypt structure and asymmetry. Figure 6.6c shows a tendency of '12' crypts towards a lower asymmetry index (4 ± 1 %) compared to the '122' crypts (6 ± 2 %) and the 'other' (9 ± 2 %). Statistical comparison, though, reveals no statistically significant differences between the groups ('12' vs '122': p = 0.279, '122' vs 'other': p = 0.387, and '12' vs 'other': p = 0.017). Similarly, the symmetry ratio shown in Figure 6.6d reveals that crypts in the '12' group are more symmetric (71 ± 6 %) than those in the '122' group (35 ± 7 %), which tend to be more symmetric than those in the 'other' group (26 ± 3 %). Interestingly, the symmetry ratio achieves better differentiation between the three measurement distributions, with statistically significant differences between the '12' vs 'other' and '12' vs '122' groups ('12' vs '122': p = 0.015, '122' vs 'other': p = 0.253, and '12' vs 'other': p = 0.000). As mentioned in section '5.2.3 Global morphological characteristics', the

asymmetry index takes into account all of the leaf branches and calculates their geodesic length difference from an estimated mean geodesic length of the entire budding crypt. The symmetry ratio takes an alternative approach of computing the ratio between only the shortest and the longest of the leaf branches geodesic lengths, disregarding the rest. Therefore, the asymmetry index is more conservative than the symmetry ratio as the former averages out the asymmetries between the branches, and the later punishes the most extreme among the asymmetries. Taken together, the measurements of the asymmetry index and the symmetry ratio suggest that '12' crypts tend to be close to perfectly symmetric and more symmetric than '122', which tend to be similar or slightly more symmetric than 'other'.

Figure 6.7 shows the Pearson's correlation coefficients (PCC) in the form of a correlation matrix for the global morphological measurements, facilitating a qualitative interpretation of the morphological characteristic's inter-relationships. The PCC's have been computed by collecting the measurements of all budding crypts (AOM/DSS group) without differentiating between technical and biological replications. Thus, the analysis that follows does not serve either as a comparison between health and disease nor as a hypothesis-testing approach for crypt budding. Instead, by revealing trends within the budding crypt morphology, it provides insights that may subsequently be experimentally tested by designing dedicated experimental procedures.

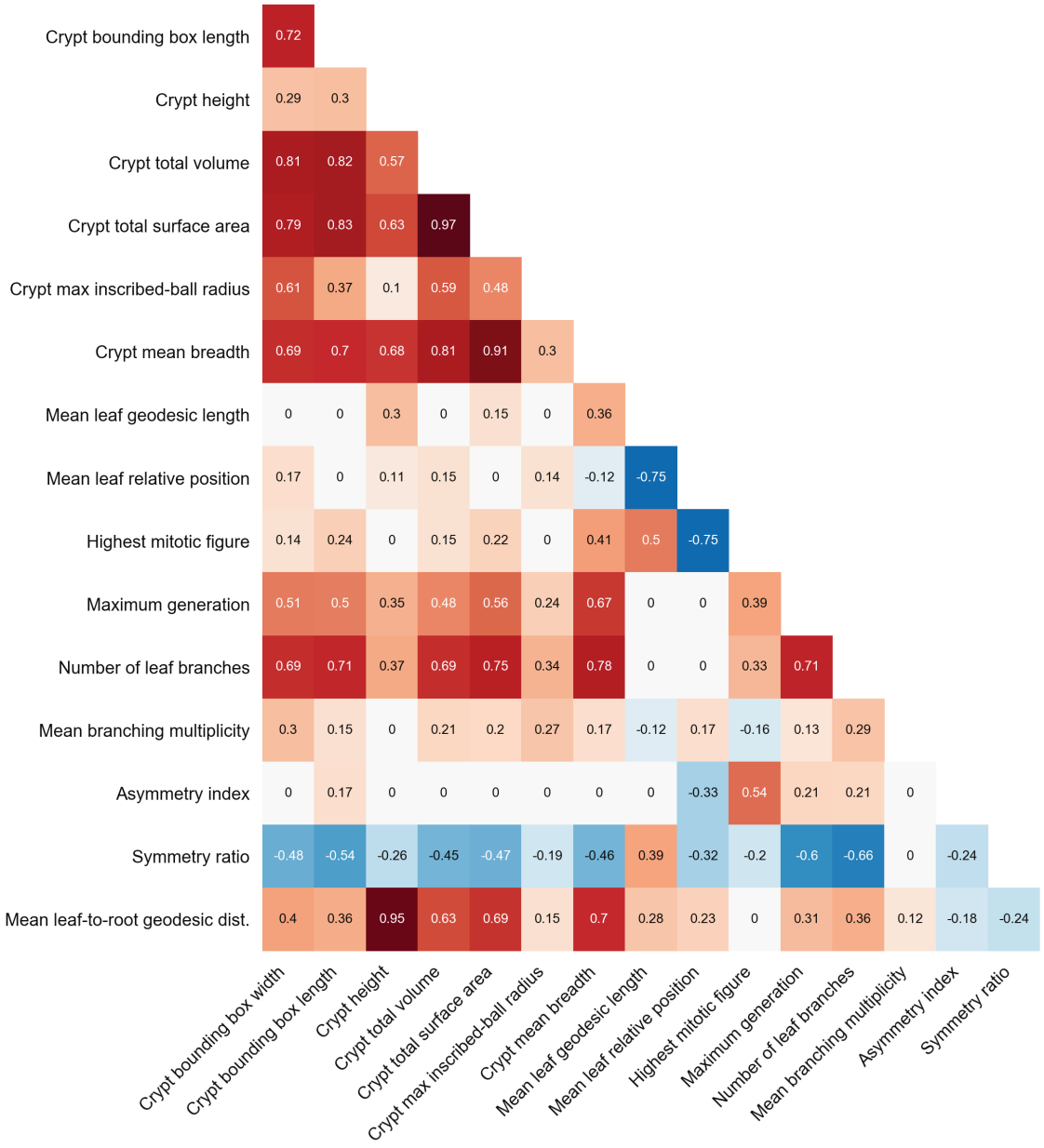


Figure 6.7: Global crypt measurements correlation matrix

The correlation matrix enables a qualitative evaluation of relationships between the global crypt measurements. Each box shows the Pearson correlation coefficient (PCC) between the measurements at the corresponding row and column. The PCC ranges from -1 to +1, with -1 indicating strong anti-correlation, 0 no correlation, and +1 strong correlation (see also section 5.3). Any conclusions made from this correlation matrix are purely indicative of relationships and should be interpreted with care before they are validated with dedicated experimental approaches.

The PCCs between the total volume and the bounding box width and length are high at 0.81 and 0.82, respectively. Similarly, the PCC between the total surface area and the bounding box width and length are high at 0.79 and 0.83, respectively, as well as the PCC between the total volume

and the total surface area at 0.97. Taken together, these high correlations show that larger crypts (larger in terms of volume and surface area) tend to extend further to the sides throughout the mucosa. Interestingly though, the crypt height correlates poorly with the bounding box width and length, with PCCs at 0.29 and 0.30, respectively. This means that taller crypts (those at places where the mucosal thickness is higher) do not expand sidewise. They are also not thicker, with a PCC between crypt height and maximum inscribed radius of 0.1, and uncorrelated with all topological characteristics (leaf relative position, highest mitotic figure, maximum generation, number of leaf branches, mean branching multiplicity, asymmetry index and symmetry ratio). These observations support a previous conclusion that crypts in the tall class were not dissimilar to those in the short class. In fact, the crypt height only correlates highly with the mean breadth (mean calliper diameter across all directions) and the mean leaf-to-root geodesic distance, which follow naturally due to geometrical reasons.

Another interesting observation is the general behaviour of the maximum inscribed ball radius. It only correlates moderately with the total volume and total surface area, with PCC at 0.59 and 0.48, respectively. This shows that the crypt's branches width is not constant, thus wherever the widest part is increased (larger maximum inscribed ball radius) this happens only locally rather than throughout the entire lumen. Apart from the bounding box width, the maximum inscribed ball radius correlates poorly with all other measurements, suggesting that the width of the crypt's branches does not have a strong relationship with any of the pathological characteristics manifested here (HMF, number of leaf branches, branching multiplicity, AI, SR).

The crypt mean breadth correlates highly with the bounding box width, length, crypt height, total volume and surface area resulting from the crypt's geometry. Importantly, it correlates highly with the maximum generation and number of leaf branches, with a PCC of 0.67 and 0.78, respectively. This suggests that the crypt extends further to the sides with every new branch. The mean leaf geodesic length correlates poorly with the measures of the crypt's size (bounding box width, bounding box length, crypt height, total volume, total surface area, and mean breadth). This suggests that the crypt grows due to multiple branching events, rather than because of leaf branches increasing in size. A high negative correlation with the mean relative position (PCC=-0.75) shows that the further from the root a new branch is located the shorter it is.

Another interesting observation is that the highest mitotic figure does not correlate well with the maximum generation, with a PCC of 0.39. This suggests that new branches do not affect the branching point of the old ones. This is also supported by the low correlation between the highest mitotic figure and the number of leaf branches with a PCC of 0.33, which suggests that new branches do not affect the branching point of the uppermost one. Taken together, these

observations support that the branching point of a branch created by a budding event does not move past its parent branch. Specifically, a budding event at a branch with generation g would create a branch with generation $g + 1$. According to the concept of independently growing crypts suggested by Edwards and Chapman [153] the new branch (or daughter crypt) would continue growing and moving upwards, which means that its branching point would decrease. The observations made here suggest that this branching point would not decrease further to reach branches with generations less than g.

The number of branches is highly correlated with the bounding box width and length, total volume, total surface area and the mean breadth with a PCC of 0.69, 0.71, 0.69, 0.75 and 0.78, respectively. In contrast, it poorly correlates with the crypt height and the maximum inscribed-ball radius with a PCC of 0.37 and 0.34, respectively. These suggest that crypts grow wider and larger but not taller as a result of repeated budding, with branches of unchanged width.

The number of branches is highly anti-correlated with the symmetry ratio with a PCC of -0.66, suggesting that crypts with more branches tend to be more asymmetric. Interestingly though, it poorly correlates with the asymmetry index. This shows that the two measures of asymmetry do not agree with each other, which is also supported by the low correlation among them with a PCC of -0.24. In fact, as discussed earlier, the asymmetry index seems to counter-balance asymmetries from multiple branches, yielding a more conservative estimation. Notably, the asymmetry index poorly correlates with all measurements apart from the HMF, which yields a moderate correlation, with a PCC of 0.54. In contrast, the symmetry ratio is more sensitive to the morphological characteristics of the bounding box width, length, total volume, total surface area and mean breadth, with a PCC of -0.48, -0.54, -0.45, -0.47 and -0.46, respectively. Importantly, it shows a high anti-correlation with the maximum generation and the number of leaf branches, with a PCC of -0.6 and -0.66, suggesting that it is a better measure of asymmetry when compared to the asymmetry index.

One of the most interesting observations in this correlation matrix is that the mean branching multiplicity poorly correlates with all of the other measures. This means that every time the crypt branches it typically splits in two (as shown by a mean branching multiplicity of two), remaining unaffected by all other factors. This supports the belief that asymmetric budding is a result of abnormal fission, i.e. deregulation of a bifurcation process (see Chapter 1).

Finally, the mean leaf-to-root distance is almost perfectly correlated with the crypt height, with a PCC of 0.95. This is a consequence of the fact that the leaf branch bases are always located at the interface between the mucosal and muscle layer. As the crypt height is imposed by the mucosal thickness, the same holds for the distance between the crypt's root and the leaf branch bases,

resulting in a very high correlation between the leaf-to-root distance and the crypt height. Another manifestation of this geometrical constraint is the low correlation between the mean leaf-to-root distance and the mean leaf geodesic length, or the leaf relative position with a PCC of 0.28 and 0.23, respectively. It shows that the geodesic distance between the leaf branch bases and the crypt root remains unchanged independently of the leaf geodesic length or relative position.

Taken together, the correlations between the global measurements draw a picture of a budding crypt that grows sidewise as a result of repeated bifurcations, yielding branches of similar width, growing independently, and having their bases fixed at the interface of the mucosal and muscle layers. Next, a closer look at the branches aggregated by their generation better elucidates the structure of budding crypts.

Table 6.1: Representative budding crypt global morphological measurements

The measurements of the global morphological characteristics for the AOM/DSS group at the end of cycle 1 of the animal study detailed in section 2.1.1. The mean ± standard deviation of each measurement, estimated following the process of section 5.3, is given for the short and tall crypts ($<$ 211 μm and $\geq$ 211 μm in crypt height, respectively). The results of statistical comparison with a permutation test are also given in the form of the corresponding p-value. Statistical significance was evaluated as detailed in section 5.3. In brief, mean value differences when tests yield p-values greater than 0.05 were deemed not statistically significant). In the case of the asymmetry index and the symmetry ratio, statistical comparison was also performed between the three structure descriptors' distributions. In this case, the critical levels for statistical significance were adjusted via Bonferroni correction to $\frac{0.05}{2}$: The relevant measurements of the asymmetry index (*) are 4 ± 1 %, 6 ± 2 % and 9 ± 2 % for the '12', the '122' and the 'other' descriptors, respectively. The statistical comparison revealed no statistically significant differences between the groups ('12' vs '122': p = 0.279, '122' vs 'other': p = 0.387 and '12' vs 'other': p = 0.017). Similarly, the symmetry ratio (**) measurements were 71 ± 6 %, 35 ± 7 % and 26 ± 3 % for the '12', the '122' and the 'other' descriptors respectively. The relevant comparisons revealed statistically significant differences between the '12' vs 'other' and '12' vs '122' groups ('12' vs '122': p = 0.015, '122' vs 'other': p = 0.253 and '12' vs 'other': p = 0.000).

Measure	Crypt height class: Short	Crypt height class: Tall	Permutation test p-value
Volumetric			
Bounding box width [µm]	39 ± 4	49 ± 3	0.113
Bounding box length [µm]	63 ± 7	75 ± 4	0.212
Crypt height [µm]	160 ± 9	255 ± 7	0.000
Total volume [µm^3]	34000 ± 4000	62000 ± 6000	0.011
Total surface area [µm^2]	13000 ± 2000	24000 ± 2000	0.004
Max inscribed ball radius [µm]	7 ± 0	7 ± 0	0.535
Mean breadth [µm]	190 ± 20	290 ± 20	0.002
Mean leaf geodesic length [µm]	60 ± 10	70 ± 10	0.314
Mean leaf relative position [-]	0.65 ± 0.04	0.73 ± 0.03	0.180
Highest mitotic figure [-]	0.47 ± 0.06	0.40 ± 0.04	0.381
Hierarchical			
Structure descriptor	‘12’, ‘122’	‘12’, ‘122’	N/A
Maximum generation	1 to 3	1 to 4	N/A
Number of leaf branches	3 ± 1	3 ± 1	N/A
Mean branching multiplicity [-]	2 ± 0	2 ± 0	N/A
Asymmetry index [%] *	9 ± 3	6 ± 1	0.379
Symmetry ratio [%] **	50 ± 9	40 ± 6	0.387

6.2.2 Branch-wise phenotyping of budding crypts

Global phenotyping (see section 6.2.1) builds the big-picture of the 3D shape of budding crypts during the early developmental stages of CRC. To better understand the branching behaviour, which is essential for computational models, it is necessary to examine individual branches. This section deals with the branch-wise phenotyping of the AOM/DSS budding crypts, assigning a representative value on each of the measures of section ‘5.2.2 Branch morphological characteristics’, which are summarised in Table 6.2.

Leaf branches play a pivotal role in budding crypts since they represent the individual (daughter) crypts that make up the budding crypt structure. Thus, the analysis that follows focuses on comparing the morphological characteristics of leaf branches to non-leaf branches at each hierarchical level (generation). This allows an in-depth understanding of the budding crypt

structure and identification of the crucial morphological differences between leaf and non-leaf branches. As shown in section 6.2.1, the predominant budding structures ('12' and '122') include generations 0 to 2 forming a tree structure such as that shown in Figure 6.5c&d. Both budding structures begin from a tree trunk at generation 0 and branch from generations 1 onwards, which means that there are no leaf branches with generation 0. Plots in this section include the morphological measurements of branches with generations 0 to 2. Generation 0 only includes non-leaf branches (the trunk), and those with generations 1 and 2 are arranged in pairs of leaf/non-leaf branches, along with the corresponding statistical comparison of their means (see permutation test, section 5.3).

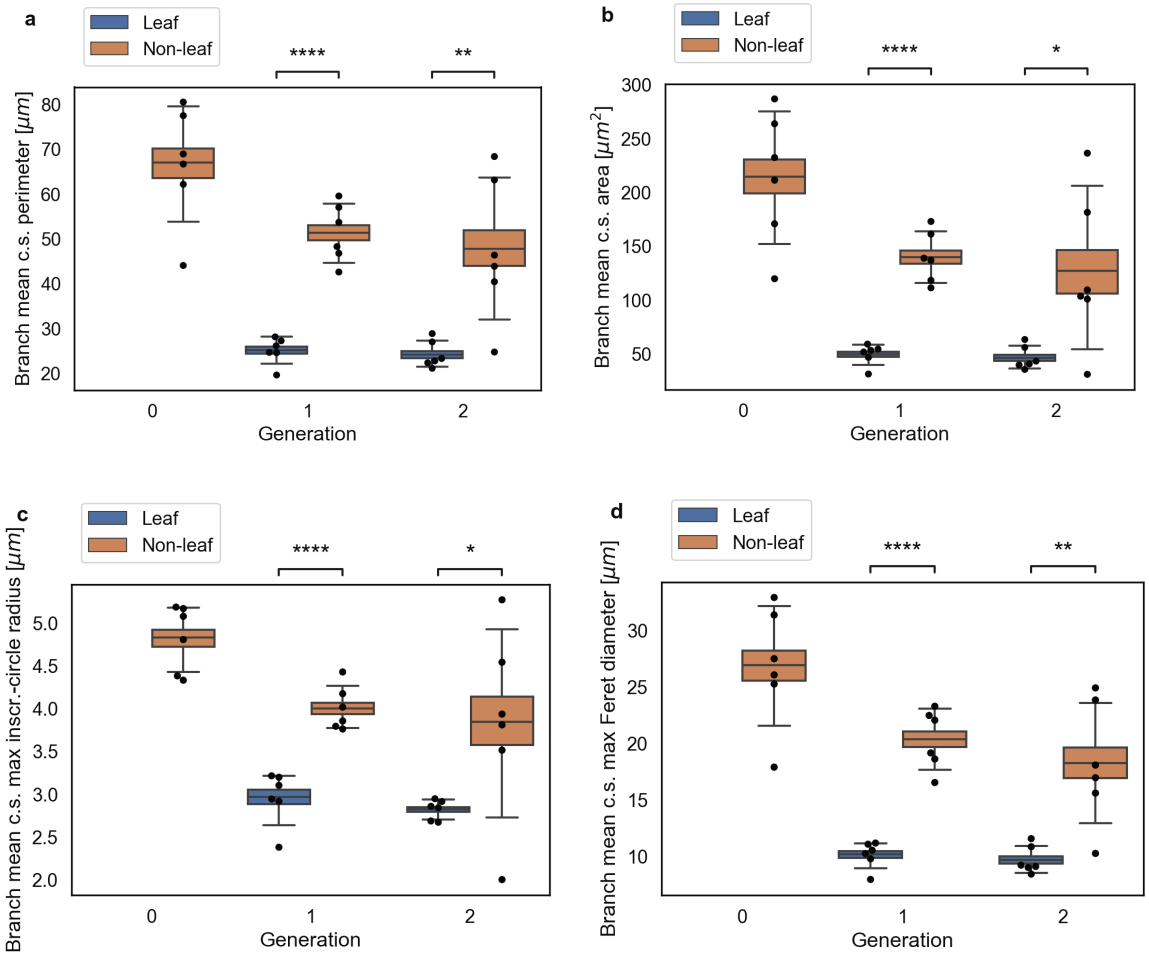


Figure 6.8: Branch cross-section size measurements

(a) The mean cross-section perimeter of non-leaf branches declines from generation 0 to 2 (67 ± 5 µm, 51 ± 2 µm, 48 ± 6 µm), but those are constantly larger compared to the mean cross-section perimeters of leaf branches, which remain constant at 25 ± 1 µm and 24 ± 1 µm for generations 1 and 2, respectively. (b) The mean cross-section area follows a similar trend with 210 ± 20 µm^2, 140 ± 10 µm^2 and 130 ± 30 µm^2 for the non-leaf branches of generation 0 to 2, respectively, and 49 ± 4 µm^2 and 46 ± 4

μm^2 for the leaf branches of generation 1 and 2, respectively. (c) The mean cross-section maximum inscribed circle radius drops from 5 ± 0 µm at generation 0, to 4 ± 0 µm at generation 1 and 2 for non-leaf branches, while it remains constant at 3 ± 0 µm for leaf branches of generations 1 and 2. (d) The mean cross-section maximum Feret diameter declines from generation 0 to 2 (27 ± 2 µm, 20 ± 1 µm, 18 ± 2 µm) for non-leaf branches, and remains constant at 10 ± 0 µm for leaf branches of generation 1 and 2. All differences between the leaf and non-leaf measurement distribution means within generations 1 and 2, respectively, are statistically significant. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles, respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). Box-and-whiskers plots appear in pairs for each generation, with those on the left (blue) corresponding to leaf branches and those on the right (orange) corresponding to non-leaf branches. The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

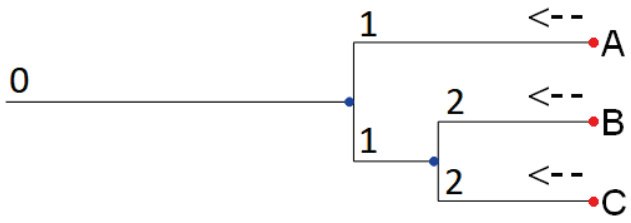


Figure 6.9: Migrating cell flows

A stem cell niche at the base of leaf branches (A, B, C) produces cells that differentiate and migrate upwards towards the colonic lumen. The migrating cell flows from individual branches are added up at each branching point, thus requiring larger cross-section perimeters with decreasing generations to accommodate them. Numbers above the branches correspond to their generation, arrows show the cell flows, blue disks show the branching points, and red disks show the stem cell niches.

The first of the categories of the branch morphological characteristics detailed in section 5.2.2 refers to 2D cross-sections taken along each budding crypt. Figure 6.8 shows the characteristics that quantify the size of these cross-sections (perimeter, area, maximum-inscribed circle radius and maximum Feret diameter), combined as the means across all cross-sections of each branch.

The cross-section perimeter may be considered proportional to the number of epithelial cells surrounding the lumen at each cross-section[6]. Figure 6.8a shows that the mean cross-section perimeter of non-leaf branches is equal to 67 ± 5 μm, 51 ± 2 μm and 48 ± 6 μm for generations 0, 1 and 2, respectively. These are constantly larger compared to the mean cross-section perimeters of leaf branches, which remain constant at 25 ± 1 μm and 24 ± 1 μm for generations 1 and 2, respectively. The mean cross-section area (Figure 6.8b) follows a similar trend with 210 ± 20 μm^2, 140 ± 10 μm^2 and 130 ± 30 μm^2 for the non-leaf branches of generation 0 to 2, respectively, and 49 ± 4 μm^2 and 46 ± 4 μm^2 for the leaf branches of generation 1 and 2, respectively. The mean cross-section maximum inscribed circle radius (Figure 6.8c) drops from 5 ± 0 μm at generation 0 to 4 ± 0 μm at generation 1 and 2 for non-leaf branches, while it remains constant at 3 ± 0 μm for leaf branches of generations 1 and 2. Finally, the mean cross-section maximum Feret diameter (Figure 6.8d) declines from generation 0 to 2 (27 ± 2 μm, 20 ± 1 μm, 18 ± 2 μm) for non-leaf branches, and remains constant at 10 ± 0 μm for leaf branches of generation 1 and 2. All differences between the leaf and non-leaf measurement distribution means within generations 1 and 2, respectively were statistically significant in Figure 6.8. Taken together, these measurements begin to draw a picture of leaf branches with a cross-section size that remains unaffected by the generation. These leaf branches are attached to a central crypt body with cross-sections that increase in size with decreasing generation (from the bottom up). This abstract picture of the budding crypt structure, which is further supported by the volumetric measurements discussed later-on, supports the assumption that there is a stem cell niche at the base of leaf branches producing cells that differentiate and migrate upwards towards the colonic lumen [6, 22, 52]. In turn, a speculation can be made that migrating cell flows (Figure 6.9) from individual branches are added up at each branching point, thus requiring larger cross-section perimeters with decreasing generations (or equivalently larger surface areas discussed later-on) to accommodate them.

[6] This assumes that the size of the epithelial cells is constant, which is a very coarse approximation of reality, since several types (absorptive, goblet, etc.) of cells exist within the crypt epithelium [24].

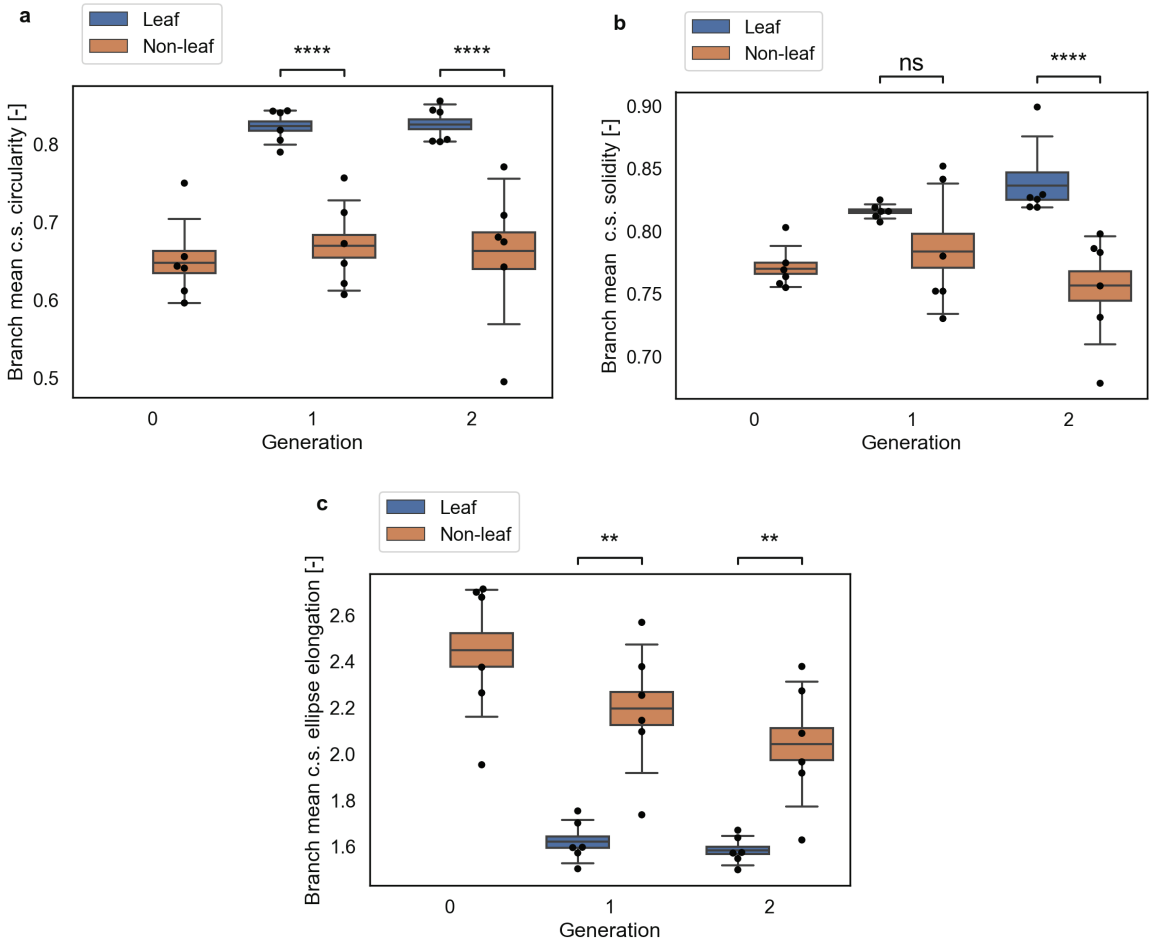


Figure 6.10: Branch cross-section shape measurements

(a) The mean cross-section circularity of the non-leaf branches (0.65 ± 0.02, 0.67 ± 0.02 and 0.66 ± 0.03 for generations 0 to 2, respectively) is consistently lower than that of leaf branches (0.82 ± 0.01 and 0.83 ± 0.01 for generations 1 and 2, respectively). The difference between the mean cross-section circularity of leaf and non-leaf branches of individual generations are statistically significant. (b) The mean cross-section solidity remains relatively constant at 0.8 ± 0.1, 0.78 ± 0.02 and 0.76 ± 0.02 for non-leaf branches with generations 0 to 2 while the leaf branches of generations 1 and 2 have a mean cross-section solidity of 0.82 ± 0.00 and 0.84 ± 0.01, respectively. The difference between the mean cross-section solidity of leaf and non-leaf branches of generation 1 is statistically significant, whereas it is not statistically significant for those of generation 2. (c) The mean cross-section ellipse elongation of non-leaf branches is equal to 2.5 ± 0.1, 2.2 ± 0.1 and 2.0 ± 0.1 for generations 0 to 2, respectively and the mean cross-section ellipse elongation of the leaf branches is equal to 1.62 ± 0.03 and 1.58 ± 0.02 for generations 1 and 2, respectively. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication

means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles, respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). Box-and-whiskers plots appear in pairs for each generation, with those on the left (blue) corresponding to leaf branches and those on the right (orange) corresponding to non-leaf branches. The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

Figure 6.10 shows measures that quantify the shape of the branch lumen cross-sections. The mean cross-section circularity (Figure 6.10a) of the non-leaf branches (0.65 ± 0.02, 0.67 ± 0.02 and 0.66 ± 0.03 for generations 0 to 2, respectively) is consistently lower compared to that of leaf branches (0.82 ± 0.01 and 0.83 ± 0.01 for generations 1 and 2, respectively). These measurements show that the leaf branches tend to have (almost) circular cross-sections, whereas the central crypt body tends to be less-circular. This is also supported by the mean cross-section ellipse elongation (Figure 6.10c), which draws a picture of slender non-leaf branches (2.5 ± 0.1, 2.2 ± 0.1 and 2.0 ± 0.1 for generations 0 to 2, respectively) with one of their dimensions about twice the size of the other. In contrast, the mean cross-section ellipse elongation of the leaf branches is equal to 1.62 ± 0.03 and 1.58 ± 0.02 for generations 1 and 2, respectively, which means that the leaf branches have less elongated cross-sections with one of their dimensions about 1.6 times the size of the other. The mean cross-section solidity (Figure 6.10b), on the other hand, is less distinct between leaf and non-leaf branches of different generations. For the non-leaf branches, the mean cross-section solidity remains relatively constant at 0.8 ± 0.1, 0.78 ± 0.02 and 0.76 ± 0.02 for generations 0 to 2, respectively. The mean cross-section solidity of leaf branches is equal to 0.82 ± 0.00 and 0.84 ± 0.01 for generation 1 and 2, respectively. The difference between the mean cross-section solidity of leaf and non-leaf branches of generation 1 is statistically significant, whereas it is not statistically significant for those of generation 2.These measurements are consistent with a slightly elongated, ellipse-like shape for the non-leaf branches, also supported by the mean cross-section circularity and mean cross-section ellipse elongation. However, the fact that the mean cross-section solidity of the non-leaf branches does not follow the same decreasing trend as the mean cross-section elongation shows that the cross-sectional shapes of non-leaf branches are over-simplified by the fitted ellipse representations. This means that if modelling the exact branch shape is of interest, a closer investigation of the cross-section shapes needs to be employed.

Figure 6.11 shows measurements of the branches' size (volume, surface area, mean breadth and maximum inscribed-ball radius). It is readily seen that the trunk (branch with generation 0) is

larger than all other branches in generations 1 and 2, both in terms of volume and surface area. Figure 6.11a shows a decreasing trend for the branch volume of 21000 ± 3000 μm^3, 5600 ± 400 μm^3 and 2500 ± 400 μm^3 for non-leaf branches with generations 0 to 2, respectively, and a branch volume of 1700 ± 200 μm^3 and 1100 ± 200 μm^3 for leaf branches with generations 1 and 2, respectively. The difference between the volumes of leaf and non-leaf branches of individual generations are statistically significant. Similarly, the surface area shown in Figure 6.11b follows a decreasing trend of 7600 ± 800 μm^2, 2500 ± 200 μm^2 and 1300 ± 100 μm^2 for non-leaf branches with generations 0 to 2, respectively, and 1020 ± 90 μm^2 and 670 ± 80 μm^2 for leaf branches with generations 1 and 2, respectively. The difference between the surface areas of leaf and non-leaf branches of individual generations are also statistically significant. Figure 6.11c shows that the mean breadth of non-leaf branches also drops with increasing generations (85 ± 6 μm, 40 ± 1 μm and 30 ± 1 μm for generations 0 to 2, respectively), as is the case for the mean breadth of leaf branches (28 ± 2 μm and 20 ± 1 μm for generations 1 and 2, respectively). The difference between the mean breadth of leaf and non-leaf branches of individual generations are statistically significant. Taken together, the volume, surface area and mean breadth measurements reveal an overall size of individual branches that decreases with generation. This supports the statement made earlier, that differentiated cells flow through the branches' epithelium from the base to the colonic lumen. Interestingly though, although the cross-section perimeter of the leaf branches remains constant, as discussed earlier, their surface area decreases with increasing generations. This means that their length should be decreasing as well, which will be discussed later-on by examining the branches' geodesic length.

Figure 6.11d shows that the maximum inscribed-ball radius also decreases slightly with generations (6 ± 0 μm, 5 ± 0 μm, 4 ± 0 μm for non-leaf branches with generations 0 to 2, respectively, and 4 ± 0 μm, 3 ± 0 μm for leaf branches with generations 1 and 2, respectively). The difference between the maximum inscribed-ball radius of leaf and non-leaf branches of generation 1 is statistically significant, unlike the difference between those of generation 2. Note, however, that although the bootstrapping method (see section 5.3) yield an estimation of zero standard deviation for the maximum inscribed-ball radius of non-leaf branches with generation equal to 2, the variability in the mean values of the biological replications (solid disks on Figure 6.11d) is exceedingly large. This shows that the morphological deformation of these branches is advanced, with large variability in their shapes. In conjunction with the observation made earlier that the cross-sectional shapes are over-simplified by elliptical representations, this observation underlines the necessity for improved modelling of the branch's shape when it comes to computational modelling of the crypt.

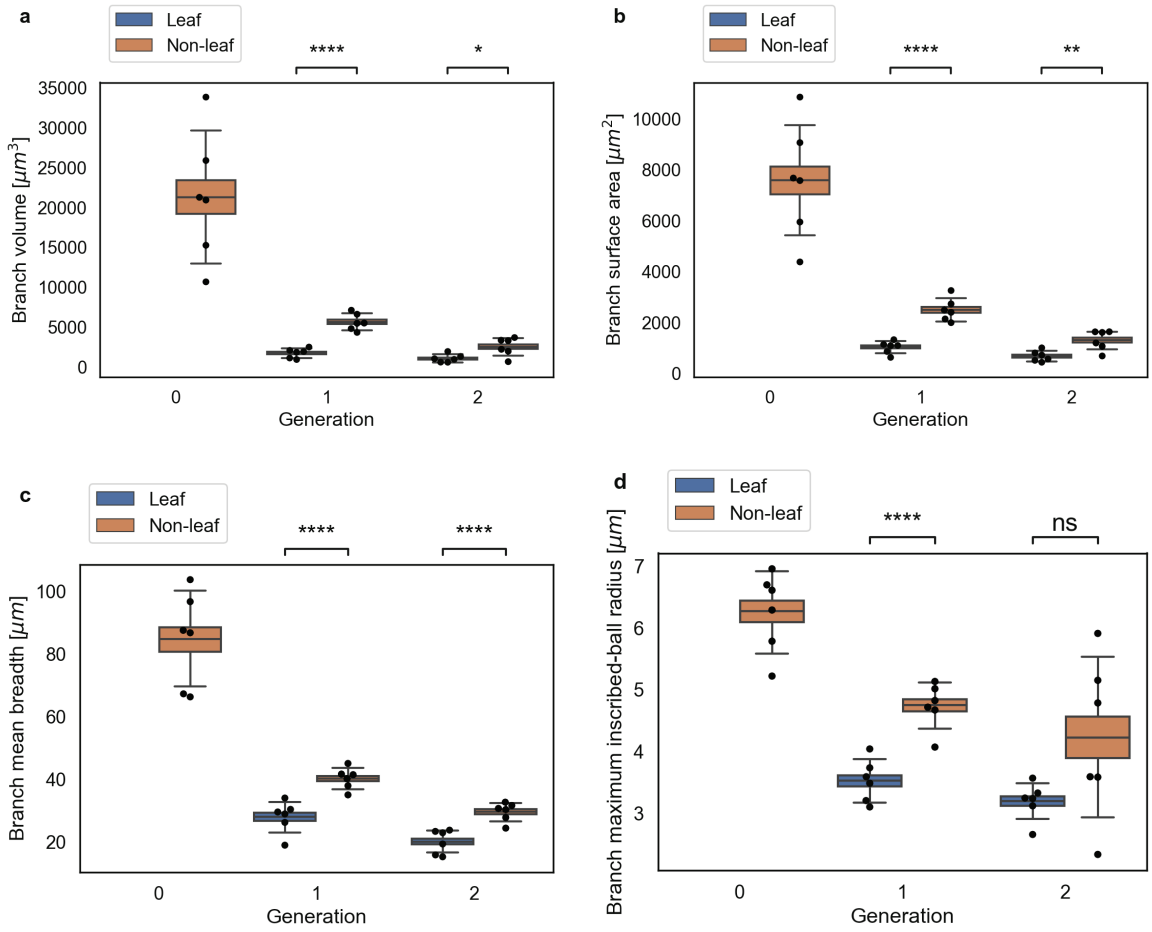


Figure 6.11: Branch volumetric measurements

(a) The crypt volume has a decreasing trend of 21000 ± 3000 µm^3, 5600 ± 400 µm^3 and 2500 ± 400 µm^3 for non-leaf branches with generations 0 to 2, respectively, and 1700 ± 200 µm^3 and 1100 ± 200 µm^3 for leaf branches with generations 1 and 2, respectively. The difference between the volumes of leaf and non-leaf branches of individual generations are statistically significant. (b) The surface area follows a decreasing trend of 7600 ± 800 µm^2, 2500 ± 200 µm^2 and 1300 ± 100 µm^2 for non-leaf branches with generations 0 to 2, respectively, and 1020 ± 90 µm^2 and 670 ± 80 µm^2 for leaf branches with generations 1 and 2, respectively. The difference between the surface areas of leaf and non-leaf branches of individual generations are also statistically significant. Note that in a and b, where the coloured box-plots are too narrow to distinguish the two category colours, the leaf branches (blue) are on the left and the non-leaf branches (orange) are on the right similarly to the rest of the plots. (c) The mean breadth of non-leaf branches also drops with increasing generations (85 ± 6 µm, 40 ± 1 µm and 30 ± 1 µm for generations 0 to 2, respectively) as is the case for the mean breadth of leaf branches (28 ± 2 µm and 20 ± 1 µm for generations 1 and 2, respectively). The difference between the mean breadth of leaf

and non-leaf branches of individual generations are statistically significant. (d) The maximum inscribed-ball radius also decreases slightly with generations (6 ± 0 μm, 5 ± 0 μm, 4 ± 0 μm for non-leaf branches with generations 0 to 2, respectively, and 4 ± 0 μm, 3 ± 0 μm for leaf branches with generations 1 and 2, respectively). The difference between the maximum inscribed-ball radius of leaf and non-leaf branches of generation 1 is statistically significant, unlike the difference between those of generation 2. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). Box-and-whiskers plots appear in pairs for each generation, with those on the left (blue) corresponding to leaf branches and those on the right (orange) corresponding to non-leaf branches. The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

Figure 6.12 shows the morphological characteristics that quantify the several distances within the budding crypt (branch Euclidean length, leaf geodesic length, non-leaf geodesic distance to closest leaf and leaf geodesic distance to root). Figure 6.12a shows that the branch Euclidean length decreases with generation (107 ± 7 μm, 52 ± 3 μm and 32 ± 4 μm for non-leaf branches of generations 0 to 2, respectively, and 61 ± 7 μm, 44 ± 4 μm for leaf branches with generations 1 and 2, respectively). The differences between the Euclidean length of leaf and non-leaf branches of individual generations are not statistically significant. Figure 6.12b shows the corresponding values of the geodesic length for leaf branches of generations 1 and 2 (80 ± 10 μm and 54 ± 5 μm). This shows that the Euclidean length underestimates the actual (geodesic) length of the leaf branches, which is a consequence of their curvature that is also quantitatively evaluated by the straightness discussed later-on. Therefore, the Euclidean length is not an appropriate measure of the branch's length when the native shape of the crypts is preserved, i.e. the curvature of the crypts is retained by avoiding to excise and straighten them as is the case here. Figure 6.12c shows the geodesic distance to the closest leaf measured at the top of the non-leaf branches of different generations. It quantifies the geodesic distance from the closest stem cell niche to the top of a given branch. For non-leaf branches with generation 0, this value (260 ± 10 μm) is the shortest distance that a differentiated cell needs to travel from the crypt base where it was born to the colonic lumen where it eventually dies. For generations 1 and 2 the non-leaf geodesic

distance to closest leaf is equal to 130 ± 10 μm and 90 ± 10 μm, respectively. The difference between generation 0 and the other two generations is statistically significant, whereas the difference between generations 1 and 2 is not statistically significant[7]. The fact that the trunk has twice or more the distance to the closest stem cell niche compared to the non-leaf branches of generations 1 and 2 leads to the following speculation: that, branches of generations 1 and 2 have direct access to newly born cells, enabling them to renew their epithelial layer with younger cells compared to those that end up in the epithelial layer of the trunk, which might die before being shed into the colonic lumen. This speculation might underpin the mechanisms that lead to dead cell deposition in the crypts at early-stage colorectal cancer, as demonstrated for instance by the existence of crypt abscesses [271]. The deposition of dead cells in the trunk could also explain why the trunk is larger compared to the rest of the branches, its irregular shape and the increasing grade of dysplasia during the progression of CRC. This line of thought is also supported by the large leaf geodesic distance to root (270 ± 10 μm and 320 ± 10 μm for leaf branches with generations 1 and 2, respectively) shown in Figure 6.12d, which is the distance a differentiated cell needs to travel before it is shed into the colonic lumen.

[7] Adjusted with Bonferroni correction for multiple (n=2) comparisons. See section 5.1.4.

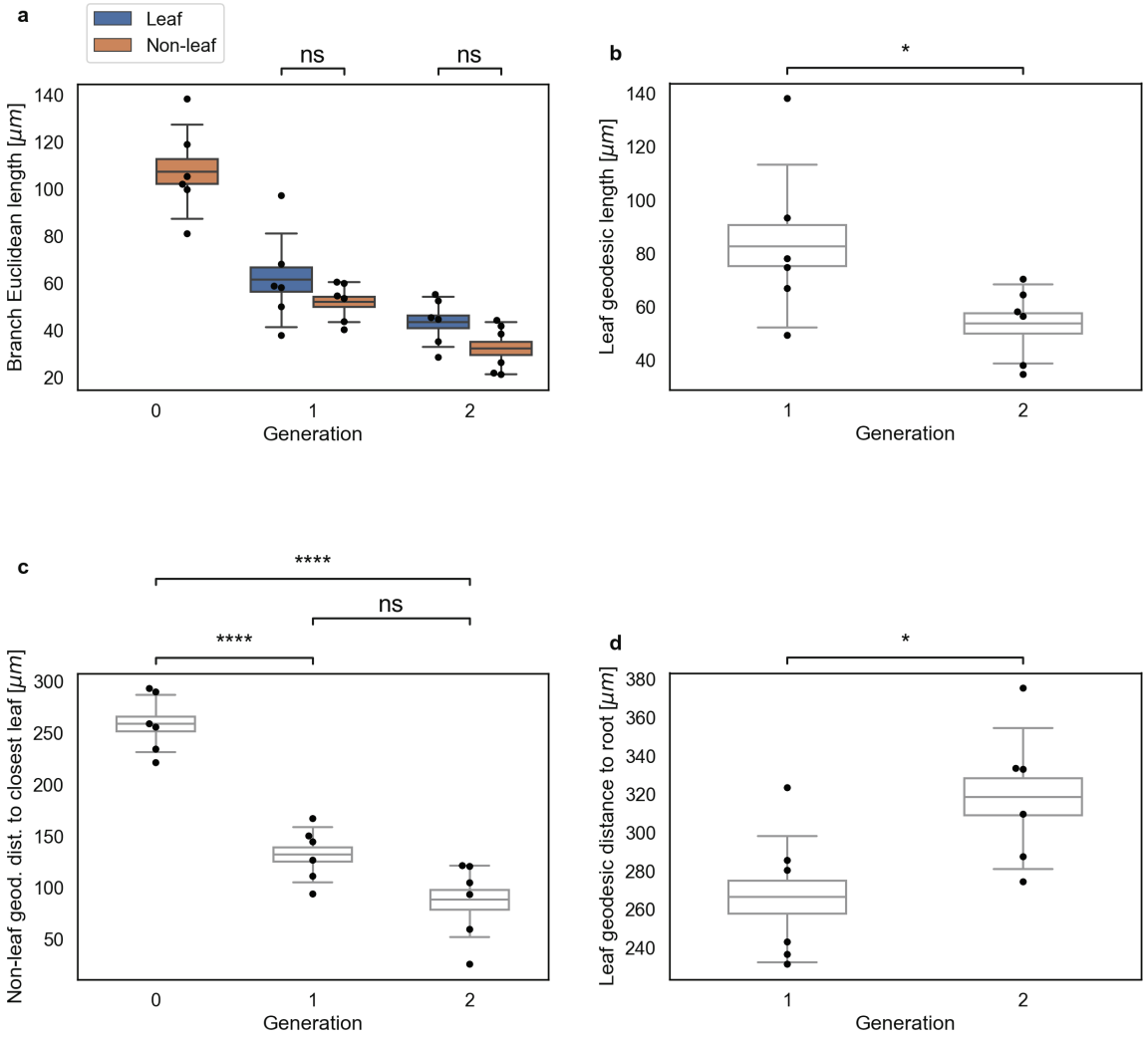


Figure 6.12: Branch distance measurements

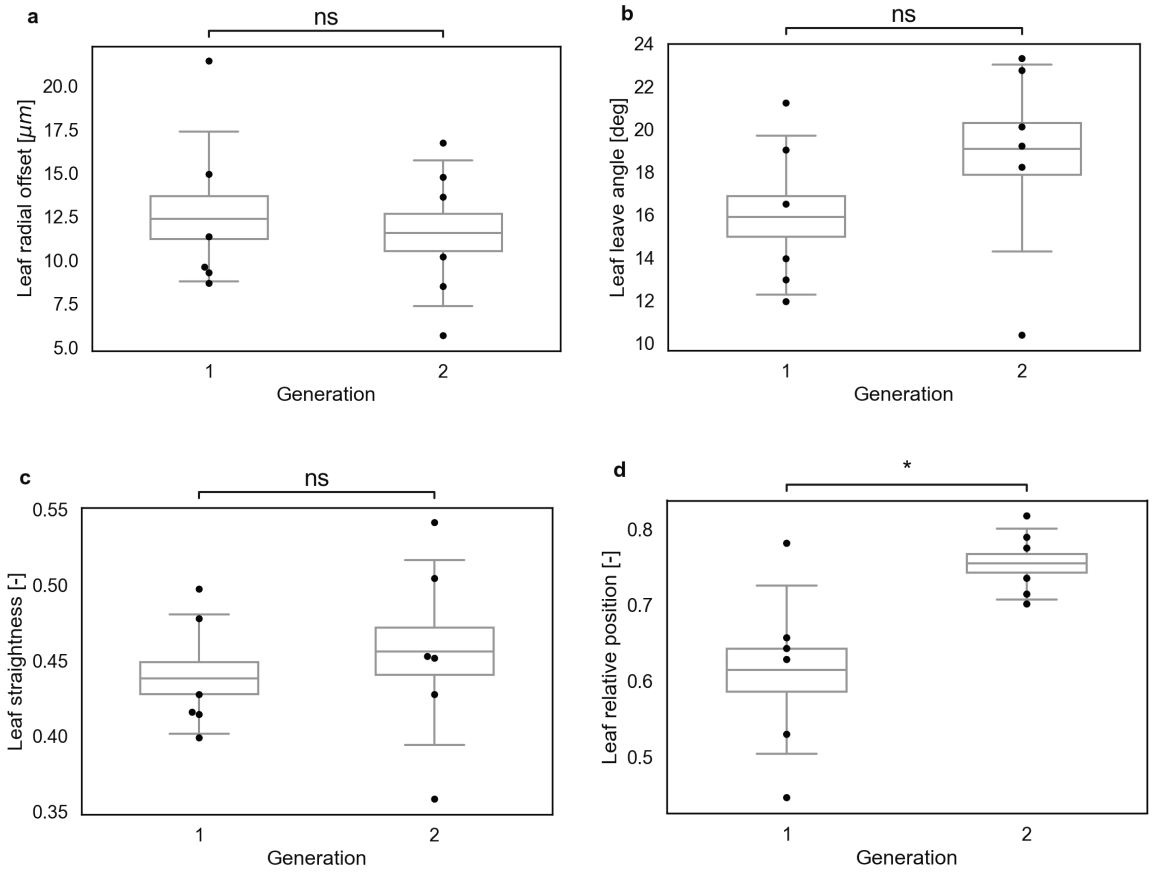
(a) The branch Euclidean length decreases with generation (107 ± 7 µm, 52 ± 3 µm and 32 ± 4 µm for non-leaf branches of generations 0 to 2, respectively, and 61 ± 7 µm, 44 ± 4 µm for leaf branches with generations 1 and 2, respectively). The differences between the Euclidean length of leaf and non-leaf branches of individual generations are not statistically significant. (b) The geodesic length for leaf branches of generations 1 and 2 is equal to 80 ± 10 µm and 54 ± 5 µm, respectively. (c) The geodesic distance to the closest leaf branch is equal to 260 ± 10 µm, 130 ± 10 µm and 90 ± 10 µm for generations 0 to 2, respectively. The difference between generation 0 and the other two generations is statistically significant, whereas the difference between generations 1 and 2 is not statistically significant. (d) The leaf geodesic distance to root is equal to 270 ± 10 µm and 320 ± 10 µm for leaf branches with generations 1 and 2 respectively. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline

corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). In a, box-and-whiskers plots appear in pairs for each generation, with those on the left (blue) corresponding to leaf branches and those on the right (orange) corresponding to non-leaf branches. In b and d the 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$. In c, the critical levels for statistical significance have been adjusted with Bonferroni correction for repeated comparisons (×2), thus the 'star' notation reads: 'ns' for $0.025 < p \leq 0.5$, '****' for $p \leq 0.00005$.

Figure 6.13 shows measures of the branches' orientation and hierarchical measures (leaf radial offset, leaf leave angle, leaf straightness, leave relative position and non-leaf number of children), which are also summarised in the schematic of Figure 6.14. The radial offset of leaf branches is equal to 13 ± 2 μm and 12 ± 2 μm for generations 1 and 2, respectively (difference not statistically significant) as shown in Figure 6.13a. Given that the mean maximum inscribed ball radius was found equal to 3 ± 0 μm for both generations (i.e. leaf branch diameter equal to 6 ± 0 μm), this means that leaf branches extent outwards by a distance roughly twice their cross-section diameter. When the leaf branches are approximated with a straight line connecting their parent and child nodes (see section 5.2.1), this radial offset results in a leave angle of 16 ± 1 deg and 19 ± 2 deg, respectively (difference not statistically significant), shown in Figure 6.13b. This is a simplification since the branches are not straight. Their straightness (Figure 6.13c) is equal to 0.44 ± 0.01 and 0.46 ± 0.02 for leaf branches with generation 1 and 2, respectively (difference not statistically significant), which means that only 45 % of the variability in the branch's mid-line can be explained by a straight line, i.e. there is a bent to the leaf branches. This explains why the Euclidean distance underestimates the geodesic distance, as shown earlier. Figure 6.13d shows the leaf relative position (0.61 ± 0.04 and 0.76 ± 0.02 for generations 1 and 2, respectively), i.e. the branching point of the leaf branches relative to the crypt height. The difference between the leaf relative position of generations 1 and 2 is statistically significant. The fact that the relative position increases with increasing generation, in conjunction with the fact that the leaf branch geodesic length decreases with increasing generation is in agreement with the experimental observation that all leaf branch bases are located at a similar depth within the mucosa (at the interface of the mucosal and muscle layer)[8]. This informs the schematic in Figure 6.14 where the

[8] Observation that was systematically validated during the manual segmentation procedure.

bases of the leaf branches (stem cell niches) are fixed at the bottom of the crypt (relative crypt height equal to 1) and their branching points are set to their respective relative positions at 0.61 and 0.76 for generations 1 and 2, respectively. Finally, the mean number of children (Figure 6.13e) shows that branches of all generations (0 to 2) result in two branches every time they split (bifurcation). One out of the six biological replications has branches of generation 2 with a single child (mean number of children equal to 1). A number of children equal to 1 only emerges as a special case, where two branches merge together creating an 'isle' of tissue (see section 5.2.1). This suggests that the simplified tree structure discussed in this section begins to break from generation 3 onwards into an unstructured form, characteristic of dysplasia. Note, however, that the biological implications related to such rare cases remain to be validated, as they are likely to be a consequence of the limited spatial resolution and, as a result, errors during the segmentation procedure (see '5.1.3 3D segmentation and manual refinement').



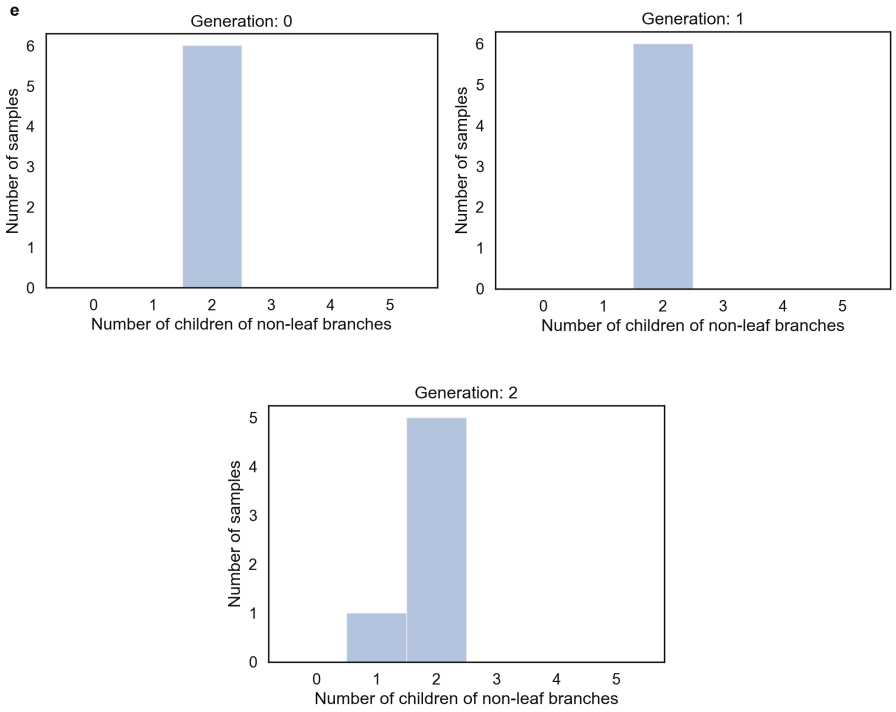


Figure 6.13: Branch orientation and hierarchical measures

(a) The radial offset of leaf branches is equal to 13 ± 2 µm and 12 ± 2 µm for generations 1 and 2, respectively (difference not statistically significant). (b) The leaf branches' leave angle is equal to 16 ± 1 deg and 19 ± 2 deg for generations 1 and 2, respectively (difference not statistically significant). (c) The leaf branches' straightness is equal to 0.44 ± 0.01 and 0.46 ± 0.02 for generation 1 and 2, respectively (difference not statistically significant). (d) The leaf relative position is equal to 0.61 ± 0.04 and 0.76 ± 0.02 for generations 1 and 2, respectively. The difference between the leaf relative position of generations 1 and 2 is statistically significant. (e) The mean number of children shows that branches of all generations (0 to 2) result in two branches every time they split (bifurcation). One out of the six biological replications has branches of generation 2 with a single child (mean number of children equal to 1), which is a special case where two branches merge together creating an 'isle' of tissue. The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

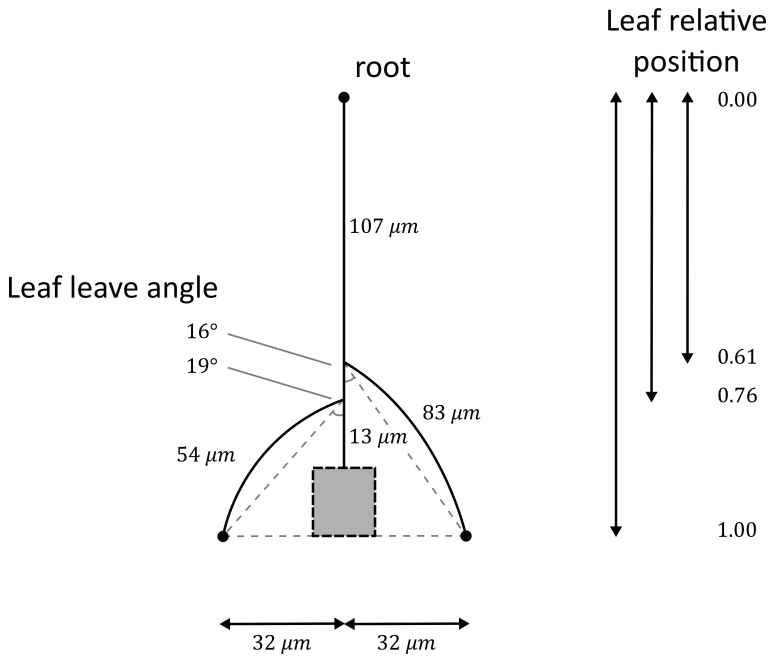


Figure 6.14: Budding crypt structure schematic

An abstract depiction of the budding crypt structure. The typical crypt bifurcates twice. Further branching events where not modelled (grey rectangle) as they were found to be a minority compared to the two main branching structures ('12' and '122'). Non-leaf branches are shown as straight lines with their corresponding Euclidean lengths, whereas leaf branches are shown as curved lines along with their geodesic lengths. The two leaf branches leave their parent branch at an angle and they are not straight. Their base is found at the bottom of the budding crypt at a radial distance away from their branching point that is about twice their diameter. The volumetric and other types of branch measures that are not shown on this schematic are summarised in Table 6.2.

The schematic in Figure 6.14 shows an abstract depiction of the budding crypt structure revealed by the analysis of sections 6.2.1 and 6.2.2, along with key branch measurements. The typical crypt bifurcates twice creating leaf branches with generations 1 and 2. Further branching events producing branches with generations 3 and above where not modelled as they were found to be a minority compared to the two main branching structures ('12' and '122'). The two leaf branches leave their parent branch at an angle, and they are not straight. Their base is found at the bottom of the budding crypt at a radial distance away from their branching point that is about twice their width. Although only Euclidean distances were measured for non-leaf branches, the leaf geodesic distance to root and the geodesic distance to closest leaf were also measured for the leaf branches. The complete description of the typical budding crypt is detailed in Table 6.2.

Table 6.2 Representative budding crypt branch morphological measurements

The measurements of the branch morphological characteristics for the AOM/DSS group at the end of cycle 1 of the exploratory animal study of section 2.1.1. The mean ± standard deviation of each measurement, estimated following the process of section 5.3, is given for leaf (L) and non-leaf (NL) branches with generations (Gen) 0 to 2. The results of statistical comparison with a permutation test are also given in the form of the corresponding p-value. Statistical significance was evaluated as detailed in section 5.3. In brief, mean value differences when tests yield p-values greater than 0.05 were deemed not statistically significant. In the special case of the geodesic distance to the closest leaf (**), where multiple statistical comparisons were performed, the critical values were adjusted with the Bonferroni correction for a repeated number of comparisons n=2. The p-value for the comparison between the geodesic distance to the closest leaf for Gen 0 vs Gen 1 was found equal to 0.000 (significant), for Gen 1 vs Gen 2 was equal to 0.037 (not significant) and for Gen 0 vs Gen 2 was equal to 0.000 (significant).

Measure	Type	Gen 0	Gen 1	Gen 2	Gen 1 L vs NL p-val.	Gen 2 L vs NL p-val.	Gen 1 vs Gen 2 p-val.
2D cross-sectional							
Perimeter (P) [µm]	L NL	N/A 67 ± 5	25 ± 1 51 ± 2	24 ± 1 48 ± 6	0.000	0.004	N/A
Area (A) [µm^2]	L NL	N/A 210 ± 20	49 ± 4 140 ± 10	46 ± 4 130 ± 30	0.000	0.013	N/A
Circularity ($4\pi \frac{A}{P^2}$) [-]	L NL	N/A 0.65 ± 0.02	0.82 ± 0.01 0.67 ± 0.02	0.83 ± 0.01 0.66 ± 0.03	0.000	0.000	N/A
Solidity [-]	L NL	N/A 0.8 ± 0.1	0.82 ± 0.00 0.78 ± 0.02	0.84 ± 0.01 0.76 ± 0.02	0.156	0.000	N/A
Inertia ellipse elongation [-]	L NL	N/A 2.5 ± 0.1	1.62 ± 0.03 2.2 ± 0.1	1.58 ± 0.02 2.0 ± 0.1	0.002	0.004	N/A
Maximum Feret diameter [µm]	L NL	N/A 27 ± 2	10 ± 0 20 ± 1	10 ± 0 18 ± 2	0.000	0.004	N/A
Maximum inscribed circle radius [µm]	L NL	N/A 5 ± 0	3 ± 0 4 ± 0	3 ± 0 4 ± 0	0.000	0.043	N/A
Volumetric							
Volume [µm^3]	L NL	N/A 21000 ± 3000	1700 ± 200 5600 ± 400	1100 ± 200 2500 ± 400	0.000	0.022	N/A

Measure	Type	Gen 0	Gen 1	Gen 2	Gen 1 L vs NL p-val.	Gen 2 L vs NL p-val.	Gen 1 vs Gen 2 p-val.
Surface area [µm²]	L NL	N/A 7600 ± 800	1020 ± 90 2500 ± 200	670 ± 80 1300 ± 100	0.000	0.009	N/A
Mean breadth [µm]	L NL	N/A 85 ± 6	28 ± 2 40 ± 1	20 ± 1 30 ± 1	0.000	0.000	N/A
Max inscribed ball radius [µm]	L NL	N/A 6 ± 0	4 ± 0 5 ± 0	3 ± 0 4 ± 0	0.000	0.009	N/A
Abstraction							
Euclidean length [µm]	L NL	N/A 107 ± 7	61 ± 7 52 ± 3	44 ± 4 32 ± 4	0.331	0.078	N/A
Geodesic distance to closest leaf [µm]	L NL	N/A 260 ± 10	N/A 130 ± 10	N/A 90 ± 10	N/A	N/A	**
Leaf geodesic length [µm]	L NL	N/A N/A	80 ± 10 N/A	54 ± 5 N/A	N/A	N/A	0.037
Leaf geodesic distance to root [µm]	L NL	N/A N/A	270 ± 10 N/A	320 ± 10 N/A	N/A	N/A	0.0346
Relative position [-]	L NL	N/A N/A	0.61 ± 0.04 N/A	0.76 ± 0.02 N/A	N/A	N/A	0.015
Radial offset [µm]	L NL	N/A N/A	13 ± 2 N/A	12 ± 2 N/A	N/A	N/A	0.727
Straightness [-]	L NL	N/A N/A	0.44 ± 0.01 N/A	0.46 ± 0.02 N/A	N/A	N/A	0.578
Number of children	L NL	N/A 2 ± 0	N/A 2 ± 0	N/A 1 to 2	N/A	N/A	N/A
Leaf leave angle [deg]	L NL	N/A N/A	16 ± 1 N/A	19 ± 2 N/A	N/A	N/A	0.223

6.3 Key findings

Chapter 6 outlines the measurements following a new methodology for objective morphological quantification of crypt budding in 3D. This was achieved in incremental steps of granularity: First, the crypts' morphology was quantified globally (section 6.2.1). Next, a more detailed quantification was performed by focusing on individual branches (section6.2.2).

The developed digital volume unrolling method of 3D cycloramas has a central role in this process: The unrolled colons were used to identify crypts, which were subsequently segmented in 3D. Additionally, the colonic layer thickness and cross-section areas were computed using the 3D

cyclorama method. These results are reported in Appendix B: 'Global characterisation of the colonic structure'. Branch-wise quantification became feasible by introducing the topology graph. This mathematical tool allows to link the morphological measurements of individual branches to their specific location within the branching structure (budding crypt) and describe their topological inter-relationships. Morphological quantification of individual branches advances the currently employed practices by building a complete 3D picture of the budding structure of the crypts. This facilitates an improved understanding of the underlying biological mechanisms that drive crypt budding, and provides detailed quantitative measurements that could better inform computational model designs.

The analysis performed on tissue samples from the exploratory animal study detailed in Chapter 2: 'Application of an animal model for colorectal cancer' revealed insights that might be worth investigating further: Non-destructive 3D imaging of the colonic structure was a crucial facilitator of the developed methodology. It enabled to preserve the native colonic shape and avoid introducing structural deformations to the tissue, and consequently the colonic crypts. This serves two purposes, firstly the crypt shape remains unchanged (except any deformations that might be induced due to chemical fixation, namely shrinkage [109, 272]), but it also retains the spatial location and orientation of the crypts within the tissue. An unchanged crypt shape facilitates accurate 3D quantification of volumetric and shape measurements such as the branch volume, surface area, cross-section circularity, straightness, etc. Preserving the spatial location and orientation was useful when crypts needed to be traced back to their location within the tissue to explain an observed clustering of short and tall crypts. A new insight emerged from this analysis that suggests that the crypt height is directly related/determined by its location within the tissue (crypts at mucosal folds are shorter than others, but they are otherwise not dissimilar).

Branch-wise quantification using the topology graph enabled a detailed description of the budding crypt morphology. The developed systematic approach provides the tools to build a complete 3D description of the budding crypts during the early-stages of colorectal cancer for the first time. The quantified phenotype yield outcomes that are supported by the current understanding of CRC development, and it also revealed insights that might lead to new questions and further studies: the structure descriptor defined here emanates directly from the topology graph. It conveniently represents the crypt structure in a single stream of numbers, facilitating comparison among structurally dissimilar crypts. The structure descriptor confirmed the current understanding of crypt budding as a bifurcation process [14, 113] (two child branches are spawned at every budding event), and therefore the attribution of abnormal budding to an impaired fissioning process [13]. This reinforces the evidence that supports crypt budding as a precursor of ACF, the earliest biomarkers of CRC [17-19]. The structure descriptor also revealed

two predominant budding structures, ‘12’ and ‘122’. The first corresponds to a single bifurcation, whereas the second corresponds to two consecutive bifurcations. A third group of budding structures named ‘other’ was defined to collect the rest of the budding structures, which correspond to exceedingly deformed crypts. Such budding structures were also observed by others [6, 9, 22], but their morphology was not quantitatively evaluated to a high detail. Such findings and quantitative measurements are invaluable to theorists that build computational models of crypt budding, where experimental data needed for validation are often scarce [6, 27].

The topology graph enabled the differentiation between leaf branches, which are assumed to have a stem cell niche at their base, and non-leaf branches. When they indeed have a stem cell niche at their base, the leaf branches are elsewhere referred to as daughter crypts [6, 22]. Here this nomenclature was purposefully avoided because the existence of stem cell niches was not validated at the time of the analysis. Nevertheless, differentiation between leaf and non-leaf branches revealed that the crypts tend to branch on the lower half of the crypt (towards the base) with the highest branch being located close to the middle. This outcome is tightly related to the ‘APC:Wnt counter-like mechanism’ proposed by Boman *et al.* [24]. They postulated that gradients of factors that regulate cell proliferation along the crypt axis (APC, Wnt) may, in turn, determine the most likely location of branching or ‘sweet spot’ [24, 39]. Therefore, the approach developed here that enables quantification of the branch relative position, directly determines the sweet spot (branching point) and could therefore in principle infer the relative concentrations of the APC vs Wnt factors according to the arguments of Boman *et al.* [24].

The topology graph, along with the differentiation between leaf and non-leaf branches, also enabled the extension of the asymmetry index [6] and the symmetry ratio [22] to include more than two leaf branches. Quantification of the crypt asymmetry using this (extended) asymmetry index and the symmetry ratio [22] revealed that ‘12’ crypts tend to be close to perfectly symmetric and more symmetric than ‘122’. This led to the conclusion that ‘12’ crypts are probably undergoing the normal fissioning process, whereas ‘122’ represent the diseased budding case. This is in line with the hypothesised aberrant crypt budding concept of Tan *et al.* [6] where a “parental crypt starts budding before the original daughter crypt has finished its crypt production sequence”. The aberrant crypt budding concept was also supported by a very high correlation between the leaf-to-root distance and the crypt height, which suggests that all leaf branches have their base (stem cell niche) on the mucosal/muscle interface.

In terms of the quantification of symmetry, it was experimentally found that the symmetry ratio achieves better differentiation of budding crypts than the asymmetry index. This was attributed to the asymmetry index being more conservative as it averages out the asymmetries between the

branches, rather than punishing the most extreme among the asymmetries. Therefore, the symmetry ratio was deemed a better measure of asymmetry, especially for crypts with more than two leaf branches.

Several interesting observations were made regarding the shape of the branches. Firstly, the maximum inscribed circle radius, i.e. the widest part of the lumen, was constant across all branches and independent of their generation. At the same time, the cross-section shape of non-leaf branches was too complex to be represented by an ellipse. On the other hand, leaf branches had an elliptical cross-section shape profile. These observations draw a picture of non-leaf branches with a challenging shape, and a fluctuating width (as quantified by the maximum inscribed circle radius); whereas leaf branches have a simpler and more consistent shape. Moreover, non-leaf branches were found to be larger compared to leaf branches of the same generation. This led to the speculation that migrating cell flows originating in stem cell niches are added up at each branching point requiring branches with larger cross-section perimeters (or equivalently larger surface areas) at decreasing generations. This speculation was also supported by measurements of the branches' volume, surface area and mean breadth that revealed an overall size of individual branches that decreases with generation.

Further to simplified thoughts and speculation regarding the budding crypt function, the measurements above and the detailed phenotype put together in section 6.2 build a complete picture of the budding crypt structure, and pave the way for improved computational models. What remains is firstly for the above-mentioned observations to be verified and estimated better in controlled experiments with large numbers of animals. Next, their biological relevance needs to be confirmed by linking the crypt structure with concentrations of proteins, cells of specific types, etc., relating to the particular study at hand. As mentioned earlier, the methodology described in Chapter 2: 'Application of an animal model for colorectal cancer' is suited for morphological quantification. However, the workflows and tools developed in Chapter 3 to 5, i.e. preparation and non-destructive imaging and structural quantification of FFPE tissue samples, enable further functional studies to be performed. Specifically, the FFPE tissue samples may be processed via histological/histochemical workflows to obtain functional information at selected sites within the 3D volume.

6.4 Limitations and outlook

Purpose of this chapter was to showcase the utility of the newly developed tools for objective morphological quantification of colonic crypts, in an attempt to support the development of computational models of crypt budding at the early-stages of CRC. To this end, colonic tissues

from an exploratory animal study were used (see section '2.1 Animal model materials and methods'). A basic animal model was used for its simplicity and reproducibility, but that was not customised to test a particular hypothesis. Therefore, the pathology in the produced samples depicts the broad background of the disease, and it does not serve for studying a specific biological question. As a result, the phenotype quantified in section 6.2 should only be interpreted as a rough description of the disease background. The validity of the measurements is also limited by the small number of animals used (6× per group). In this account, advanced statistical analysis (bootstrapping and permutation tests) had to be employed to improve the measure's estimations. Still, considerable variability in the final statistical distributions made it challenging to obtain statistically significant differences in some cases. Thus, the number of animals would need to be revised in future studies and adapted to the particular study at hand.

This is also true for the selection of sampling sites. In this study, samples from random colonic sites (proximal, middle and distal) had been grouped together. However, it is well known that the morphology of both the colonic tissue as a whole, as well as the morphology of the colonic crypts, are site-dependent [6, 9, 22]. Other studies might be interested in the morphological changes over time, as it is generally the case for time-dependent studies of other related factors (e.g. epithelial stem cell concentrations [52], number of ACF [86], bifurcation frequency [11], etc.). To this end, identification of appropriate culling time-points depending on the animal model at hand would be necessary. As far as the animal study applied here is concerned, qualitative observations revealed excessive morphological changes at cycle 2, where identification of budding crypts as individual entities was almost impossible (the mucosal structure was practically lost). At cycle 3 the mucosal structure was lost entirely, with gross tumours and generalised dysplasia being the dominant characteristics. Therefore, for studies of the earliest stages of colorectal cancer that focus on morphological changes of the budding colonic crypts, the protocols would need to be adjusted such that more culling points are introduced close to, and around cycle 1.

From the technical point of view, as is the case with any imaging modality, the quantification accuracy is limited by the available spatial resolution. The datasets used in section 6.2 were imaged at a pixel size of 2.2 µm, which is appropriate to depict the gross structure of a crypt but too large to depict cellular information. Although a better pixel size of 0.33 µm is attainable, which is comparable to the pixel size of a high-resolution histology image (e.g. Olympus VS110 at 20x optical magnification yields a pixel size of 0.32 µm [234]), this is accompanied with a limited field of view. This is a consequence of the fact that a fixed number of pixels on the detector is set up with a smaller pixel size, thus capturing a smaller FOV. Although, one could perform several high-resolution scans and subsequently stitch them together to obtain an extended field of view, technical limitations, including computer memory, bandwidth, and computational resources,

ultimately impose a trade-off between the selected pixel size and the resulting field of view. In this case, the pixel size was selected such that the entire colon (4 cm in diameter) would fit in the field of view and obtain an overall image of the sample. In order to achieve better spatial resolution, by selecting a pixel size of 0.33 µm, this would require that a much smaller volume is imaged (7^3 times smaller) that would not depict the entire colon but a small sub-volume.

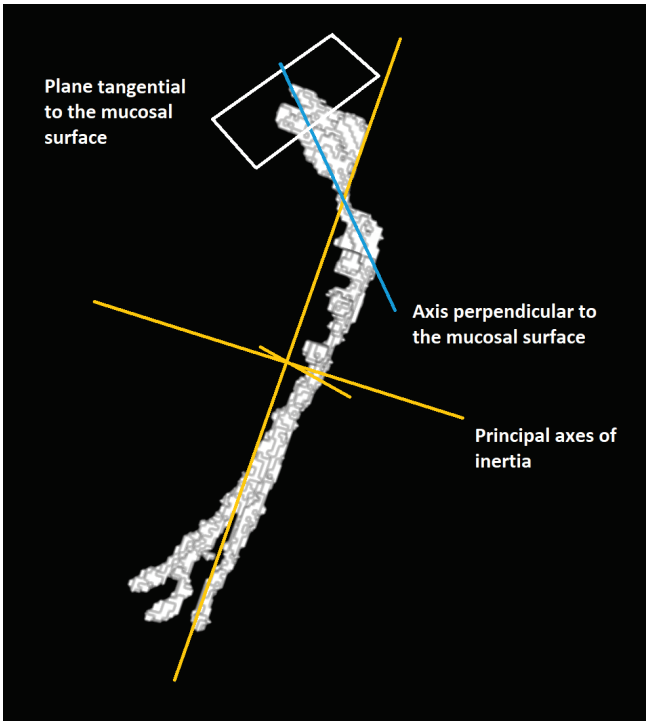


Figure 6.15: Sub-optimal alignment case

The alignment using the moments of inertia is sub-optimal in cases where the crypt branches are exceedingly curvy (small straightness), extending heavily to one side. This results in a principal axis of inertia (long yellow line) that deviates from the central axis of the trunk branch, resulting in an aligned crypt that has it trunk at an angle with respect to the axis perpendicular to the mucosal surface.

Imaging the entire colonic diameter allowed to depict several thousands of crypts in a single scan, which after 3D segmentation could be traced back to their original location in the tissue. This allowed the identification and isolation of individual crypts into separate image stacks. Those were aligned automatically by computing their moments of inertia, resulting in consistently aligned crypts without any user input. However, the alignment is sub-optimal when the crypt branches are exceedingly curvy (small straightness), extending heavily to one side. This results in a principal axis of inertia that deviates from the trunk's central axis (see Figure 6.15), resulting in an aligned crypt with its trunk at an angle with respect to the axis perpendicular to the mucosal surface. This alignment tilt would affect the cross-sectional morphological measurements to some

extent. Therefore, this should be considered when the cross-sectional morphological measures of exceedingly abnormal crypts (many branches, with low straightness) are compared.

A final remark is worth to be made regarding the applicability of the 3D quantification method to other structures further from the colonic crypts: The methodology detailed in Chapter 5: 'Tools and workflows for 3D morphological quantification of budding structures' is in principle applicable to branching networks with at least one entrance (root node) and at least one exit (leaf node). Such network structures are found in several biological tissues such as airways in the lung [273], blood vessel networks [274], and neural networks [275]. Given a binary segmentation of the structure of interest, the developed software will automatically isolate, align, create the topology graph and quantify all described measures. These may then subsequently be analysed and interrogated regarding their biological significance depending on the structure they represent.

Chapter 7 Summary, key findings and outputs

The work presented here revolved around a central concept of developing the engineering tools necessary for 3D imaging and analysis, for the morphological quantification of budding structures. The motivation for this main objective was the study of early-stage colorectal cancer (CRC), where crypt budding is one of the first morphological alterations observed in the colon.

Several factors were identified as possible obstacles in the study of 3D morphology and pathology of budding crypts in section '1.1.2 The need for new tools in the study of crypt budding'. These focused on the drawbacks of existing methods, which include the fact that they are typically not suited for concurrent functional and morphological studies of a large number of crypts. It was also underlined that existing methods that include large numbers of crypts are not well suited for *in situ* studies of the 3D morphology of crypt budding. A lack of a systematic approach for morphological quantification of budding structures in 3D was also identified, with the preeminent work in this direction being restricted in basic 2D measures (see Tan *et al.* [6]). To date, the role of the 3D morphology of budding crypts in CRC has not been (extensively) studied, not even from the computational modelling point of view. In addition to the exceedingly challenging nature of 3D morphological modelling of biological structures, an underlying theme in the literature was the lack of experimental data for model validation and further development. To this end, with CRC as a principal motivating application, the necessary methods, tools and workflows were developed to non-destructively image murine colons and systematically quantify crypt budding in 3D.

Murine colon samples were induced with early-stage CRC in Chapter 2: 'Application of an animal model for colorectal cancer'. The sample preparation procedure was designed such that the resulting colons retain their natural tubular shape, and to avoid introducing morphological deformations to the budding crypts. The colons were embedded in paraffin wax so that well-established histological methods (e.g. H&E histology, histochemistry) can be readily applied after non-destructive 3D imaging, within a newly developed workflow to study the crypt budding morphology. Chapter 2 fulfils specific aim i. in section '1.5.2 Specific aims': *Application of an animal model for CRC: To employ a well-established animal model in order to produce FFPE animal colons to be subsequently imaged and quantified. The model needs to recapitulate aspects of CRC so that crypt budding can be observed.*

X-ray propagation-based phase-contrast imaging was identified as an appropriate non-destructive 3D imaging modality to capture the colonic microstructure. In Chapter 3: 'Soft tissue X-ray µCT imaging', the necessary technical requirements were collected, and the imaging protocol design process was simplified through a newly developed flowchart. The flowchart enables the user to

derive an optimal X-ray propagation-based phase-contrast imaging protocol for synchrotron and laboratory sources that enforces these requirements and yields their desired spatial resolution of the final output, i.e. the phase-retrieved CT reconstruction. This flowchart yields an optimal imaging protocol, irrespective of the material of the sample being imaged. Therefore, it applies in principle to any kind of sample, and in particular those that do not significantly attenuate the X-ray beam, where phase-contrast imaging is particularly beneficial. Chapter 3 fulfils specific aim ii. in section '1.5.2 Specific aims': *X-ray phase-contrast µCT imaging of the budding crypt microstructure: To develop a method that enables X-ray phase-contrast µCT imaging of murine colon samples that is easy to apply by non-experts in physics or engineering.*

Once the cylindrical FFPE samples were imaged in 3D, their tubular structure was visualised in the form of image stacks of tomographic reconstructions. A computational tool (3D cyclorama) was thus developed to digitally unroll the geometrically challenging natural shape of the colons in order to facilitate the identification of budding crypts. The tool samples the 3D data of each colon as a series of onion-like internal surfaces at different depth levels. This enables quick counting of large numbers of crypts (a few thousands) and quick and effective tracing through the onion-like layers in search of branching points of individual crypts within the colonic mucosa. Since its development, the 3D cyclorama software has found applications in several problems, including flattening a jet-engine part with dual curvature, unrolling the microvillus membrane of a placental gland, unrolling an ancient lead scroll and flattening a pharmaceutical film (see Appendix A.2). It was thus decided that it is made available to the public as a plugin within Fiji/ImageJ: https://imagej.github.io/plugins/3d-cyclorama so that the broader scientific and industrial communities can benefit from this capability. Chapter 4 fulfils specific aim iii. in section '1.5.2 Specific aims': *Digital colon unrolling to aid budding crypt identification: To develop a method that facilitates the identification of budding crypts within the µCT scans in an efficient and easy way.*

In Chapter 5: 'Tools and workflows for 3D morphological quantification of budding structures', the identified crypts were isolated by means of a developed workflow that used seed-points defined on the 3D cycloramas to segment the crypt lumens from the 3D µCT data. This provided binary segmentations of individual crypts, which were used as input data for 3D morphological quantification. A systematic approach to identify and quantify individual branches within the branching crypts was developed, which enables the budding crypt morphology to be quantified with unprecedented granularity. More specifically, a topology graph (mathematical structure that captures relationships between individual branches) enabled the representation of the crypt structure so that measurements on individual branches could be allocated to their corresponding location within the budding crypt. Through the topology graph, it was possible to define the

structure descriptor (a single stream of numbers that represents the crypt structure), which in turn enabled the comparison between dissimilar budding crypts.

These mathematical tools and workflows were subsequently employed in Chapter 6: 'Quantitative morphometry of crypt budding in a mouse model' to build a complete phenotype of the 3D structure of budding crypts during the early developmental stages of CRC. Although based on the exploratory animal study with a limited number of animals (×6 per treatment group), the quantified phenotype supports certain aspects of the current understanding of the CRC development. These include the belief that crypt budding originates in impaired fissioning, i.e. it is a bifurcation process, and also the fact that terminal branches constitute colonic crypts, i.e. stem cell niches are located at their bases. Furthering the currently known facts about the budding crypt morphology, the derived phenotype suggests that the mean crypt bifurcates twice before the mucosa goes into a state known as dysplasia. The main aspects of the budding crypt phenotype are summarised in Figure 6.14, which shows the specific crypt topology and branch lengths for the average budding crypt. A few examples have also been mentioned where the derived phenotype could be used to confirm or further the works of others. These include the computation of the 'sweet spot' for branching proposed by Boman *et al.*[24], and extension of the asymmetry index [6] and the symmetry ratio [22] in 3D. These revealed that the derived phenotype is in line with the hypothesised aberrant crypt budding concept of Tan *et al.* [6] where a "parental crypt starts budding before the original daughter crypt has finished its crypt production sequence". These findings and the highly detailed phenotype with quantitative measurements of individual branches are invaluable to theorists that build computational models of crypt budding, where experimental data needed for validation are often scarce [6, 27]. Chapter 5 and Chapter 6 combined fulfil specific aim iv. in section '1.5.2 Specific aims': *Morphological quantification of crypt budding in 3D: To develop a method that enables objective quantification of the structure of identified budding crypts.*

Table 7.1: Key findings and outputs

Chapter	Key findings & outputs
Chapter 1	• Lack of appropriate tools for 3D morphological quantification of crypt budding. • X-ray phase-contrast μCT imaging followed by conventional histology could be a good candidate.
Chapter 2	• Animal model for CRC and sample preparation protocol that preserves the colon's natural shape.
Chapter 3	• A flowchart that enables lay users to design propagation-based in-line phase-contrast imaging protocols yielding optimal spatial resolution.
Chapter 4	• 3D cyclorama method for digital unrolling of deformed tubes (intact colons). • International collaboration for applications in the field of archaeology. • Fiji/ImageJ plugin.
Chapter 5	• A workflow to isolate colonic crypts from the μCT data using seed points defined on 3D cycloramas. • Mathematical tools (topology graph) for systematic morphological quantification of branching structures (colonic crypts) in 3D.
Chapter 6	• Detailed morphological phenotyping of the colonic crypts in the animal model of Chapter 2 using the tools and workflows of Chapter 5.

Chapter 8 Discussion

Advanced 3D imaging techniques, including X-ray µCT, are gaining ground in sciences as they provide in-depth insight and data-rich representations of the sample under investigation. Such experimental outputs are only conducive to research when useful information can be extracted out of the recorded data. When it comes to structural analysis, morphological quantification quickly becomes exceedingly difficult with challenging geometries of features within the sample. This problem extends to a wide range of applications, including extraction of text from curved/rolled documents (e.g. text on books or wrinkled paper [235-240], ancient scrolls [241-244]), and extraction of cracks within materials (e.g. cylindrical batteries under failure [246, 247], and industrial parts examined for fatigue life faults [5, 276, 277]).

In the biomedical field, geometries of features within human and animal tissues are typically even more complex. Examples range from vast bone cell networks formed by osteocytes [248, 249] that are involved in disease including osteoporosis and osteoarthritis, to convoluted networks of blood vessels and airways in the lung [272, 278]. Although progress related to the above-mentioned examples has been achieved using X-ray µCT, other cases that could benefit from this advanced 3D imaging technology have not been confronted yet. A prime example is the study of the budding crypts within the colon during early-stage colorectal cancer.

Chapter 1 discusses the need for 3D imaging of the colonic crypts for studies that include morphological quantification and conventional histology/immunohistochemistry. The work presented here has been motivated by this demand for new tools and workflows to enable the study of crypt budding in 3D. Yet, the developed approaches are not confined to CRC, nor to biological structures. Beginning from Chapter 3 ‘Soft tissue X-ray µCT imaging’ the proposed optimal protocol design flowchart aims to bridge a gap between X-ray µCT technology developers and users with non-technical backgrounds. The approach for designing a protocol that yields optimal image resolution is applicable to low-absorbing samples, including soft tissues, plastic materials, etc. Next, Chapter 4 ‘Colonic 3D cyclorama: a method for digital volume unrolling’ deals with simplifying challenging geometries of samples imaged with any kind of 3D imaging modality, including confocal laser scanning microscopy, electron tomography and light sheet microscopy, with applications spanning an extensive range of fields, including archaeology, the manufacturing industry and biomedical research. This approach enables unrolling of tubular structures such as colons and industrial pipes, but also flattening sheets, such as pharmaceutical patches and ancient clay tablets. Finally, Chapter 5 ‘Tools and workflows for 3D morphological quantification of budding structures’ employs abstract mathematical tools such as a topology graph that facilitate

the quantification of branching structures, including colonic crypts but also applicable to blood vessel and airway networks, plant root networks, etc.

Special care has been taken so that the design, nomenclature and implementation of the developed tools is generic enough so that it can be applied to a range of different examples, including those discussed above. Nonetheless, these tools and workflows were developed and showcased using an animal model for CRC (see Chapter 2), where crypt budding plays a key role in understanding the disease. Notably, though, a few considerations are needed before the presented workflows can be fully adopted in CRC research.

8.1 X-ray phase-contrast imaging from the synchrotron into the lab

The quantified samples have been imaged at a synchrotron source. Such facilities provide highly brilliant X-rays (high X-ray flux and spatial coherence), suitable for phase-contrast imaging at extremely high throughput. To put this in perspective, each scan at DLS took about 3.5 minutes, whereas a similar scan in the lab would take several hours. This high throughput is pivotal in order to conduct studies with many samples and enable statistical hypothesis testing to study particular aspects of a disease.

However, availability to synchrotron sources is limited, and beamtime is granted through a competitive peer-reviewed proposal process. Admittedly, it would be more convenient if this modality would be available to researchers on demand through laboratory sources. Although this is in principle not impossible, it is exceedingly difficult to achieve on commercial systems. When applying the flowchart of Figure 3.4 to design an optimal phase-contrast imaging protocol on a ZEISS Xradia Versa system for a desired feature size of $\alpha_{p,des} = 10\ \mu m$ and an effective energy of 20.0 keV (~60 keVp), it yield a magnification objective $M_o = 20$, no binning $M_b = 1$, a source-to-object distance $R_1 = 30\ mm$ and an object-to-detector distance $R_2 = 60\ mm$. Yet, this protocol was impossible to apply experimentally due to extremely low photon-count on the detector (limited X-ray flux). Gradually relaxing the protocol by reducing the source-to-object distance to $R_1 = 14\ mm$ and an object-to-detector distance $R_2 = 20\ mm$ and setting a high exposure time of 20 s per projection for 3201 projections, resulted in a protocol capable of imaging the FFPE murine colon in the lab (see Figure 8.1). The high-resolution µCT scan (pixel size = 0.33 µm) revealed the tissue microstructure down to individual epithelial cells. The colonic crypts are readily visible, and the crypt epithelium is distinguished by its boundary. This scan took, however, a total of 18.5 hours, which is impractical for several reasons. Not only it does not allow for high-throughput imaging of many samples, but also access to laboratory imagers is typically costly. As discussed in section ‘8.2 On the feasibility of using X-ray CT for the study of colorectal cancer’, it

remains to be elucidated how X-ray exposure affects the several proteins and other elements in the tissue that might be relevant to the study of CRC.

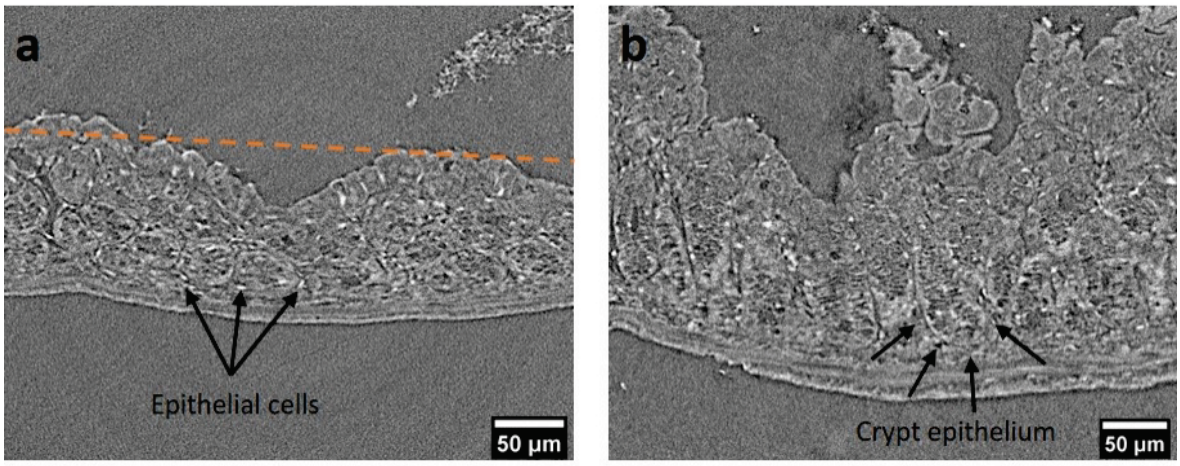


Figure 8.1: High-resolution imaging of a FFPE murine colon in the lab

Low-absorbing edge-enhanced (see Figure 1.27) X-ray µCT scan obtained using ZEISS Xradia Versa system. Laboratory X-ray µCT of soft tissue embedded in paraffin wax is possible but impractical (scan time ~ 18.5 hours). (a) A top cross-section of a healthy murine colon reveals the tissue microstructure down to individual epithelial cells. Colonic crypts are not clearly visible in this view because the cross-section does not slice them longitudinally. (b) A cross-section of the µCT volume along the dashed line in a better reveals the colonic crypts. The crypt epithelium is directly distinguished by its boundary and the colonic microstructure is portrayed in great detail.

Albeit the currently faced difficulties related to achieving phase-contrast X-ray µCT in the lab, recent technical and theoretical advances suggest a promising future for this technology. Improvements in both X-ray flux and spatial coherence can be achieved by the fast-developing liquid-metal-jet type of X-ray sources [279]. This technology uses a jet of melted metal as an anode, upon which electrons are projected to create X-rays. The superior thermal properties of such sources (high-speed metal flow enables rapid cooling), allow for more than 100× brighter sources compared to standard reflection source types [174]. However, a typical limitation of liquid-metal-jet sources is that high X-ray beam power results in vaporisation of the liquid metal [280]. As the X-ray beam power and the X-ray energy are tightly related to each other, increasing the X-ray energy in order to reduce absorption by the tissue is particularly challenging. Pushing the limits of the liquid-metal-jet source technology, recent improvements in the achievable X-ray energies already supported the first high-resolution imaging of mammography and angiography phantoms [280], paving the way for the use of liquid-metal-jet sources for biomedical applications.

A better X-ray beam alone is not enough to achieve phase-contrast imaging, though. Careful selection of the experimental settings is also necessary, with the balance between the Fresnel fringe width and the detector pixel size being crucial, as discussed extensively in Chapter 3 'Soft tissue X-ray µCT imaging'. Considering limitations imposed by the X-ray beam and also the resulting image quality (typically quantified via the signal-to-noise ratio) poses a much more difficult optimisation problem, which does not have a unique solution [281]. Thus, phase-contrast imaging in the lab will unavoidably require the user to carefully select the experimental settings for the sample at hand. Theoretical approaches that model such systems [281-283], as well as simplified design algorithms such as those developed in section '3.1 Soft tissue imaging in practice' assist in protocol design, bringing laboratory phase-contrast imaging one step closer to practical realisation.

8.2 On the feasibility of using X-ray CT for the study of colorectal cancer

Suggesting that non-destructive X-ray µCT imaging may be used prior to conventional histology/immunohistochemistry for the study of CRC mutely implies using a correlative approach involving compatible technologies. A brief proof-of-concept for such an approach was presented in section '2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining', where a sample was stained with H&E after imaging with X-ray µCT to prove that the two modalities are not mutually exclusive. However, H&E staining does not cover the entire range of histological assays that might be relevant to CRC research. As discussed in detail in sections '1.2 Colorectal cancer and crypt budding' and '1.3 Standard approaches for crypt budding research', a wide range of factors might be of interest, including the concentration and distribution of stem, $Muc2^+$ and $cKit^+$ cells and levels of proteins such as β-catenin, E-cadherin and survivin throughout the crypt epithelium. These can be effectively visualised via conventional immunohistochemistry techniques but the capability of doing so after X-ray irradiation, to ensure that these components are not adversely affected by the X-rays, remains to be validated.

Promising steps in this direction have been recently taken by Zeller-Plumhoff *et al.* [284, 285], who employed synchrotron µCT followed by immunohistochemistry to study the oxygenation of murine skeletal muscle. Jones *et al.* [286] applied a similar approach to study fibrotic plaques in human lung samples stained by *Movat's pentachrome* after µCT imaging. Soft-tissue optimised µCT imaging in the lab followed by immunohistochemistry was employed by Robinson *et al.* [287] to assess the lymphatic function of human FFPE lung tissue. Taken together, these recent advances in the emerging field of 3D X-ray histology [288] and the relevant correlative approaches suggest that a similar correlative workflow to study crypt budding should be feasible, with no significant impediments in foresight. As the sample preparation workflow outlined in sections

‘2.1.2 Tissue sample acquisition and preparation’ and ‘2.1.3 Paraffin sectioning and haematoxylin & eosin histology staining’ results in FFPE samples suitable for X-ray µCT imaging followed by paraffin sectioning and staining, correlative approaches that involve both modalities are in principle directly applicable.

As discussed in section ‘1.3 Standard approaches for crypt budding research’, specific research questions might require either studying the tissue as a whole (global), or focusing on a small volume of interest (local). Although destructive, available methods for local imaging (e.g. microdissection, scanning electron microscopy, confocal imaging) are well-suited for the job at hand. On the other hand, global methods such as isolating the crypts via mucosal digestion are exceedingly invasive. As a result, the crypt’s initial orientation and location within the tissue is lost entirely by the time the crypts can be imaged. Tan *et al.* explicitly state that crypts from young mice are too fragile to be processed with mucosal digestion [6], thus limiting its applicability in certain cases. In contrast, the adopted process of preparing the colons in the form of FFPE sausages that preserves their natural shape is gentler to the tissue. Embracing the sample preparation method proposed here, which results in cylindrical wax blocks of 4 mm in diameter was only possible due to the development of two essential methods: i. µCT imaging for the whole tissue in the FoV, and ii. digital processing of the resulting 3D data.

i. X-ray µCT is gaining ground for imaging of soft tissues [169, 289-291]. As conventional attenuation-contrast X-ray imaging is almost ineffective on soft tissues, which barely absorb X-rays, advanced approaches need to be employed. A method to enhance X-ray absorption is by chemically changing the material properties of the tissue. This may be achieved by increasing its electron density with a diffusible staining agent, which eventually yields high-contrast images [292, 293]. Today, iodine-based staining solutions are among the most widespread X-ray contrast agents [109] due to their low toxicity and ease of use [293]. They can be used after any fixation (e.g. with formalin), and with either alcoholic or aqueous solvents. Nevertheless, X-ray contrast staining is an indirect imaging approach, which involves changing the sample’s chemical composition and eventually imaging the concentration of the staining agent. This poses concerns regarding the compatibility of such methods with a correlative imaging workflow mentioned earlier. Also, tissue shrinkage due to staining is a common issue that typically needs to be minimised [109, 292, 293], which becomes relevant when accurate morphological quantification is of interest.

X-ray µCT imaging of unstained soft tissues becomes possible with phase-contrast imaging. These methods exploit the change in the phase of X-ray waves when they interact with the sample’s internal and boundary features (see section ‘1.4 Advanced imaging approaches in research’). X-ray

phase-contrast imaging approaches typically require complex experimental setups that involve the use of elements such as interferometers, twin zone plates, Fresnel zone plates, or diffractive/refractive optics [173]. These are often exceedingly difficult to implement due to several reasons, including the need for high mechanical stability of the elements, limited FoV and loss of beam intensity as it passes through the element material [289]. The simplest of the X-ray phase-contrast imaging methods, in-line or propagation-based phase-contrast imaging, does not require any additional elements compared to the conventional attenuation-contrast imaging. This method becomes feasible when the X-ray propagation is large enough to allow the physical phenomenon to develop, resulting in pronounced edges on the radiographies, called Fresnel fringes. This might entice experimentalists to make an uninformed selection of the propagation distance, as it seems to (always) improve the phase-contrast when increased. To aid in the process of selecting the optimal propagation distance, a protocol design flow chart was developed here (see Chapter 3 'Soft tissue X-ray µCT imaging'). It enables researchers to design a protocol that is optimal with respect to spatial resolution, taking into account technical limitations of the setup being used. This makes propagation-based phase-contrast imaging more accessible to experimentalists interested in using the technique as a non-destructive imaging tool within a correlative workflow.

Such a workflow was implemented by combining µCT imaging, digital unrolling and quantitative morphometry of crypt budding, which preserves the samples in a state ready for histology and immunohistochemistry. The imaging protocol developed for this purpose (see section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons') was designed so that it fits the entire colon cross-section in the µCT slices, by compromising the recorded spatial resolution. Although achieving higher spatial resolution, which could be directly comparable to conventional histology, is possible (see section '3.3.1 High-resolution synchrotron phase-contrast CT of murine colons'), this comes with a compromise in the available FoV. This approach would not provide an overview of the sample, but it would be suitable to study the microstructure of selected crypts of interest at a much higher spatial resolution.

The second essential method (ii.) that enables the study of crypt budding in cylindrical FFPE intact colons is that of digital processing of the µCT 3D data. Processing such data is no trivial task, with the challenging natural geometry of the colon being a crucial obstacle. To identify the budding crypts within the tissue, a method was developed to digitally unroll the colon first (see Chapter 4). In contrast with traditional techniques, which either involve cutting the colon open or extracting the crypts from the mucosa (see section '1.3 Standard approaches for crypt budding research'), digital unrolling is the key enabling technology that allows preserving the colon intact. It ensures that the colonic morphology is not affected further to any impact that chemical fixation and

embedding might have (tissue shrinkage due to dehydration is a common effect of this process [109, 272]). Digital unrolling then aids in quick and effective identification of budding and non-budding crypts, enabling the study of several crypts at once (in the order of thousands in each scan). At the same time, it is also possible to focus on a selected region of interest within the tissue. Therefore, it is no longer necessary to choose either a global or a local technique mentioned earlier, as both become concurrently available. The fact that µCT imaging is non-destructive allows a global (low-resolution) scan to be performed initially and be used to inform the choice of sub-volumes of interest for subsequent high-resolution scans and/or histology.

The second pillar of digital processing was the developed methods for morphological quantification of budding crypts in 3D. After identifying the budding crypts using 3D cycloramas, the crypts had to be digitally isolated from the rest of the tissue. The process involves a segmentation step, where relevant pixels are marked with a label that signifies whether they belong to the crypt or not. Section '5.1 Workflow for identification and 3D segmentation of budding' proposes a semi-automatic workflow to segment the crypt lumens through a region-growth algorithm. This method has, however, two major disadvantages: a. it is exceedingly time-consuming since it involves manual refinement of the segmentation (takes about 3 hours for each scan comprising ~20 crypts), b. it is only applicable to segment the crypt lumens, thus disregarding the epithelium. Of course, the crypt epithelium is visible in the µCT data and especially well-pronounced in high-resolution scans (see Figure 3.10 and Figure 8.1). However, manually segmenting the epithelium is so labour-intensive that it becomes practically infeasible to perform for a large number of crypts. In cases where the microstructure of the crypt needs to be segmented in high granularity (separate labels for lumen, epithelium, cells, etc.) the problem is even more difficult. It requires that the labels are allowed to be disconnected (e.g. individual cells do not lie in a single connected region). The problem is in this case termed pixel classification, and the different regions are called classes instead of labels [294]. Solutions to the general problem of image classification are quite broad with methods ranging from simple thresholding of the grey value, to advanced techniques such as those employing artificial neural networks [294-296]. As these methods are tightly related to the data they are applied to, the selection of the best approach often depends on the employed imaging modality. Consequently, researchers tend to develop custom methods for the datasets at hand, with 3D datasets of soft tissue being among them [297-299]. Machine learning methods use a small set of manually segmented data to train a classifier that segments the rest of the dataset automatically. As machine learning is becoming increasingly popular, packages such as the trainable Weka segmentation in ImageJ/Fiji [300] or the standalone package Ilastik [301] are gradually emerging.

Considering what mentioned above, the proposed approach of preserving the colons in unstained FFPE tissue blocks, employing non-destructive X-ray phase-contrast imaging and subsequent digital unrolling and quantification seems to be a viable workflow for the study of crypt budding in CRC. It not only inflicts minimal tissue deformation during the sample preparation step, but it also allows for histology/immunohistochemistry to be performed post-imaging. Yet, there is still much to be done, primarily towards expediting the digital segmentation procedure and efficiently extracting the 3D structures from the 3D image data, in order to make this workflow attractive enough to biomedical researchers.

8.3 On the 3D morphology of budding crypts and its relevance to CRC

X-ray µCT imaging provides a 3D structural representation of the sample. It is thus suitable for morphological quantification of the internal features of the tissue. This was showcased by segmenting the crypts within the murine colon samples and quantifying their size and topology by means of an elaborate set of morphological characteristics defined for this purpose (see section '5.2 Definition of budding and non-budding crypt characteristics'). The fundamental goal was to develop tools able to thoroughly phenotype budding and non-budding crypts, which is remarkably needed for computational model development. Thus, each crypt was decomposed to its constitutive branches, which were quantified separately. The individual branches were then reassembled using the topology graph, which represents how individual branches are interconnected, providing a detailed representation of the entire budding or non-budding crypt.

This approach provided for the first time a method to draw a detailed and comprehensive picture of budding crypts in their natural 3D form. It not only quantifies the morphology of individual branches, but it also enables to associate this information with the location of each branch within the budding crypt. This detailed description provides the capability to test concepts put forward in the literature that involve the quantification of the crypt structure. Boman *et al.* [24] for instance, proposed that the distribution of the APC protein and components of the Wnt signalling pathway along the crypt axis should determine a 'sweet spot' that is the most likely location for branching. This can be directly quantified by the branch relative position (see Table 5.1), which measures the branch's parent node position relative to the crypt's height. To quantify the asymmetry of budding crypts, either the asymmetry index proposed by Tan *et al.* [6] or the symmetry ratio proposed by Langlands *et al.* [22] can also be measured. The two measures were extended for more than two branches and it was experimentally found that the symmetry ratio achieves better differentiation of budding crypts than the asymmetry index.

In terms of morphological phenotyping, the work presented in Chapter 5 'Tools and workflows for 3D morphological quantification of budding structures', extends the work of Tan *et al.*, who were the first to systematically evaluate the morphology of crypt budding [6]. The authors quantified basic measures of bifurcating crypts, including the branch length and width, and raised the question of asymmetry as a characteristic of abnormal budding. Here, the developed tools enable the study of crypts with more than two branches, extending the relevant measures in 3D and providing a means of comparison between heterogeneous structures such as crypts with different number of branches. The topology graph plays a pivotal role in this comparison, as it assigns each branch with a generation, which associates morphological measures with their position within the budding crypt. This enabled the detailed quantification of complex budding structures observed or hypothesised in the literature [6, 9, 22] but not quantitatively evaluated so far. For instance, Wong *et al.* have observed frequent asymmetric budding in crypts excised from adenomas [19]. In an attempt to explain the mechanism behind these observations, Tan *et al.* hypothesised that this is a result of perturbation of the normal crypt production sequence (fission), where "the parental crypt starts budding before the original daughter crypt has finished its crypt production sequence" [6]. This hypothesis is consistent with the results presented here (see Chapter 6), where it was possible to confirm that budding is primarily a result of repeated bifurcation.

The analysis performed in Chapter 6 'Quantitative morphometry of crypt budding in a mouse model' is in line with the current understanding of crypt budding during CRC development. More specifically, it is believed that carcinogenesis begins with a series of mutations in a stem cell of a normal crypt that develops into a monocryptal lesion. This in turn results in tissue disorganisation characteristic of dysplasia as a consequence of abnormal and excessive crypt fission [13, 19, 20, 39, 62, 83]. Wong *et al.* [19] advocated abnormal crypt fission or budding as a putative mechanism behind tumour initiation and growth in human CRC [9, 52, 84]. It is believed that through this process of fissioning, mutated monocryptal lesions develop into ACF [39, 86] that eventually become adenomas. Though, throughout the literature, these processes have been studied mainly from the functional point of view, i.e. focusing at the role of the signalling pathways and the deregulation of their normal function. Thus, their impact on the colonic structure, i.e. crypt budding and tissue disorganisation, has been only qualitatively described. Furthering our understanding in this direction, the methods described here provided quantitative data related to the number of times each crypt bifurcates, the shape of the resulting branches, the budding crypt topology and relevant sizes. As mentioned above, the measurements presented in Chapter 6 support the well-established notion that crypt budding is a result of repetitive bifurcation (fissioning) process [13, 14, 113] via quantitatively validating that each budding event produces two new branches on average. Measurements of individual branch sizes (branch

volume, surface area, etc.) supported the concept of migrating cell flows originating in stem cell niches at the bottom of leaf branches and being added up at each branching point. Importantly though, as these branch sizes are now quantitatively evaluated and identified within the broader picture of the budding crypt, they become relevant to theorists who develop quantitative computational models for crypt budding (see section '1.3.4 Computational models of the intestine').

Recently emerging concepts in the field of computational modelling put forward the role of biomechanical processes in crypt budding [6, 22, 23, 90, 124, 154]. For instance, Pin *et al.* argued that fission occurs due to stiffness differentials introduced by neighbouring concentrations of stem cells and $cKit^+$ cells in the colon [90, 154]. This notion challenged the belief that the initiation of crypt fission was related to the crypt size, as proposed by earlier models [10, 25, 26, 153]. Studies focusing on biomechanical stiffness differentials typically employ *in vitro* (2D) crypt organoids, thus the findings remain to be validated *in situ* in animal tissue. In this case, a systematic approach would be needed to quantify and compare the crypt morphology in 3D, which is now possible following the workflows in Chapter 5. These workflows pave the way to systematic phenotyping on budding crypts in 3D. Viewed within the context of non-destructive imaging of FFPE murine colons, followed by quantitative morphometry and 3D histology put forward here, the developed methods describe a valid approach to the structural and functional study of early-stage CRC. Further to their theoretical description, the quantification methods have been applied to an AOM/DSS murine model, and the results have been presented and analysed in Chapter 6. As recent works in the field of computational modelling of the colon begin to focus on systems of multiple crypts [27, 302], these measurements could be readily used to inform existing or new models.

8.4 Workflow limitations

Albeit the specific limitations and shortcomings of the individual techniques outlined above, it is worth examining the shortcomings of the workflow as a whole. Considering that the end goal of this workflow is the accurate structural quantification of the budding crypts, it is pivotal that the deformation imposed to the tissue by the sample preparation protocol is minimised. Here, the well-established method of chemical fixation via formalin and embedding in paraffin wax was employed, in order to allow for subsequent histological and immunohistochemical assays. However, it is a well-known fact that tissue shrinks in volume during this process [109, 303]. While it has been shown that tissue shrinkage due to formalin fixation is homogeneous for lung tissue [272, 304], a systematic study for the colonic sausages employed here remains to be performed. Firstly, the most appropriate fixative should be identified for the study at hand, where criteria

might include tissue shrinkage, compatibility with subsequent histological/immunohistochemical methods, and fixation time. B. Metscher for instance, in a study that considers 9 common fixatives, argues that Bouin's solution is better than formalin in terms of the criteria mentioned above, despite formalin being the most common tissue fixative for µCT imaging [305]. Once the fixation solution has been selected and optimised, the tissue volume shrinkage should be estimated so that subsequent measurements on the µCT data can be adjusted for shrinkage.

Another limitation of the sample preparation method employed here is the fact that it is an *ex vivo* technique. As a result, the crypt's evolution can only be observed at a fixed point in time, which means that time-related studies can only be conducted by observing tissues of different animals culled at progressive time points. Any outcomes from such experimental approaches will always be susceptible to mismatch of experimental conditions, biological variability and statistical errors. To observe the evolution of individual crypts in live animals, recently developed techniques involve surgically installing an abdominal imaging window, and subsequent *in vivo* imaging of the crypt using a customised two-photon microscope [11, 135]. Imaging can be performed over several days, thus allowing the study of the evolution of individual crypts over time. Of course, this approach only allows for the study of a small number of crypts but it provides information delated to intestinal stem cells, crypt size, composition and dynamics, with a single-cell resolution [11]. Such methods could be used in conjunction with the workflow proposed here to validate hypotheses emerging from certain structural observations such as the existence of cell-flows discussed earlier, for instance. It could also be used for exact measurements of the crypt size *in vivo*, and subsequent calibration of the crypt size as measured on the µCT data.

With respect to the exploratory animal study employed here, specific limitations were discussed in section 6.4. Key issues can be summarised as follows: a small number of animals (6× per group) was used as a proof-of-concept and the animal model was also not customised to test a particular hypothesis. This has implications in both the culling times and the sampling sites, which would need to be customised to suit specific hypothesis of future studies.

For the X-ray phase-contrast imaging protocol design flowchart key issues can be summarised as follows. The set of requirements taken into account does not consider limitations of X-ray flux (which is common in laboratory sources) or any criteria related to signal noise (see section '3.1.2 Minimal requirements for phase-contrast imaging'). As discussed above, considering such constraints poses an optimisation problem that does not have a unique solution.

For the 3D cyclorama method, the main drawback is related to the adopted mapping approach through the DTW algorithm. This results in deformations distributed throughout the mapped volume, though without considering any criterion for optimality. Alternative mapping approaches

of manifold embedding could be used to impose such desired properties for the mapping of the 3D re-slicing surfaces. For instance, isometric mapping [253] could be used for minimal deformation or moving least squares [254] for maximal smoothness (see section 4.3.1).

Last, but arguably the most important limitation of the developed workflow, is the fact that it yields a small throughput. Access to synchrotron sources for µCT imaging is not only granted through a competitive process, it is also limited in terms of beamtime (a few days granted every six months). On the other hand, imaging in the lab is exceedingly slow, with scan times in the order of 18-20 hours per sample (see section '8.1 X-ray phase-contrast imaging from the synchrotron into the lab'). Even after imaging, the process of 3D image segmentation is labour-intensive and exceedingly time-consuming (see section '8.2 On the feasibility of using X-ray CT for the study of colorectal cancer'). Finally, although not employed here, a subsequent step of serial sectioning would be appropriate to complete the structural information provided by µCT with 3D functional information provided by immunohistochemistry. Such approaches are recently emerging for lung tissue [306], where 3D image registration is a major issue that is being confronted. Specifically, paraffin sectioning introduces artefacts such as tissue elongation in the direction of cutting, streaks and cracks [123], which are accounted for through a semi-automated and labour-intensive image registration workflow in the work of Lawson *et al.* [306]. This workflow involves pairing features found in both the µCT and histology slices, and subsequent non-rigid transformation of the histology image to match the µCT data of individual slices in the 3D image stack. Taken together, all the time-consuming constitutive parts of the workflow proposed here, make it infeasible to apply repeatedly in a short time frame. Instead, this workflow would be more suitable in the context of a meticulously designed experiment over a time scale of several (6 or more) months.

8.5 Summary

The work presented here focuses on the development of new tools and workflows for the study of crypt budding, as a specific case of the broader problem of X-ray µCT imaging and extracting useful information from 3D imaging data. A gap has been identified between the X-ray µCT technology developers and users, and an attempt has been made to bridge this gap by developing tools and workflows that facilitate the design of X-ray phase-contrast imaging protocols, digital unrolling of tubular structures and structural quantification of budding structures in 3D. X-ray phase-contrast imaging was straight forward for synchrotron sources, but much more challenging in the lab, where limiting factors include the limited X-ray flux and FoV.

The 3D cyclorama was the key enabling technology for *in situ* quantification of the colonic crypts. Digital unrolling via a simple mapping approach (through the DTW algorithm) was feasible here but more advanced methods were proposed for improved mapping approaches in terms of minimisation of the imposed image deformation.

3D quantification of crypt budding provided for the first time a method to draw a detailed and comprehensive picture of budding crypts *in situ*, in 3D. The methods developed here extended the work of Tan *et al*. [6], who were the first to systematically evaluate the morphology of crypt budding, and introduced new measures for topological and morphological quantification of budding structures in general.

Regarding the workflow, which integrates X-ray imaging, digital unrolling and structural quantification of the budding crypts, the need for compatible correlative approaches has been underlined (X-ray imaging followed by immunohistochemistry), though with no significant impediments in foresight. Tissue shrinkage during fixation was identified as an important issue that needs to be addressed for accurate morphological quantification, as well as appropriate animal model design for the particular scientific question at hand for future studies. Finally, bottlenecks in gaining access to synchrotron sources, in the process of 3D image segmentation, and 3D image registration after paraffin sectioning, were identified as the key limitations for routine application of the developed workflow.

Chapter 9 Conclusion and outlook

Employing CRC as a principal motivating application, the necessary methods, tools and workflows were developed here to non-destructively image murine colon tissue and systematically quantify crypt budding in 3D. An animal model for CRC was employed and crypt budding was quantified using the newly developed tools, resulting in a complete phenotype and a generic methodology that can be used by researchers in the future. The several tools and methods developed here may be combined in a workflow that involves the use of an animal model, preparation of FFPE intact colons in cylindrical blocks, non-destructive X-ray µCT imaging, digital unrolling and quantitative morphometry of murine colonic crypts. The FFPE colons may then be sectioned and stained following conventional histology or immunohistochemistry procedures resulting in both structural (through µCT) and functional (through histology) information about the sample.

The effect of biomechanical forces on crypt budding is lately receiving increased interest (see section '1.2 Colorectal cancer and crypt budding'). These forces result in biomechanical stiffness that is built-up as a consequence of the arrangement and numbers of different cell types (e.g. $cKit^+$ and $Muc2^+$). An alternative approach suggests that the distribution of proteins, including APC and proteins in the Wnt pathway, along the crypt epithelium, determines the 'sweet spot' for branching. Therefore, the 'Holy Grail' in understanding crypt budding during early-stage CRC appears to be related to elucidating the mechanisms that regulate the distributions of cells and proteins along the crypt epithelium. Outstanding results in this direction have been achieved by confocal imaging of excised crypts [6] and crypt organoids [22]. However, imaging crypts in such a high spatial resolution imposes technical limitations (exceedingly small pixel size results in small FoV, small sample size required because staining needs to penetrate the tissue, etc.). These limitations practically enforce the techniques to be applied to tiny pieces of tissue or even isolated crypts or organoids. In this context, the workflows developed here could be used initially to identify budding crypts of interest within the intact colon, which could be subsequently digitally isolated or physically excised so that appropriate imaging of proteins can be performed.

Functional information with single-cell resolution has also been achieved with *in vivo* two-photon imaging of murine intestinal crypts [11, 135]. Such methods are invaluable in the study of crypt budding, as they avoid sample-preparation techniques that might affect the morphology of the tissue, which is a major concern when chemical fixation is involved. Thus, *in vivo* imaging could be combined with the workflow developed here to correct the µCT measurements for tissue deformation due to the sample preparation method, i.e. due to chemical fixation and paraffin embedding. Not only an exact measurement of the crypt dimensions can be achieved *in vivo*, but

also its function (protein concentration) can be observed over a time window of several days. These observations would therefore be useful for validation of hypotheses that arise for the analysis of the µCT data emerging from the workflow developed here.

Readily useful outputs of the work presented here are the measurements of the budding crypts in Chapter 6: 'Quantitative morphometry of crypt budding in a mouse model'. The measurements of the morphological characteristics of crypts and individual branches constitute a highly detailed phenotype of crypt budding in an AOM&DSS mouse model for CRC. They can be used to inform existing, or develop new computational models for crypt budding that may extend in 3D. For the first time, theorised concepts of crypt budding (e.g. the concept of repeated bifurcation leading to dysplasia) and simple morphological characteristics have been quantified systematically, in 3D and with minimally invasive sample preparation procedures. Therefore, both the methods in Chapter 5 and the computed phenotype in Chapter 6 could be of significant importance in the field of computational modelling of crypt formation, development and budding. These models could be employed to compute the biomechanical stiffness in the budding crypts, which in turn could be used to derive the concentrations of different cell types [23, 27]. Such an approach follows the reverse order of the methods mentioned above, making predictions using the best available model of the underlying biological processes. When these models accurately recapitulate reality, they often yield insights that lead to unforeseen hypotheses that would have not been formulated otherwise, paving the way to ground-breaking discoveries.

Further to the quantification of crypt budding, the methodology of Chapter 5 is in principle applicable to different branching structures abundant in the biomedical world. Airway and blood vessel networks in the lung [273, 274], neural networks in the brain [275], even studying the growth of plant roots [307, 308], are some examples where branching structures are under investigation by researchers. Although imaging these structures in 3D is becoming increasingly more popular in this process, extracting useful information is often not trivial. The methods described in Chapter 5 provide the capability of performing detailed morphological quantification, down to individual branch level, which could facilitate deeper understanding of the underlying (biological) processes.

The methods and tools developed here extend further to a broad range of applications, which span a large range of scientific fields. Application areas of X-ray µCT include archaeology and palaeontology, biology, biomedical and life sciences, engineering and physical sciences or material science [3-5, 165]. As a result, it can be rather difficult to attain a certain level of knowledge and/or experience in engineering, material science, mathematics, physics and X-ray imaging at the same time, particularly for researchers with different academic backgrounds, and when the

literature becomes too technical. Here the key elements of X-ray imaging essential in understanding propagation-based phase-contrast imaging were portrayed, and the minimal set of technical requirements to achieve the best experimentally attainable spatial resolution was derived. This approach revealed that there was an optimal X-ray propagation distance that yield the best attainable spatial resolution. It was therefore experimentally confirmed here that the spatial resolution only follows the theoretical prediction when the presented set of minimal technical requirements were satisfied. The provided protocol design flowchart was capable of predicting this optimal X-ray propagation distance without the need for experimental derivation. It can thus be used by researchers from different fields, to design X-ray phase-contrast imaging protocols and expedite their experiments, both using laboratory and synchrotron X-ray CT setups.

Another method with a broad range of applications developed here is the digital volume unrolling using 3D cycloramas. Although primarily developed to unroll murine colons, the method could provide novel insights into applications where unrolling of deformed cylindrical tubes of non-uniform thickness is necessary. A specific example put forward as a case-study in Appendix A.2.2, is the virtual unrolling of an ancient lead amulet. As the glyphs were inscribed within the lead sheet, the retrieved text within the unrolled volume extended over several slices of the 3D cyclorama. This is a prime example where a volumetric approach of scroll unrolling is greatly beneficial over approaches that deem the scroll as a single surface [236, 237, 239, 241-244]. Through this collaborative work with the Australian Centre for Neutron Scattering and MACQUARIE University museum of ancient cultures, it was possible to retrieve text inscribed on the rolled lead sheet. Scholars at Rutgers University and the University of Cambridge identified the scripture language as ancient Mandaic, which is infrequently found in intact scrolls, establishing the MU4106 lead amulet as a rare ancient artefact.

As the 3D cyclorama method is capable of flattening curved sheets, it was possible to extend the collaboration with the Australian group and employ 3D cycloramas to flatten clay envelopes and tablets. The conventional approach of presenting such artefacts in the literature involves the arrangement of orthogonal photographic views of the envelop or tablet in the shape of a cross [309]. Imaging the artefacts in 3D and flattening them using the 3D cyclorama method provides an alternative way to present and decipher these ancient findings. Similarly to the curved envelopes and tablets, 3D cycloramas were used to digitally flatten jet engine parts with dual curvature for a leading aero-engine manufacturer in the UK. The goal was to achieve a 10-fold reduction in the time taken for non-destructive evaluation of manufacturing defects through CT, with digital flattening being a crucial step in the process. The flattening capabilities of the 3D cyclorama method have been showcased here in a case-study for this purpose, where a pharmaceutical film made using additive manufacturing techniques has been digitally flattened and examined for

deposition defects (see Appendix A.2.1). Taken together, these examples show the great potential of the 3D cyclorama method to facilitate research and industrial applications with a context much broader than digital unrolling of murine colons. To this end, the software has been reformed into an ImageJ/Fiji plugin and made available to the public through the Fiji update sites (see https://imagej.github.io/plugins/3d-cyclorama).

Appendix A 3D cyclorama extended

A.1 Validation using a digital phantom

To validate the method of Chapter 4 'Colonic 3D cyclorama: a method for digital volume unrolling', a 3D digital phantom was generated (Figure A 1 – A 4) in two steps. First, a 2D grid template was created *in silico*, which was used to create both a deformed cylindrical tube (digital phantom) and a 3D template stack (ground truth). The digital phantom (Figure A 2 & Figure A 4b) was created by rolling and stretching the 2D grid template (Figure A 1a) by means of an analytical mapping devised for this purpose (Figure A 3). The 3D template stack or ground truth (Figure A 4a) was built as a planar version of the digital phantom using the same 2D grid template. The 3D cyclorama method was then applied to unroll the phantom volume and compare the resulting 3D cyclorama (Figure A 4c) visually and quantitatively with the 3D template stack as ground truth. Next this process is described in detail.

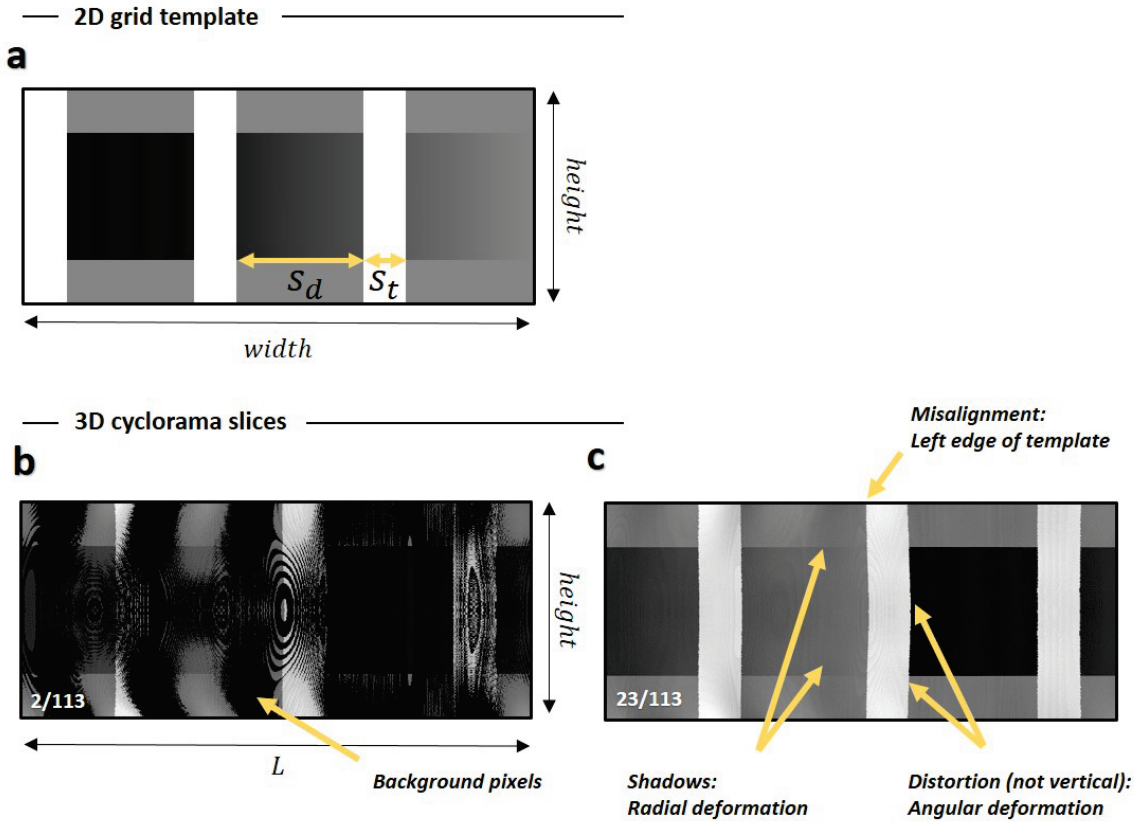


Figure A 1: 2D grid template and slices of the 3D cyclorama after unrolling the digital phantom

(a) The 2D grid template used to create the digital phantom is composed as a greyscale gradient with superimposed horizontal and vertical bands. These serve

both as visual cues and as features for quantitative evaluation of the deformation after unrolling. (b-c) Two slices of the 3D cyclorama after unrolling the digital phantom at 3% (2/113) and 21% (23/113) relative depth level. Dark patches on the 3% (2/113) slice emerge due to probing of background pixels during the mapping of superficial surfaces close to the inner boundary. Angular deformation appears as curved vertical lines on the 21% (23/113) slice, while deformation in the radial direction appears as shadows due to probing of pixels from neighbouring depths. The cycloramas were aligned with respect to the 2D template, since the cutting edge is defined azimuthally by the arbitrary first point of the contours.

First, a 2D grid template image $g_{h,j}$ was created, with pixel greyscale values according to the following rule:

$$g_{h,j} = offset + j \cdot height + h, \quad j = 0 \dots width - 1, h = 0 \dots height - 1, \qquad (A\,1)$$

where $offset > 0$ is an arbitrary positive integer and $width$ and $height$ are the image width and height in number of pixels, respectively. In a next step, two single-valued horizontal and three single-valued equidistant vertical stripes with thickness s_t and distance s_d were superimposed in order to create a rectangular grid (Figure A 1). The arbitrary offset was chosen to be positive so that after rolling of the template onto a deformed cylindrical tube (Figure A 2), it can easily be differentiated from the background (greyscale value of background = 0). The two indices j and h introduce a gradient (Figure A 1) to create features of varying greyscale values, necessary for quantitative comparison between the 3D cyclorama (unrolled version of the deformed cylindrical tube) and the 3D template (ground truth). The horizontal and vertical boundaries of the stripes create visual cues that help assessing the amount of deformation in the cycloramas in a qualitative manner. The 2D grid template image $g_{h,j}$ (Figure A 1a) was implemented using the following sizes in number of pixels: $offset = 296, width = L = 1200, height = H = 500, s_d = 300, s_t = 100$.

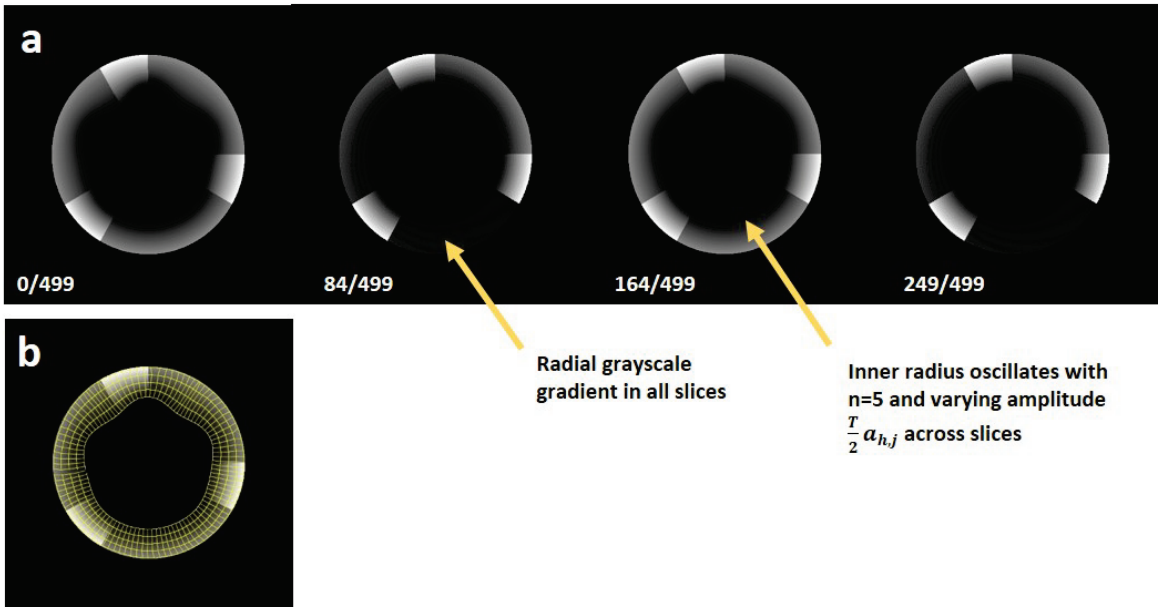


Figure A 2: Cross-sections of the 3D digital phantom

(a) Cross-sections of the digital 3D phantom at incremental heights ($0 \leq h \leq 499$). The inner radius $r_{h,j}$ is modulated by a sinusoid of changing amplitude and frequency, creating a phantom shape that is challenging to unroll. Greyscale values of the phantom also change across its wall thickness, with the outer side being brighter (larger grey values towards the outer boundary) than the inner side (smaller grey values towards the inner boundary). (b) Electric field lines and contours used to unroll the digital 3D phantom ($C = 5$ depth levels are shown here for visual clarity).

The geometry of a deformed cylindrical tube (digital phantom) was defined as follows: The outer boundary was described as a circle with constant radius across all heights of the tube (Figure A 3):

$$x_{out,j} = x_c + \lfloor R \cdot \cos(\theta_{1,j}) \rfloor , \quad (A\,2)$$

$$y_{out,j} = y_c + \lfloor R \cdot \sin(\theta_{1,j}) \rfloor , \; j = 1 \dots \lfloor 2\pi R \rfloor . \quad (A\,3)$$

with outer radius R, centre of the circle (x_c, y_c), azimuth angle $\theta_{1,j} = [0, 2\pi)$ and $\lfloor \quad \rfloor$ denoting the floor function, while the inner boundary at a certain height h was defined as follows:

$$x_{in,h,j} = x_c + \lfloor r_{h,j} \cdot \cos(\theta_{1,j}) \rfloor , \quad (A\,4)$$

$$y_{in,h,j} = y_c + \lfloor r_{h,j} \cdot \sin(\theta_{1,j}) \rfloor , \quad (A\,5)$$

$$r_{h,j} = R - \frac{T}{2} a_{h,j} , \quad (A\,6)$$

$$a_{h,j} = \left(\frac{1+\sin(n \cdot \theta_{1,j} + sin(\varphi_{1,h}))}{2} \cdot \frac{1+\sin(\varphi_{2,h})}{2} \cdot \frac{1+sin(\theta_{2,j})}{2} \right) + 1, \; j = 1 \dots width, \; h = 1 \dots height , n \in \mathbb{N} , \quad (A\,7)$$

$$r_h := \min_j r_{h,j} = R - T \leq r_{h,j} \leq \max_j r_{h,j} = R - \frac{T}{2},\ j = 1 \dots width,\ h = 1 \dots height, \qquad (A\ 8)$$

with the real maximal tube thickness $0 < T < R$ and $width = \lfloor 2\pi R \rfloor$. The real amplitude $a_{h,j} \in [1,2]$ modulates the inner radius $r_{h,j}$ of the tube (r_h: minimal inner radius at height h). For $a_{h,j}$, changes with angles $\theta_{1,j} = [0, 2\pi)$ and $\theta_{2,j} = [-\pi, \pi)$ are both linear azimuth angle sequences with number of elements equal to $width$, while $\varphi_{1,h} = [0, \pi)$ and $\varphi_{2,h} = [-\pi, 0)$ are polar angle sequences with number of elements equal to $height$. The inner radius r_h takes values in the range of $\left[R - T, R - \frac{T}{2}\right]$, yielding a varying tube thickness of $R - r_{h,j} = \frac{T}{2} a_{h,j} \in \left[\frac{T}{2}, T\right]$, as if the tube was radially 'stretched' and 'compressed', leading to an azimuthally changing tube thickness. Finally, the integer n, which is multiplied with $\theta_{1,j}$, defines the frequency with which the inner radius oscillates (Figure A 2). In essence, the modulating amplitude $a_{h,j}$ enforces the tube's or phantom's thickness to change locally, creating a shape that is challenging to unroll. To create the phantom shown Figure A 2, the frequency $n = 5$ was arbitrarily selected. A 3D image stack $G_{h,\rho_{h,j},\theta_{h,j}}$, which represents a deformed cylindrical tube with the geometry described above, was then created by probing the grey values of the 2D grid template $g_{h,j}$ as follows:

$$G_{h,\rho_{h,j},\theta_{h,j}} = w_{\rho_{h,j}} \cdot g_{h,j},\ j = 1 \dots width,\ h = 1 \dots height,\ w_{\rho_{h,j}} \in \mathbb{R}, \qquad (A\ 9)$$

where $w_{\rho_{h,j}} = \left[1,\ R - r_{h,j}\right] = \left[1,\ \frac{T}{2} a_{h,j}\right]$, $\frac{T}{2} a_{h,j} \in \left[\frac{T}{2}, T\right]$ is a radius-dependent weight that introduces a linear scaling of the grey values of the 2D grid template across the depth of the tube (Figure A 2), which is proportional to the varying tube thickness (Figure A 3b). The radius $\rho_{h,j} = \left[r_{h,j}, R\right] = \left[R - \frac{T}{2} a_{h,j}, R\right]$ ranges from the inner to the outer boundaries of the deformed cylindrical tube $G_{h,\rho_{h,j},\theta_{h,j}}$ at height $h = [1, height]$ with the azimuthal angle $\theta_{h,j}$ defined as follows:

$$\theta_{h,j} = \theta_{1,j} + \max\left\{R - \frac{T}{2} - \rho_{h,j}, 0\right\} \cdot \left(a_{h,j} - \frac{3}{2}\right). \qquad (A\ 10)$$

Equation (A 9) defines a mapping between index j on the grid template and azimuth angle $\theta_{h,j}$ on the digital phantom, introducing a non-rigid deformation. Radial and angular deformations are introduced to the deformed tube by equations (A 6) and (A 10), respectively. These deformations mimic the case where a perfect tube with outer radius R and inner radius $R - \frac{T}{2}$ is non-rigidly deformed by stretching it inwards at specific azimuth angles. The non-rigid deformation that is introduced in this case is implemented analytically by offsetting the angle $\theta_{1,j}$ by $\max\left\{R - \frac{T}{2} - \rho_{h,j}, 0\right\} \cdot \left(a_{h,j} - \frac{3}{2}\right)$. The first factor $\max\left\{R - \frac{T}{2} - \rho_{h,j}, 0\right\}$ makes this offset non-zero when the tube's thickness $R - r_{h,j} = \frac{T}{2} a_{h,j} \in \left[\frac{T}{2}, T\right]$ is larger than $\frac{T}{2}$ and zero elsewhere, thus

introducing deformations only at the inner half of the tube's thickness (Figure A 3c). The second factor $\left(a_{h,j} - \frac{3}{2}\right)$ of this offset shifts the mean value of $a_{h,j}$ to zero, thus ensuring that the template is compressed ($\theta_{h,j}$ follows $\theta_{1,j}$) when the offset is negative and on the other hand, stretched ($\theta_{h,j}$ precedes $\theta_{1,j}$) when the offset is positive (Figure A 3c). Since the deviation of the inner radius $r_{h,j} = R - \frac{T}{2} a_{h,j}$ from the midline at radius $R - \frac{T}{2}$ is proportional to the offset $\max\left\{R - \frac{T}{2} - \rho_{h,j}, 0\right\} \cdot \left(a_{h,j} - \frac{3}{2}\right)$, the degree of compression and stretching follows a similar pattern. The digital 3D phantom was implemented with $R = 190$ pixels, $T = 114$ pixels, $n = 5$, $x_c = y_c = 287$ pixels, resulting in an image stack of 500 slices, with a dimension of 573×573 pixels^2 each.

Finally, the 3D template stack (ground truth) was created using the same 2D grid template $g_{h,j}$:

$$g_{h,j,k} = w_k \cdot g_{h,j}\,,\ k, w_k \in \mathbb{N}\,, \qquad (A\ 11)$$

where $h = [1, height]$, $j = [1, width]$, $k = [1, T]$ and $w_k = k$ is a depth-dependent weight that introduces a linear scaling of the grey values of the grid template across the depth k of the 3D template stack. The implementation of this 3D template stack had 114 slices with a dimension of 1200×500 pixels^2 each.

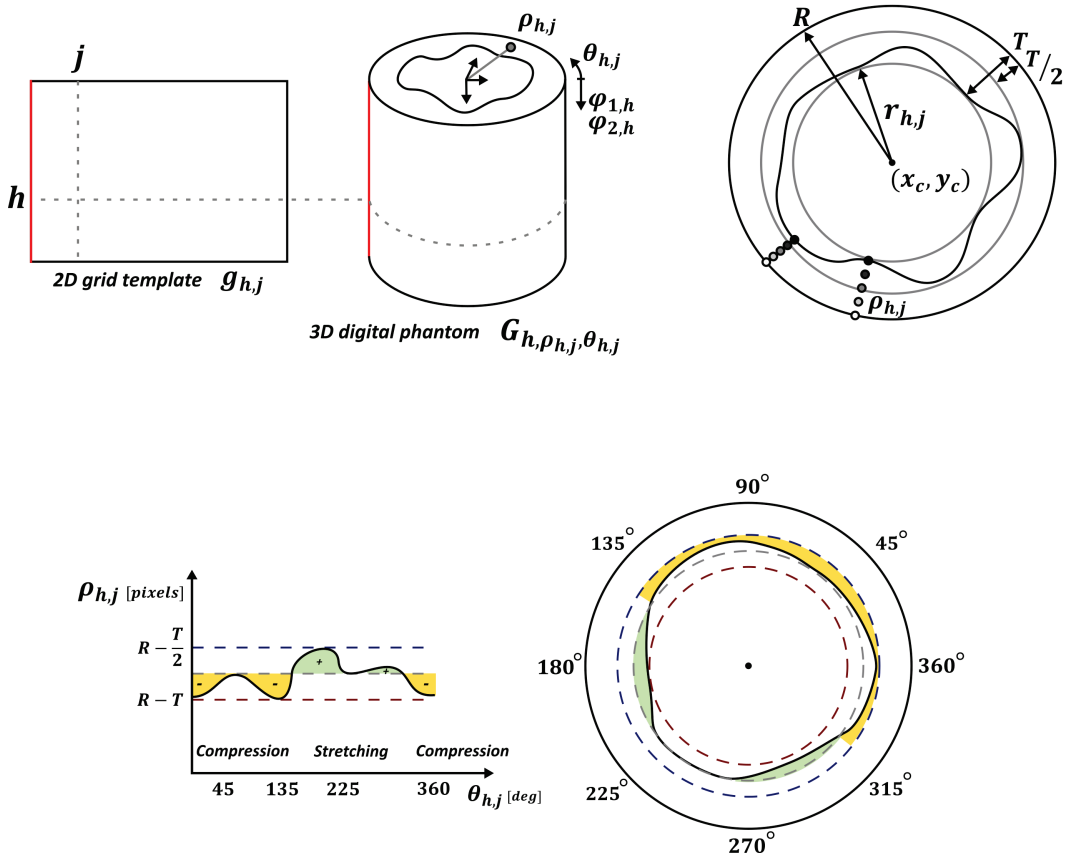


Figure A 3: Mathematical description of the 3D digital phantom

(a) A 3D digital phantom was created as deformed cylindrical tube (3D image stack $G_{h,\rho_{h,j},\theta_{h,j}}$) by rolling a 2D (rectangular) grid template image $g_{h,j}$ and mapping points (h,j) to points $\left(h,\rho_{h,j},\theta_{h,j}\right)$. (b) The wall of the digital phantom has a fixed outer radius R and an inner radius $r_{h,j}$ that follows a sinusoid of varying angular frequency and amplitude (minimal inner radius $r_h := \min_j r_{h,j} = R - T \leq r_{h,j} \leq \max_j r_{h,j} = R - \frac{T}{2}$), giving rise to a minimum and maximum thickness $\frac{T}{2}$ and T of the deformed cylindrical tube, respectively. The tube is assigned by linearly scaled greyscale values, where brightness levels are proportional to the radius $\rho_{h,j}$ at points $\left(h,\rho_{h,j},\theta_{h,j}\right)$ through sampling of the grid template $g_{h,j}$. (c) The undulating inner radius $r_{h,j} \in \left[R-T, R-\frac{T}{2}\right]$ results in a varying tube thickness $R - r_{h,j} = \frac{T}{2} a_{h,j} \in \left[\frac{T}{2}, T\right]$. The grey midline of this band at radius $\frac{1}{2}\left(R - \frac{T}{2} + r_h\right) = r_h + \frac{T}{4}$ can be considered as the boundary of a cylindrical tube that has been stretched and contracted by pulling and pushing its inner boundary, resulting in contracted ("−" : yellow) and stretched ("+": green) volumes inside the wall of the digital phantom.

This 3D template stack $g_{h,j,k}$ (Figure A 1 and Figure A 4a) was used as the ground truth to evaluate the result of the 3D cyclorama method when applied to unroll the digital 3D phantom $G_{h,\rho_{h,j},\theta_{h,j}}$. The quality of unrolling was quantified by estimating the non-rigid transform that maps the resulting 3D cyclorama onto the 3D template stack, by calculating and visualising the local volume change it introduced. To this end, *Elastix* [310], an open source collection of image registration algorithms, was used to non-rigidly register the two 3D image stacks. The software estimates the optimal transformation that maps the 3D cyclorama on the 3D template (Figure A 4) to maximise their mutual information as an image matching metric. *Elastix* [310] provides the Jacobian determinant α of the transformation at each point in space, which quantifies the volume of a transformed unit cube. In other words, the Jacobian determinant shows how much the space has been stretched or contracted at each point of the transformed 3D cyclorama to fit the 3D template, imposed by digital unrolling at each point in 3D. Identity of the Jacobian determinant means an exact match, i.e. the 3D cyclorama exactly matches the phantom, while a Jacobian determinant value $\alpha \neq 1$ means that the volume of the 3D cyclorama at that voxel is α times smaller ($\alpha < 1$) or larger ($\alpha > 1$) than the phantom. Making use of the Jacobian determinant rather than measuring the absolute difference between the 3D cyclorama and the 3D phantom, makes the quantification independent of the absolute greyscale values.

The boundary contours were defined by absolute thresholding of the digital phantom's grey values shown in Figure A 2 and Figure A 4b. Unrolling was performed using the linear mapping mode with $C = 114$ depth levels (Figure A 2b shows $C = 5$ depth levels for visual clarity), $I = 100$ point-long contours, an electric field search window of 31 points ($w = 15$), electric field line segment size $\delta = 2$ pixels, maximum number of steps $K_{max} = 57$, minimum electric field line variance $V_{min} = 10$, rigidity equal to 1, interpolation interval $m = 50$, and rotation axis origin $O = (300, 300)$. The resulting 3D cyclorama's length was $L = 1191$ pixels, the height was $H = 500$ pixels, and the depth was $C = 114$ slices. Angular deformations (see Figure A 1) appear as geometrical distortion along the horizontal axis of the cycloramas. Radial deformations appear as shadows in cycloramas as greyscale values from neighbouring depth levels *c* of the 3D phantom, resulting in a diffuse grey level distribution observed in the 2D cycloramas (Figure A 1).

A few pre-processing steps needed to be performed prior to comparing the 3D cyclorama quantitatively and visually with the 3D template stack as ground truth, for validation of the 3D cyclorama method. As contours begin azimuthally at an arbitrary first point, the 3D cyclorama is not necessarily aligned with the 3D template (Figure A 1). Thus, the two stacks were firstly manually shifted by rolling the 3D cyclorama (cropping the left part and attaching it to the right) so that the left side of the 3D cyclorama was aligned with the left side of the 3D template. Cycloramas close to the deformed tube's boundaries partially included pixels of the background

that appear as distinct dark patterns (Figure A 1b). Thus cyclorama slices (first four and last two depth levels $c = 0, 1, 2, 3$ and $c = 112, 113$) that contained background (black) pixels were removed from both the 3D cyclorama and the 3D template. Finally, as the cyclorama's length L is arbitrarily determined based on the length of the longest contour $cont_{h,c}$, it is slightly shorter than the 3D template (1191 instead of 1200 pixels). Therefore, the 3D cyclorama was scaled (*scal*e function in Fiji with bicubic interpolation) to fit the template dimensions, in order to obtain a better initialisation of the subsequent registration process and ensure convergence of the iterative registration algorithm in *Elastix*.

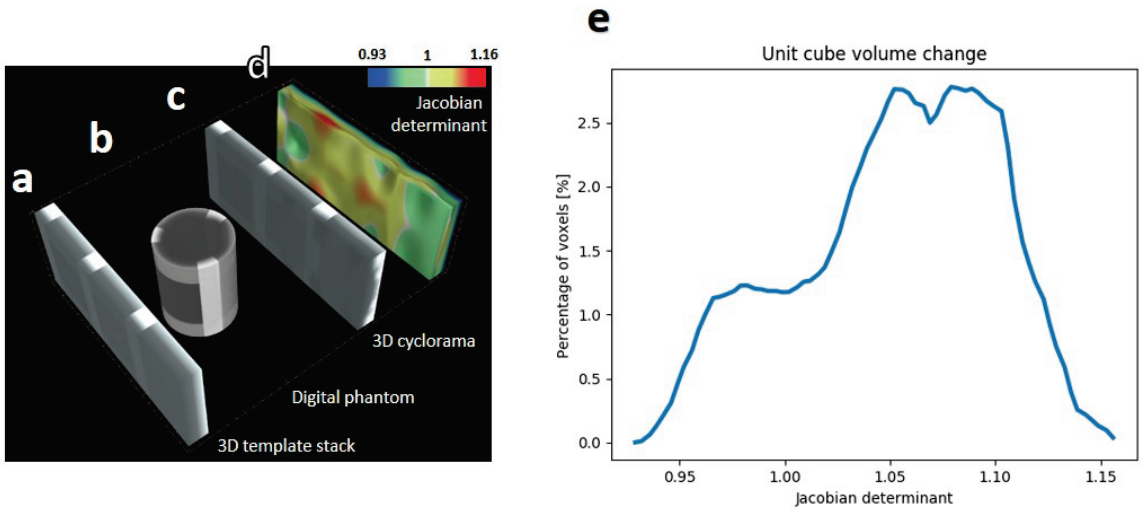


Figure A 4: Validation of the cyclorama method using a digital 3D phantom

(a) The 3D template stack (ground truth) or digital 3D phantom consists of a volume with greyscale gradients making each voxel (except those on the vertical bands) distinct. (b) The digital phantom was created by digitally rolling a 2D grid template onto a deformed cylindrical tube that is defined analytically. (c) The 3D cyclorama after unrolling, manual alignment (and scaling) to the 3D phantom. (d) The Jacobian determinant of the mapping shows the degree of volume change at each point in space. Identity of the Jacobian determinant means that the 3D cyclorama exactly matches the phantom, while a Jacobian determinant value $\alpha \neq 1$ means that the volume of the 3D cyclorama at that voxel is α times smaller (< 1) or larger (> 1) than the phantom. (e) Non-rigid registration of the 3D cyclorama onto the 3D phantom returned the Jacobian determinant at each point of the 3D template. The mean value $\alpha = 1.05$ of the Jacobian determinant distribution showed a general preference to preserve the volume. However, a skew to the right showed a trend of volume enlargement (up to 1.16 times). In this case, the maximum contraction was given by a Jacobian determinant of $\alpha = 0.93$.

The registration process returned a 3D image stack with each voxel representing the Jacobian determinant at the corresponding point of the 3D template. A visual representation of the deformation pattern (the value α of the Jacobian determinant) is provided in Figure A 4d, which shows that the deformation was larger on the side of the inner boundary of the cylindrical tube. The mean value $\alpha = 1.05$ of the Jacobian determinant distribution (Figure A 4e) shows a general preference to preserve the volume. However, a skew to the right shows a trend of volume expansion (up to 1.16 times), while the maximum contraction was given by a Jacobian determinant of $\alpha = 0.93$. The trend for volume expansion (Jacobian determinant $\alpha > 1$) shows that the 3D cyclorama method produces stacks where the volume is locally enlarged. The root of this observation is the choice of mapping each contour onto a template whose length is equal to the greatest length among all contours. As discussed in section '4.2.2 Mapping: Generation of 3D cycloramas', the mapping from a deformed tube onto an unrolled volume is not unique. Therefore, certain choices made to design the mapping method (e.g. preserving geodesic distances, mapping onto the longest contour, etc.) affect the characteristics of the unrolled volume. In the implemented case, there is a slight preference for volume expansion. Volume contraction ($\alpha < 1$) is introduced due to the radial deformation (see Figure 4.4b) imposed by the reformation of the volume into a series of re-slicing surfaces. The number of depth levels C essentially defines whether the volume will be expanded or contracted in the radial direction. This happens because one 2D cyclorama with a thickness of one pixel is created for each depth level, resulting in a 3D cyclorama of depth equal to C. Therefore, when the number of depth levels C is smaller than the maximum local thickness T, the volume of the deformed tube will be expanded radially so that the 3D cyclorama depth matches the maximum local thickness T. Similarly, the deformed tube volume will be contracted radially when C is larger than T. In the case of the digital phantom, where $C = T = 114$, the observed contraction originates in rounding errors, where the same voxel of the digital phantom is sampled by two neighbouring contours. The fact that volume expansion/contraction depends on the choice of an input parameter (number of depth levels or contours C) means that the deformation pattern would change accordingly and thus, quantitative measures derived from cycloramas must be interpreted with care (see section 4.3).

A.2 Interdisciplinary applicability: Case studies

Next, two distinct case studies are presented showcasing the capability of digital flattening of sheets (section 'A.2.1 Digital flattening of a pharmaceutical film') and scroll unrolling (section 'A.2.2 Virtual unrolling of an ancient lead amulet'). The µCT data was kindly provided by the

original authors of the respective case study, and digital flattening/unrolling was performed using the 3D cyclorama method.

A.2.1 Digital flattening of a pharmaceutical film

This case study shows how the 3D cyclorama method could be useful in the morphological assessment of 3D-printed adhesive films for drug delivery applications. Gioumouxouzis *et al.* [311] and Eleftheriadis *et al.* [312] employed fused deposition modelling (FDM 3D printing) in order to develop personalised drugs with unique properties and drug release behaviour. Eleftheriadis *et al.* subsequently evaluated the 3D-printed pharmaceuticals using µCT to compare their structure, when compared to the initial 3D design. This process becomes challenging when the 3D-printed object is flexible, as it is the case for certain pharmaceutical films that the authors are developing for time-controlled drug administration [312]. The $20 \times 20 \times 0.8$ mm^3 plain film shown in Figure A 5 is composed of poly(vinyl alcohol) as core polymer and plasticiser and is designed for unidirectional drug release. The µCT imaging resulted in a stack of $1000 \times 161 \times 901$ pixels^3 at an (isotropic) pixel size of 22 µm (volume of $22.0 \times 3.5 \times 19.8$ mm^3). Since the pharmaceutical film was curved (see Figure A 5a), it is difficult to quantify the deposition quality of the different drug layers. To this end, the 3D cyclorama method was applied to digitally flatten the film.

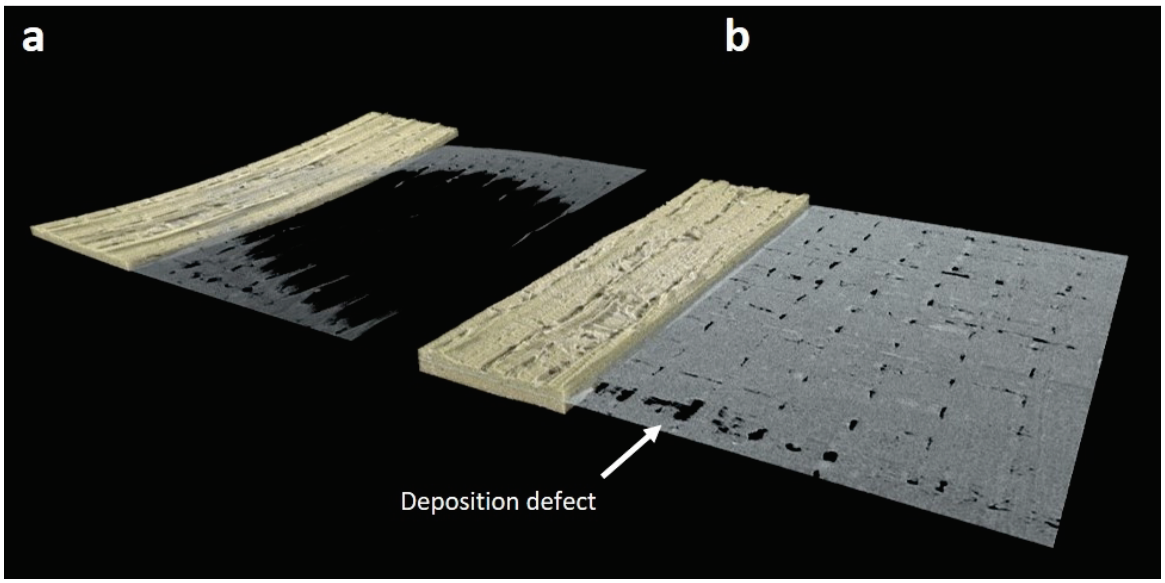


Figure A 5: 3D cycloramas to unroll a pharmaceutical film

Photorealistic 3D renderings of a pharmaceutical adhesive film for drug delivery imaged by µCT. (a) The curved nature of the 3D-printed film results in planar slices (grey) that only partially illustrate the film. (b) Flattening onto a 3D cyclorama can undo the inherent deformation and allow structural analysis of the pharmaceutical film layers, revealing deposition defects such as gaps in drug distribution (arrow).

The CT stack of 901 slices with dimensions of 1000 × 161 pixels^2 was unrolled (see Figure A 5b) using the linear mapping mode with $C = 50$ depth levels, $I = 100$ point-long contours, an electric field search window of 31 points ($w = 15$), electric field line segment size $\delta = 2$, maximum number of steps $K_{max} = 50$, minimum electric field line variance $V_{min} = 10$, rigidity equal to 1, interpolation interval $m = 50$ and rotation axis origin $O = (340, 40)$. The resulting 3D cyclorama's length was $L = 2025$ pixels, the height $H = 901$ pixels, and the depth $C = 50$ slices.

The 3D cyclorama enables the quantification of the quality of drug deposition in each slice as shown in Figure A 5b. After digital flattening, the 3D cyclorama stack enables examination of individual layers at incremental depths of the pharmaceutical film. Figure A 5b shows a deposition defect (gap in drug distribution) on a 2D cyclorama emerging out of a photorealistic rendering of the corresponding 3D cyclorama. This case study shows how the 3D cyclorama method can be employed to undo deformations in 3D image stacks, not necessarily portraying biological samples, and to qualitatively characterise structures at different layers of the sample.

A.2.2 Virtual unrolling of an ancient lead amulet

This case study shows how the 3D cyclorama method may be useful to non-destructively retrieve scripture from ancient artefacts (including lead amulets, papyri, clay envelopes and clay tablets). Raymond *et al.* [313] gained access to one of the rare lead amulets (MU4106) that remains intact to date (not physically unrolled), which resides in the collection of the museum of ancient cultures (MAC) at MACQUARIE University, Sydney, Australia. In many ancient cultures, scrolls were buried or thrown into cisterns and wells, resulting in artefacts that are preserved in very poor condition, becoming corroded and brittle [313]. To this end, Raymond *et al.* employed non-destructive neutron imaging to reveal the ancient scripture without physically unrolling the sample.

The tightly coiled scroll was scanned with the *Dingo* neutron imaging instrument at the Australian Centre for Neutron Scattering, Australian Nuclear Science and Technology Organisation (ANSTO). The high resolution scan with a voxel size of 13 µm reveals that the amulet is composed of four interweaved sheets of lead with text inscribed in their inner surface (Figure A 6a&b). After imaging, the 3D cyclorama method initially developed to unroll murine colons, was employed in a collaborative effort to virtually unroll, decipher, and translate ancient inscribed lead scrolls. As the sheet is not in the form of a deformed tube, a workaround was devised that considers the samples as several concentric deformed tubes that could be unrolled onto 3D cycloramas.

A scaled-down version (using the ImageJ scale tool, with 0.5 × 0.5 × 0.5 scaling factors for x, y and z, respectively and bilinear interpolation) of the neutron CT stack of 1921 slices with dimensions of 850 × 850 pixels^2 and an isotropic voxel size of 26 µm was unrolled using the linear mapping

mode with $C = 30$ depth levels, $I = 100$ point-long contours, an electric field search window of 21 points ($w = 10$), electric field line segment size $\delta = 5$, maximum number of steps $K_{max} = 5$, minimum electric field line variance $V_{min} = 10$, rigidity equal to 1, interpolation interval $m = 30$ and rotation axis origin $O = (434, 430)$. The resulting 3D cyclorama lengths were about (varying between the different 3D cycloramas) $L = 1300$ pixels, the height $H = 1921$ pixels, and the depth $C = 30$ slices. Figure A 6c shows the boundary and intermediate contours ($C = 3$ instead of the actual $C = 30$ depth levels shown here for visual clarity) of the outer sheet that resulted in the 3D cyclorama of Figure A 7.

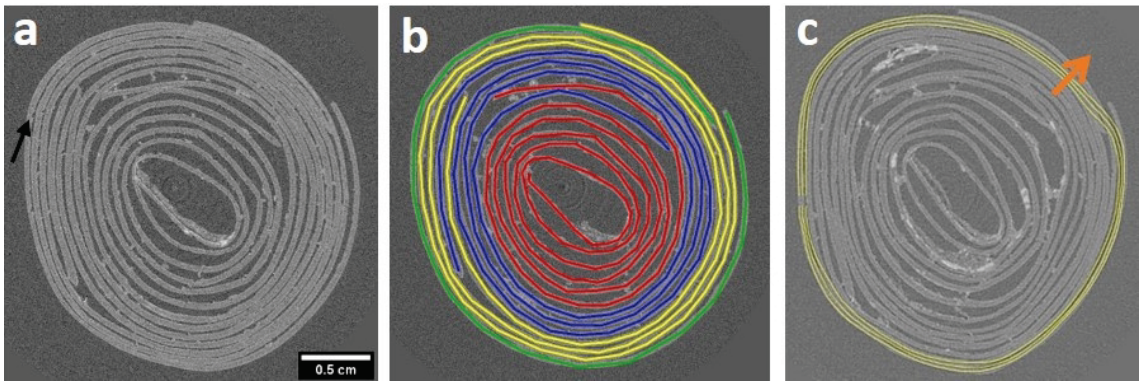


Figure A 6: Ancient scroll imaged by neutron CT

(a) A neutron CT slice of the lead amulet reveals that it is rolled in the form of a scroll. The scripture characters appear as tiny dents on the lead sheet (arrow). (b) The scroll consists of four separate interweaved sheets (red, blue, yellow, green). (c) Digital unrolling with the 3D cyclorama method was achieved by considering several concentric deformed tubes. Here, the boundary and intermediate contours ($C = 3$ instead of the actual $C = 30$ depth levels shown here for visual clarity) are shown for the outer sheet (green). The arrow shows the direction of increasing depth levels.

As the text is inscribed, it extends over a certain depth within the lead sheet. As a result, the retrieved text within the unrolled volume extends over several slices of the 3D cyclorama. Figure A 7 shows an example of a character retrieved by tracing its footprint through the 3D cyclorama slices. This is a prime example where a volumetric approach of scroll unrolling is greatly beneficial over approaches that deem the scroll as a single surface [236, 237, 239, 241-245]. By virtual unrolling all concentric tubes within the sheet, it was possible to retrieve the entire scripture. Scholars at Rutgers University and University of Cambridge identified the scripture language as Mandaic, an ancient dialect of Aramaic, and are in the process of translating the script for publication. The Mandaeans were a religious sect that emerged in the third century AD, and were adherents to the prophet, John the Baptist [313].

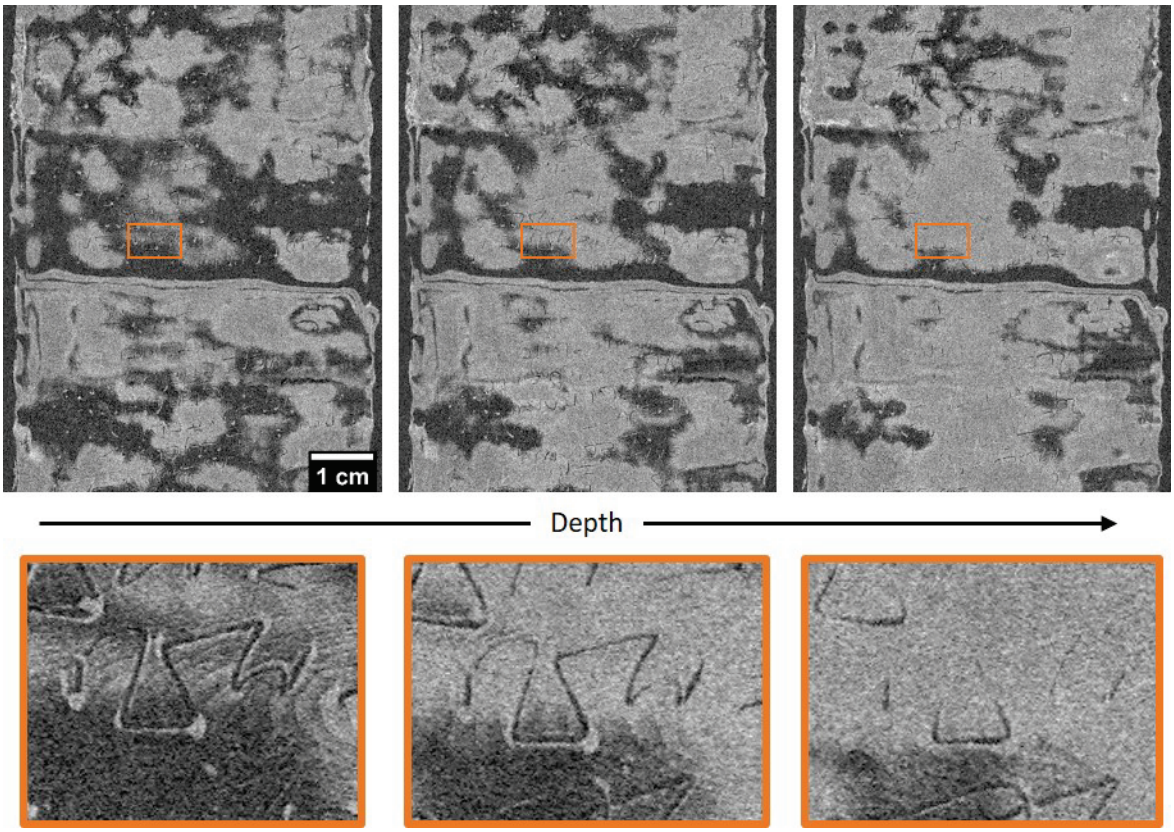


Figure A 7: 3D cycloramas to unroll an ancient scroll

At the top row, sequential 2D cycloramas with increasing depth levels from left to right reveal the unrolled outer sheet (green) of the scroll (see Figure A 6b&c). At the bottom row, enlarged versions at the corresponding orange boxes are shown. As characters are inscribed on the lead sheet, they extend over several slices of the 3D cyclorama. This is an example where a volumetric approach of scroll unrolling is beneficial over approaches that deem the scroll as a single surface.

This finding establishes the MU4106 lead amulet as a rare ancient artefact, since Mandaic scrolls are infrequently found intact. The two cutting-edge technologies of neutron 3D imaging and subsequent virtual unrolling pave the road for non-destructive studies of ancient artefacts, such as amulets or tablets enclosed in clay envelopes [309] that have long been locked in museum drawers for the sake of their preservation.

Collaboration to flatten clay envelopes and tablets using the 3D cyclorama method is currently ongoing.

Appendix B Global characterisation of the colonic structure

The majority of the work presented in this thesis amounts to the main objective of 3D quantification of the morphology of budding crypts during the early-stages of colorectal cancer. The crypt morphology is however distinctly different in the healthy colon, where budding is typically only observed in the form of (symmetric) fission. Therefore, an approach that enables the comparison between healthy and diseased colons needs to consider the global structure of the colon, and perform comparisons through macroscopic measures.

This section presents a comparison of the colonic morphology between the untreated and the AOM/DSS-treated groups of the animal study in Chapter 2 'Application of an animal model for colorectal cancer'. A list of the samples is shown in Table B 1. The quantification workflow is outlined in 'B.1 Workflow for global characterisation of the colonic structure', the comparison between the two groups is presented in 'B.2 Colonic structure morphometry in health and disease' and a discussion of the key findings is presented in appendix 'B.3 Key findings'.

Table B 1 Samples included in morphological quantification dataset

The colonic tissue samples come from 6 animals that were treated with AOM&DSS, and culled at the end of cycle 1 (AOM/DSS group), and 6 animals that were not treated with any agent (Untreated group). The mouse ID is composed of the cage that housed the animal and the animal number within the cage (see supplementary material S1). Scan ID is the specific identifier of each μCT scan performed at DLS, and cyclorama ID is the specific identifier of the unrolled volume that was used for the identification of the colonic crypts. The sample ID was given for convenience of presentation and it is composed of 'U' for the Untreated group or 'AD' for the AOM/DSS group, followed by an incremental number for each animal (biological replication).

Sample ID	Scan ID	Cyclorama ID	Mouse ID	Treatment group
U0	99037	201906121526	F4	Untreated
U1	99072	201906121828	F6	Untreated
U2	99086	201906121542	F2	Untreated
U3	99087	201906121818	F1	Untreated
U4	99090	201906141729	F5	Untreated

Sample ID	Scan ID	Cyclorama ID	Mouse ID	Treatment group
U5	99128	201906071753	F3	Untreated
AD0	99051	201906121731	A5	AOM/DSS
AD1	99056	201906121745	B6	AOM/DSS
AD2	99082	201906121758	C2	AOM/DSS
AD3	99094	201906121558	A1	AOM/DSS
AD4	99097	201906121717	A2	AOM/DSS
AD5	99115	201906121810	C4	AOM/DSS

B.1 Workflow for global characterisation of the colonic structure

Traditional methods employed to study the colonic structure, e.g. to measure the mucosal layer's thickness, typically involve a step of cutting the colon open or coiling it into a Swiss roll, (see section '1.2.1 Crypt formation and mucosal homeostasis'). In contrast, the sample preparation protocol employed here preserves the native shape of the colons. This avoids introduction of bias in the descriptive colonic measures as the colon no longer needs to be cut open or rolled, retaining the natural shape of the crypts. However, the natural shape of the colon (deformed tube of non-uniform thickness) makes quantification of the colonic structure particularly challenging. The colonic cross-sections depicted in µCT slices, imaged with synchrotron phase-contrast X-ray tomography (see section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons'), have wavy borders with a non-uniform colonic wall thickness (see Figure 5.2). The concept of a local colon thickness was introduced in Chapter 4, in order to capture the challenging geometry of the colon and to digitally unroll it. As the local thickness at a given point on the circumference portrays the thickness of the annulus within two boundary contours, it can be used to describe and estimate the mucosal and colonic wall thickness on each cross-section (see Figure 4.8b).

Specifically, the meshing step (section '4.2.1 Meshing: Generation of onion-like re-slicing surfaces') of the cyclorama method was employed in order to estimate the local thickness of the mucosal layer, the muscle layer and the colonic wall, as well as to estimate their cross-sectional area. This was applied on sub-stacks of 301 µCT slices (pixel size of 2.2 µm, corresponding 662 µm) of the µCT scans of section '3.3.2 Low-resolution synchrotron phase-contrast CT of murine colons'. The number of slices was set to 301 due to technical limitations (memory consumption for digital unrolling of 8 bit µCT stacks on a computer with 16 GB of RAM) as the size of the CT stacks was exceedingly large (1176 slices or 2.6 mm per CT image stack in total). The three contours delineating the interface between (a) mucosa and wax, (b) mucosa and muscle layer,

and (c) muscle layer and wax in Figure 5.2, were manually defined using the 'Segmented line selection tool' in Fiji to define region of interests. Computation of the electric field lines was performed with $I = 200$ point-long contours, an electric field search window of 31 points ($w = 15$), electric field line segment size $\delta = 2$, maximum number of steps $K_{max} = 50$, minimum electric field line variance $V_{min} = 50$, rigidity equal to 1 and interpolation interval $m = 20$. The mucosal cross-sectional area was derived (using the 'Measure tool' in Fiji) as the area of the annulus enclosed by contours (a) and (b) in Figure 5.2. Similarly, the muscle layer area was derived by selecting (b) and (c) as boundary contours, and the colonic wall layer area was derived by selecting (a) and (c) as boundary contours.

B.2 Colonic structure morphometry in health and disease

A macroscopic characterisation of the colonic structure was obtained by initially measuring the numbers of budding and non-budding crypts, as well as quantifying the thickness and cross-section area of the colonic tissue layers, which are summarised in Table B 2. The global morphological characteristics for the budding and non-budding crypts were measured and compared among the two treatment groups, summarised in Table B 3.

Figure B 1 shows the number of budding and non-budding crypts found within each 662 µm – long colonic section, and the resulting crypt fission index (CFI) for the two groups. A mean number of 2000 ± 100 (mean ± standard deviation) crypts were counted in the untreated group compared to 1400 ± 200 crypts in the AOM/DSS group (difference not statistically significant). Among those, a mean number of 0.5 ± 0.2 crypts were budding in the untreated group, compared to 46 ± 5 in the AOM/DSS group (difference statistically significant). These numbers amount to a CFI of 0.02 ± 0.01 % for the untreated group and 3.6 ± 0.5 % for the AOM/DSS group (difference statistically significant).

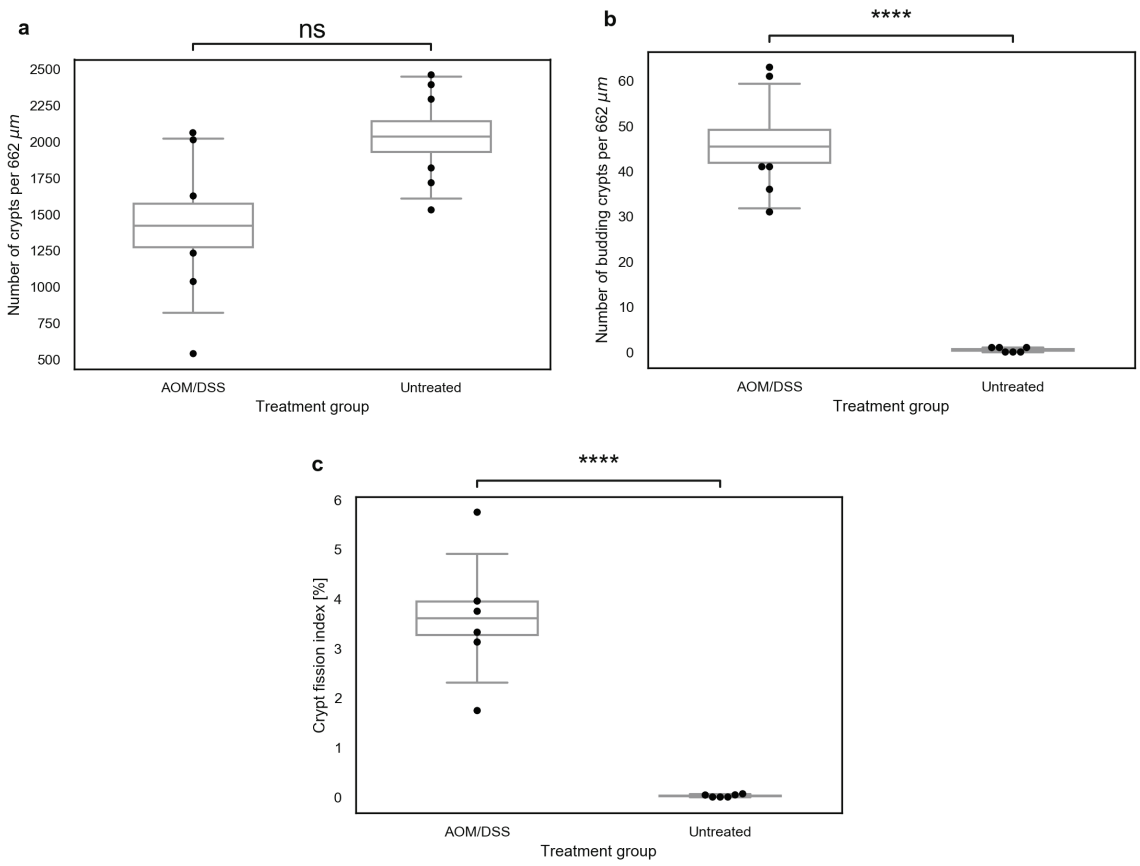


Figure B 1: Budding and non-budding crypt counts and crypt fission index

(a) The number of non-budding crypts counted in 662 µm – long colonic sections is equal to 2000 ± 100 (mean ± standard deviation) in the untreated group compared to 1400 ± 200 crypts in the AOM/DSS group (difference not statistically significant). (b) A mean of 0.5 ± 0.2 crypts were budding in the untreated group compared to 46 ± 5 in the AOM/DSS group (difference statistically significant). (c) These numbers amount to a CFI of 0.02 ± 0.01 % for the untreated group and 3.6 ± 0.5 % for the AOM/DSS group (difference statistically significant). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

Taken together, these measurements show that there is almost no budding detected among these 662 µm – long colonic sections in the untreated control group, which represents the healthy

case. The AOM/DSS group on the other hand, which represents the early-stage colorectal cancer case, shows a statistically significant increase in the number of budding crypts. An increased CFI has been reported as a precursor of ACF formation [13, 17-19], which are believed to be the earliest pre-cancerous lesions observed in the colon. The CFI decreases with age and it also varies with the colonic site that was sampled [6, 22]. In the healthy case, there is an exceedingly large drop of the CFI after the developmental phase from 8 weeks to 10 weeks, with Langlands *et al.* [22] reporting an average CFI in the colon of healthy mice of 23.8 % to 2.6 %, respectively. This yields a linear prediction of a virtually unmeasurable CFI (< 0) for mice of over 10.5 weeks old. In this respect, the estimated CFI for the untreated group (0.02 ± 0.01 %) agrees with Langlands *et al.* Tan *et al.* [6] proceeded to measure the CFI at 40 weeks of age and they found it equal to 0.7 %. Similarly *Bjerknes et al.* [63] have estimated a CFI of less than 1.7 % for mice older than 14 weeks. Therefore, the absolute value of the CFI appears to vary through the literature and thus, a direct comparison cannot be made. Instead, a statistically significant increase in the CFI of rodents treated with carcinogenic chemical agents (DMH) has been reported in the past [62], as well as between adenomas and control tissue in humans with FAP [19].

Figure B 2a,c,e show the cross-section thickness of the mucosal & muscle layer and the colonic wall, respectively, computed as the mean of the local thickness (length of electric field lines defined in section 4.2.1) measured at 200 points for each µCT slice. Figure B 2b,d,f show the cross-section area of the mucosal & muscle layer and the colonic wall, respectively. Together they quantify the morphology of the bulk of the colon and its composing tissue layers as a whole. The mucosal mean cross-section thickness was found equal to 170 ± 30 µm for the untreated group and 310 ± 10 µm for the AOM/DSS group (difference statistically significant). Similarly, the muscle mean cross-section thickness was found smaller for the untreated group at 55 ± 6 µm compared to 90 ± 6 µm of the AOM/DSS group (difference statistically significant). The colonic wall mean cross-section thickness was consequently found equal to 220 ± 30 µm and 410 ± 20 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). These show an increased thickness of the individual tissue layers and the entire colon as a whole in the AOM/DSS group compared to the untreated control, which is generally attributed to acute inflammatory response [212]. In humans, the increased thickness of the mucosal layer is a characteristic of colitis (condition that predisposes individuals to CRC) otherwise referred to, but not strictly defined as, crypt hyperplasia, mucosal hyperplasia, thickening of the colon wall, crypt elongation or epithelial hyperplasia [271]. In animals, a recent review of more than 90 IBD mouse models [115] revealed that none of them reported thickening of the muscle layer, which relates to thickening of the muscularis mucosae (colonic muscle layer) in human CRC. Here, the data-rich internal representation of the murine colon provided by 3D µCT imaging revealed for the first

time a statistically significant difference between the muscle thickness of the untreated and AOM/DSS groups in an AOM/DSS murine model.

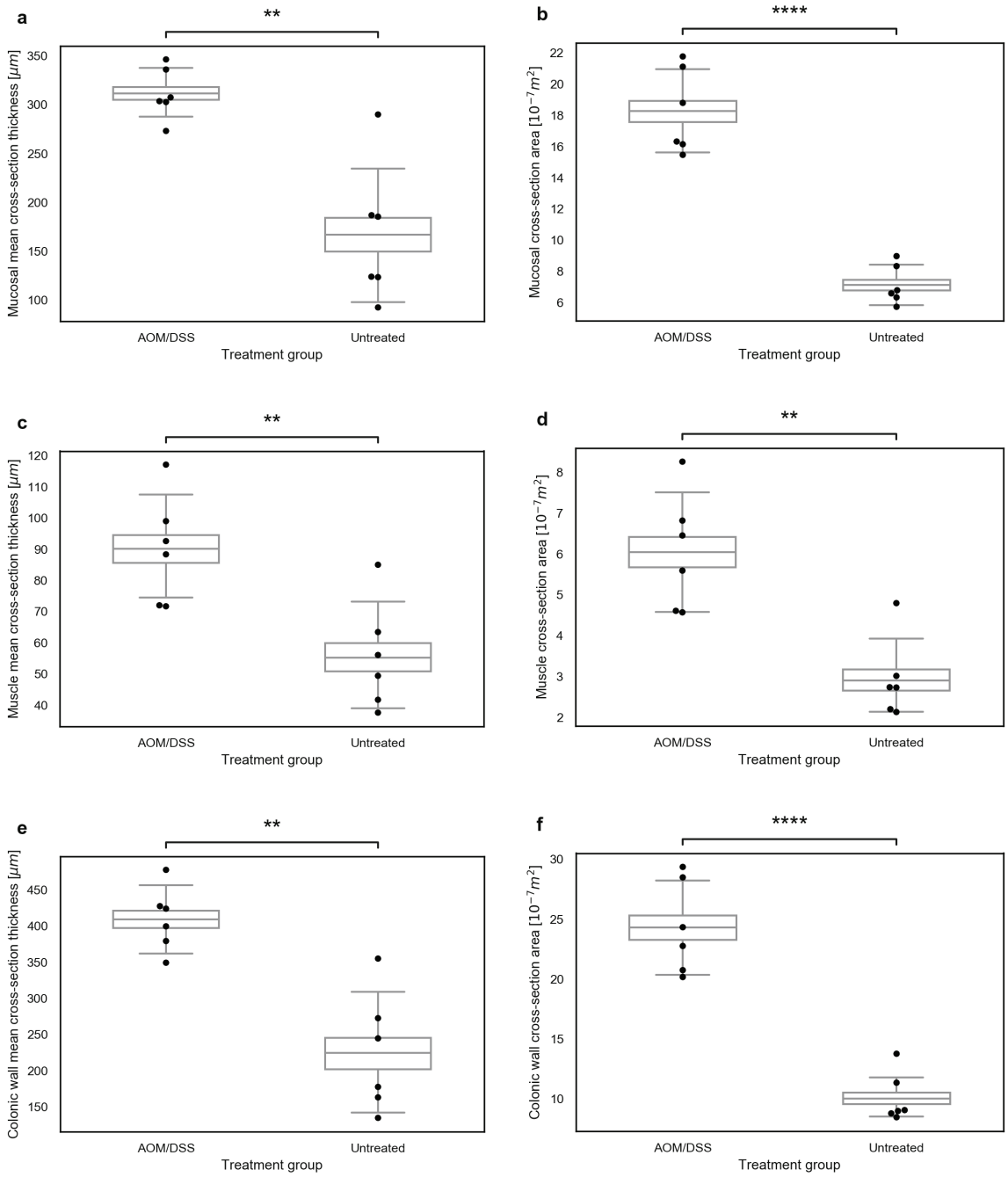


Figure B 2: Tissue layers' cross-section thickness and cross-section area measurements

The cross-section thickness of the mucosal & muscle layer and the colonic wall were computed as the mean of the local thickness (length of electric field lines defined in section 4.2.1) measured at 200 points for each µCT slice. (a) The mucosal mean cross-section thickness was found equal to 170 ± 30 µm for the untreated group and 310 ± 10 µm for the AOM/DSS group (difference statistically significant). (c) Similarly, the muscle mean cross-section thickness was found smaller for the untreated group

at 55 ± 6 μm compared to 90 ± 6 μm of the AOM/DSS group (difference statistically significant). (e) The colonic wall mean cross-section thickness was consequently found equal to 220 ± 30 μm and 410 ± 20 μm for the untreated and AOM/DSS group, respectively (difference statistically significant). The cross-section areas follow a similar pattern with the (b) mucosal, (d) muscle, and (f) colonic wall cross-section areas of the untreated group being equal to 7.1 ± 0.5 $\times 10^{-7}$ m^2 2.9 ± 0.4 $\times 10^{-7}$ m^2 and 10.1 ± 0.8 $\times 10^{-7}$ m^2, respectively, compared to 18 ± 1 $\times 10^5$ μm^2 6 ± 1 $\times 10^5$ μm^2 and 24 ± 1 $\times 10^5$ μm^2 for the AOM/DSS group (all differences are statistically significant). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

The cross-section areas follow a similar pattern with the mucosal, muscle and colonic wall cross-section areas of the untreated group being equal to 7.1 ± 0.5 $\times 10^{-7}$ m^2, 2.9 ± 0.4 $\times 10^{-7}$ m^2 and 10.1 ± 0.8 $\times 10^{-7}$ m^2, respectively, compared to 18 ± 1 $\times 10^5$ μm^2, 6 ± 1 $\times 10^5$ μm^2 and 24 ± 1 $\times 10^5$ μm^2 for the AOM/DSS group (all differences are statistically significant). The cross-section area measurements reveal the same underlying fact as the cross-section thickness measurements (presence of inflammation in the colon). However, the variability in the area measurements is better suited to reveal statistical difference with stricter confidence levels. This is readily visible in Figure B 2b,d,f where the statistical distributions of the untreated and AOM/DSS groups of the cross-section area measurements are better separated compared to those of the cross-section thickness measurements in Figure B 2a,c,e and the corresponding p-values in Table B 2 are smaller. Thus, the cross-section area might be a more appropriate measure to quantify the presence of inflammation in the different tissue layers than their thickness.

Next, the global morphological measures of the crypts are compared among the two groups and summarised in Table B 3. The crypt size is quantified by the crypt height, the crypt bounding box width and the crypt bounding box length shown in Figure B 3a,b,c, respectively. The crypt height was equal to 127 ± 9 μm for the untreated group and 210 ± 10 μm for the AOM/DSS group (difference statistically significant). Notice that, although the crypts extend across the mucosal layer, the crypt height in the two groups is smaller than the corresponding values of the mucosal thickness shown in Figure B 2a. This is because the crypt height underestimates the actual crypt

height (see section '5.2 Definition of budding and non-budding crypt characteristics'), which is better estimated by the mean leaf-to-root geodesic distance discussed later-on. The bounding box width was equal to 23 ± 1 µm and 44 ± 3 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). The bounding box length was equal to 30 ± 1 µm and 68 ± 5 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). Taken together, the size measurements show that the crypts in the AOM/DSS group tend to expand in all directions, not only increasing in height but also almost doubling in terms of bounding box width and length, sideways. This is purely due to the space taken by new branches that are spawned on the sides of the main crypt, as quantified by the leave angle in section 6.2.2.

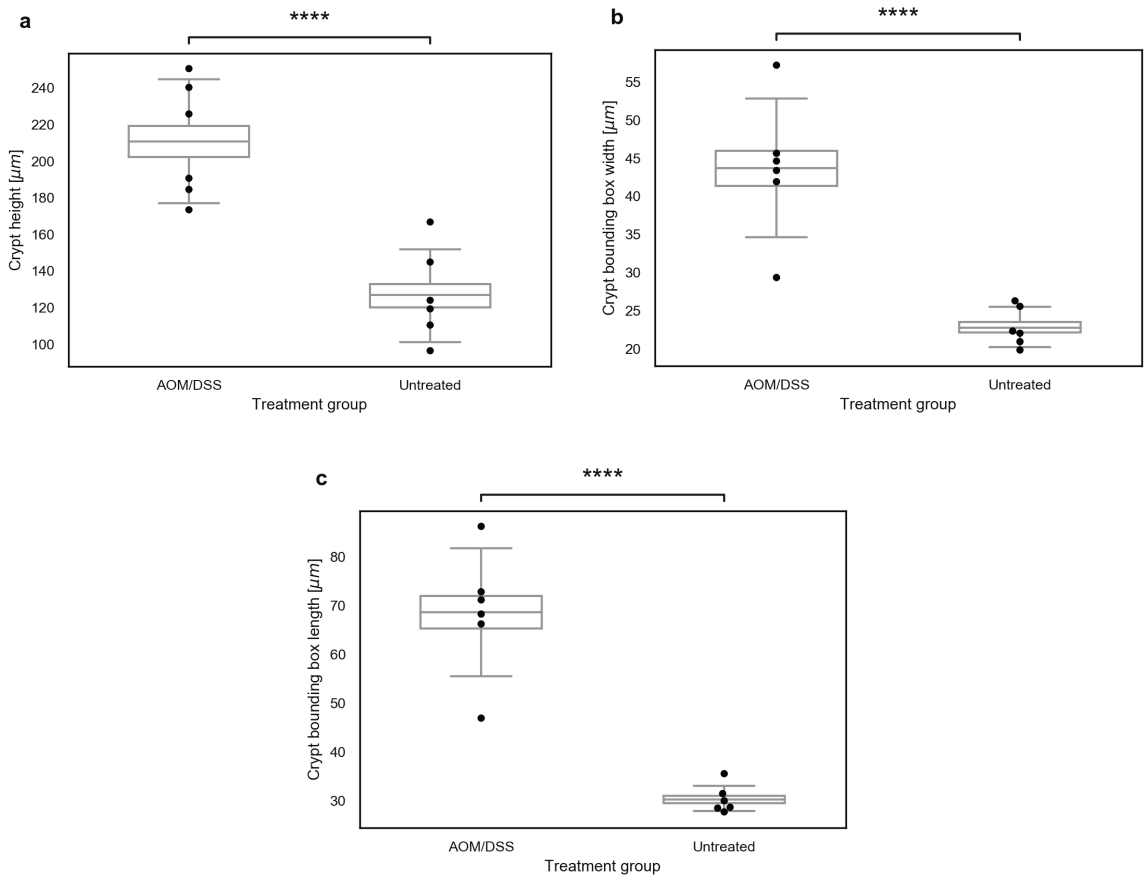


Figure B 3: Crypt height, bounding box width and length comparison between untreated and AOM/DSS group

(a) The crypt height is equal to 127 ± 9 µm for the untreated group and 210 ± 10 µm for the AOM/DSS group (difference statistically significant). (b) The bounding box width is equal to 23 ± 1 µm and 44 ± 3 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). (c) The bounding box length is equal to 30 ± 1 µm and 68 ± 5 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via

bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

A more detailed picture of the crypt lumen's size is created in Figure B 4, which shows the crypt total volume, crypt total surface area, maximum inscribed-ball radius and crypt mean breadth for the two groups. The crypt total volume is 16800 ± 700 μm^3 and 49000 ± 5000 μm^3 for the untreated and AOM/DSS group, respectively (difference statistically significant). The total surface area is 6000 ± 200 μm^2 and 19000 ± 2000 μm^2 for the untreated and AOM/DSS group, respectively (difference statistically significant). The radius of the maximum inscribed-ball is the same for the two groups (7 ± 0 μm, difference not statistically significant). Finally, the mean breadth is 87 ± 5 μm and 240 ± 20 μm for the untreated and AOM/DSS group, respectively (difference statistically significant). The total volume and total surface area measurements show that crypts in the AOM/DSS group are larger compared to those in the untreated group in terms of their lumen's volume and their epithelial layer. The fact that the maximum inscribed-ball radius is equal in the two groups however, shows that the crypt size in the AOM/DSS group is larger due to the increased number of branches (added volume and surface area of the spawned branches) rather than due to increment in the branch's width. This statement is also supported by the increased crypt mean breadth, and bounding box width and length. In addition to the added size of the new branches though, there is a size increment due to the increased mucosal thickness, which is suggested by the increased crypt height in the AOM/DSS group. This is examined further below, by looking at the lengths of individual branches.

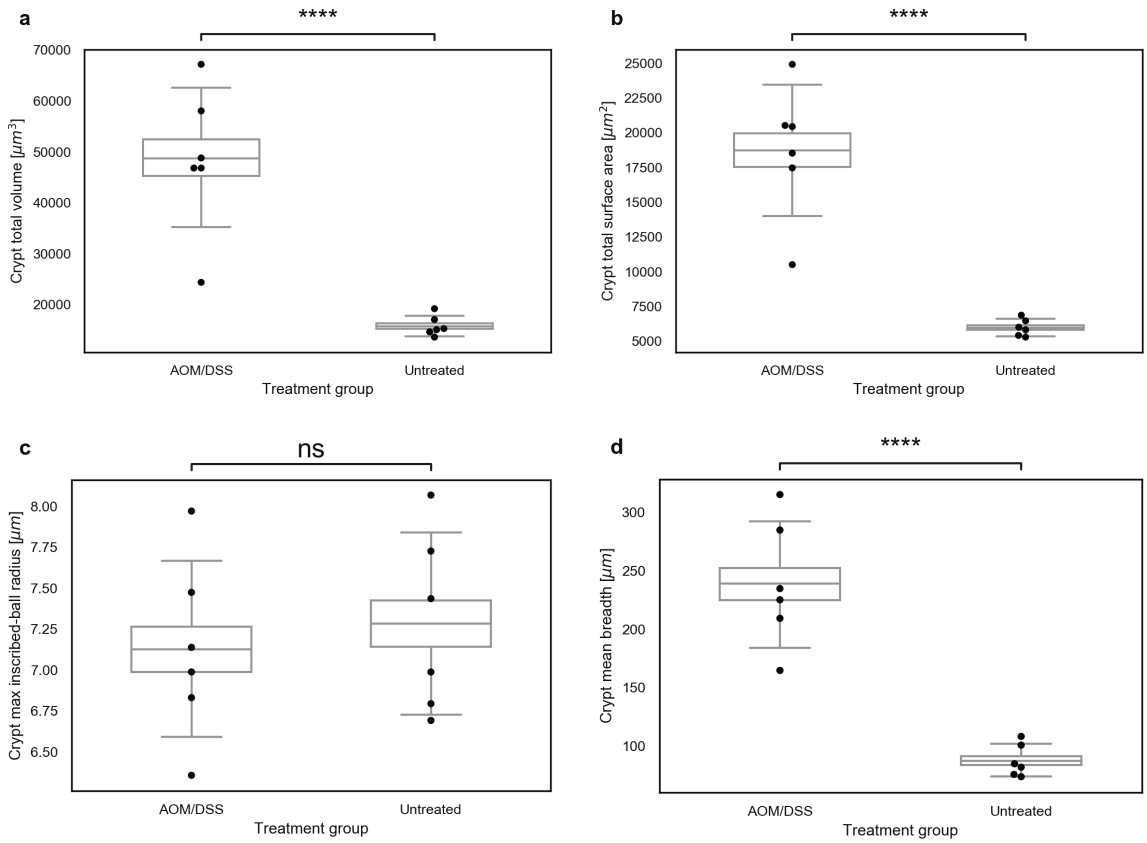


Figure B 4: Crypt total volume, surface area, maximum inscribed-ball radius and mean breadth comparison between untreated and AOM/DSS group

(a) The crypt total volume is 16800 ± 700 µm^3 and 49000 ± 5000 µm^3 for the untreated and AOM/DSS group, respectively (difference statistically significant). (b) The total surface area is 6000 ± 200 µm^2 and 19000 ± 2000 µm^2 for the untreated and AOM/DSS group, respectively (difference statistically significant). (c) The radius of the maximum inscribed-ball is the same for the two groups (7 ± 0 µm). (d) The mean breadth is 87 ± 5 µm and 240 ± 20 µm for the untreated and AOM/DSS group, respectively (difference statistically significant). The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

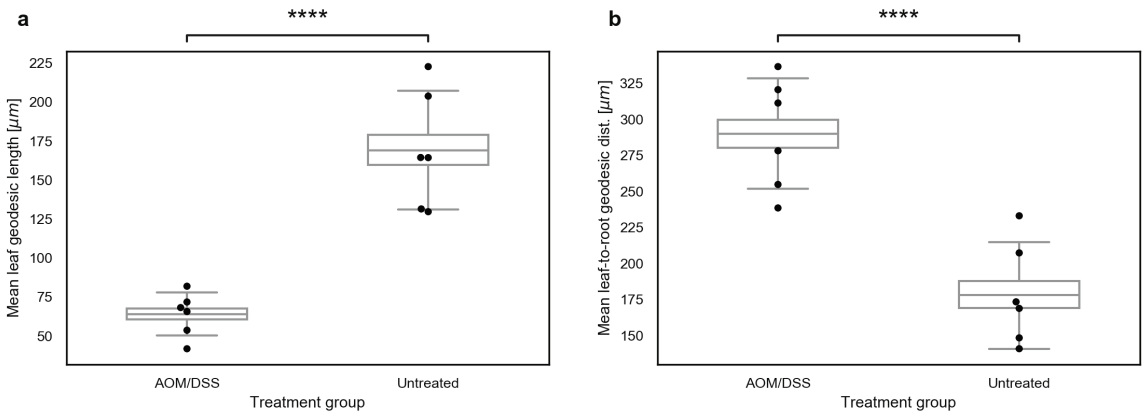


Figure B 5 Mean leaf geodesic length, leaf-to-root geodesic distance comparison between untreated and AOM/DSS group

(a) The mean leaf geodesic length is 170 ± 10 μm and 64 ± 5 μm for the untreated and AOM/DSS group, respectively (difference statistically significant). (b) The mean leaf-to-root geodesic distance which is equal to 180 ± 10 μm and 290 ± 10 μm for the untreated and AOM/DSS group, respectively (difference statistically significant). It quantifies the geodesic (actual) length of a crypt from end-to-end, i.e. from the bottoms of the leaf branches to the top of the crypt. The estimated means for each biological replication are shown as full disks (●). The sampling distributions, estimated via bootstrapping of the biological replication means are shown as box-and-whiskers plots (midline corresponds to the median value, lower and upper lines correspond to the first and third quartiles respectively and the whiskers lengths are equal to 1.5 times the inter-quartile range or IQR). The 'star' notation shows the statistical significance levels for statistical comparisons using a permutation test: 'ns' for $0.05 < p \leq 1$, '*' for $0.01 < p \leq 0.05$, '**' for $0.001 < p \leq 0.01$, '***' for $0.0001 < p \leq 0.001$, and '****' for $p \leq 0.0001$.

Figure B 5 focuses on individual branches in the untreated and AOM/DSS groups. As discussed in section 5.2.1, a crypt is composed of leaf and non-leaf branches that are distinguished upon the number of their children. Budding crypts begin with a root at the top (on the mucosal surface) upon which the trunk is attached, which splits into more branches. Those branches might split further several times until they spawn branches with no children, which are called leaf branches. The same naming convention applies on non-budding crypts, which means that the crypt is composed of a single branch that is concurrently the trunk and a leaf branch. Note that the AOM/DSS group is composed of strictly budding crypts whereas the untreated group is composed mainly of non-budding crypts. Figure B 5a,b quantify and compare the leaf branches of the untreated and AOM/DSS groups. The mean leaf geodesic length is 170 ± 10 μm and 64 ± 5 μm for

the untreated and AOM/DSS group, respectively. Note that these two are not directly comparable as the leaf geodesic length of non-budding crypts (most of the crypts in the untreated group) refers to the entire crypt, whereas for budding crypts (all of the crypts in the AOM/DSS group) it refers to the geodesic length of child branches. A more fair comparison is achieved by measuring the mean leaf-to-root geodesic distance which is equal to 180 ± 10 μm and 290 ± 10 μm for the untreated[9] and AOM/DSS group, respectively (difference statistically significant). It quantifies the geodesic (actual) length of a crypt from end-to-end, i.e. from the bottoms of the leaf branches to the top of the crypt. The mean leaf-to-root geodesic distance of the untreated group (180 ± 10 μm) approaches the mean crypt length found by Tan *et al.* [6] for the distal part of the colon (166 μm). Notably, the mean leaf-to-root geodesic distance of the untreated and AOM/DSS groups (180 ± 10 μm and 290 ± 10 μm) is directly comparable to their mucosal mean cross-section thickness (170 ± 30 μm and 310 ± 10). This suggests that the proposed method of quantifying the mucosal thickness using the electric field lines (see section 4.2.1) yields a more accurate estimation of the crypt's actual length than simply measuring the crypt's dimensions with a straight line (crypt height measure).

Table B 2 Tissue layer cross-section thickness and cross-section area

The measures for the tissue layers for the untreated and AOM/DSS group at the end of cycle 1 of the animal study detailed in section 2.1.1. The mean ± standard deviation of each measure, estimated following the process of section 5.3, is given for the two groups. The results of statistical comparison with a permutation test are also given in the form of the corresponding p-value. Statistical significance was evaluated as detailed in section 5.3. In brief, differences when tests yielded p-values greater than 0.05 were deemed not statistically significant.

Measure	Untreated	AOM/DSS	Permutation test p-value
Number of crypts per 662 μm	2000 ± 100	1400 ± 200	0.058
Number of budding crypts per 662 μm	0.5 ± 0.2	46 ± 5.0	0.000
Crypt fission index (CFI) [%]	0.02 ± 0.01	3.6 ± 0.5	0.000
Mucosal mean cross-section thickness [μm]	170 ± 30	310 ± 10	0.002
Muscle mean cross-section thickness [μm]	55 ± 6	90 ± 6	0.006

[9] The mean leaf geodesic length would be identical to the mean leaf-to-root geodesic distance for the untreated group if it was composed solely of non-budding crypts.

Measure	Untreated	AOM/DSS	Permutation test p-value
Colonic wall mean cross-section thickness [µm]	220 ± 30	410 ± 20	0.002
Mucosal cross-section area [10^{-7} m^2]	7.1 ± 0.5	18 ± 1	0.000
Muscle cross-section area [10^{-7} m^2]	2.9 ± 0.4	6 ± 1	0.004
Colonic wall cross-section area [10^{-7} m^2]	10.1 ± 0.8	24 ± 1	0.000

Table B 3 Global morphological characteristics comparison between untreated and AOM/DSS groups

The measures of the global morphological characteristics for the untreated and AOM/DSS group at the end of cycle 1 of the animal study detailed in section 2.1.1. The mean ± standard deviation of each measure, estimated following the process of section 5.3, is given for the two groups. The results of statistical comparison with a permutation test are given in the form of the resulting p-value. Statistical significance was evaluated as detailed in section 5.3. In brief, mean value differences when tests yielded p-values greater than 0.05 were deemed not statistically significant.

Measure	Untreated	AOM/DSS	Permutation test p-value
Volumetric			
Bounding box width [µm]	23 ± 1	44 ± 3	0.000
Bounding box length [µm]	30 ± 1	68 ± 5	0.000
Crypt height [µm]	127 ± 9	210 ± 10	0.000
Total volume [$µm^3$]	15800 ± 700	49000 ± 5000	0.000
Total surface area [$µm^2$]	6000 ± 200	19000 ± 2000	0.000
Max inscribed ball radius [µm]	7 ± 0	7 ± 0	0.630
Mean breadth [µm]	87 ± 5	240 ± 20	0.000
Mean leaf geodesic length [µm]	170 ± 10	64 ± 5	0.000
Mean leaf-to-root geodesic distance [µm]	180 ± 10	290 ± 10	0.000

B.3 Key findings

Digital unrolling using the 3D cyclorama method proved invaluable in three aspects: i. It was possible to measure the CFI quickly and efficiently, simply by counting crypts on a single digital image stack. This enabled an accurate estimation by counting crypts in an order of thousands in very little time. ii. Measurements of the local thickness and cross-section area proved that the latter is a better measure to differentiate between the untreated control and the AOM&DSS-treated group in terms of inflammation in the tissue. This better differentiation was achieved because the variability in the cross-section area was reduced compared to that of the local thickness, as the thickness of the healthy colon is naturally not constant. It would be interesting to examine how a coarse estimation of the cross-section area with more traditional methods (measuring a few distances on a histology cross-section) could compare to this finding. iii. An important last aspect was revealed when the mucosal local thickness was found in good agreement with the crypt leaf-to-root geodesic distance (crypt actual length). As the local thickness is curved (defined upon the electric field lines introduced in section 4.2.1), it is naturally larger than a straight line traditionally used to measure the mucosal thickness [115]. This shows that the newly introduced local thickness is an appropriate measure to estimate the crypt length by measuring the mucosal thickness, and it yields a better estimation compared to measuring the mucosal thickness with a straight line.

Bibliography

[1] D. Rhoten, “Interdisciplinary research: Trend or transition,” *Items and Issues,* vol. 5, pp. 6-11, 2004.

[2] L. M. Campbell, “Overcoming obstacles to interdisciplinary research,” *Conservation biology,* vol. 19, pp. 574-577, 2005.

[3] M. B. Thorsten, *Computed tomography: From photon statistics to modern cone-beam CT*: Springer, Verlag, Berlin, Heidelberg, 2008.

[4] J. Hsieh, *Computed Tomography - Principles, Design, Artifacts, and Recent Advances*, Third ed.: Society of Photo-Optical Instrumentation Engineers (SPIE), Bellingham, Washington, USA, 2015.

[5] L. De Chiffre, S. Carmignato, J. P. Kruth *et al.*, “Industrial applications of computed tomography,” *CIRP Annals,* vol. 63, pp. 655-677, 2014.

[6] C. W. Tan, Y. Hirokawa, B. S. Gardiner *et al.*, “Colon Cryptogenesis: Asymmetric Budding,” *Plos One,* vol. 8, pp. e78519, 2013.

[7] R. M. Clarke, “The effect of growth and of fasting on the number of villi and crypts in the small intestine of the albino rat,” *Journal of anatomy,* vol. 112, pp. 27, 1972.

[8] W. H. Stclair, and J. W. Osborne, “Crypt Fission and Crypt Number in the Small and Large Bowel of Postnatal Rats,” *Cell and Tissue Kinetics,* vol. 18, pp. 255-262, 1985.

[9] H. S. Wasan, H. S. Park, K. C. Liu *et al.*, “APC in the regulation of intestinal crypt fission,” *Journal of Pathology,* vol. 185, pp. 246-255, 1998.

[10] J. Totafurno, M. Bjerknes, and H. Cheng, “The Crypt Cycle - Crypt and Villus Production in the Adult Intestinal Epithelium,” *Biophysical Journal,* vol. 52, pp. 279-294, 1987.

[11] L. Bruens, S. I. J. Ellenbroek, J. van Rheenen *et al.*, “In Vivo Imaging Reveals Existence of Crypt Fission and Fusion in Adult Mouse Intestine,” *Gastroenterology,* vol. 153, pp. 674-677, 2017.

[12] A. M. Baker, C. Gabbutt, M. J. Williams *et al.*, “Crypt fusion as a homeostatic mechanism in the human colon,” *Gut,* vol. 68, pp. 1986-1993, 2019.

[13] D. Kristt, K. Bryan, and R. Gal, “Colonic aberrant crypts may originate from impaired fissioning: Relevance to increased risk of neoplasia,” *Human Pathology,* vol. 30, pp. 1449-1458, 1999.

[14] L. C. Greaves, S. L. Preston, P. J. Tadrous *et al.*, “Mitochondrial DNA mutations are established in human colonic stem cells, and mutated clones expand by crypt fission,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 103, pp. 714-719, 2006.

[15] S. L. Preston, W. M. Wong, A. O. O. Chan *et al.*, “Bottom-up histogenesis of colorectal adenomas: Origin in the monocryptal adenoma and initial expansion by crypt fission,” *Cancer Research,* vol. 63, pp. 3819-3825, 2003.

[16] A. G. Renehan, S. T. O'Dwyer, N. J. Haboubi *et al.*, “Early cellular events in colorectal carcinogenesis,” *Colorectal disease,* vol. 4, pp. 76-89, 2002.

[17] M. J. Wargovich, V. R. Brown, and J. Morris, “Aberrant crypt foci: the case for inclusion as a biomarker for colon cancer,” *Cancers (Basel),* vol. 2, pp. 1705-1716, 2010.

[18] M. Lopez-Ceron, and M. Pellise, “Biology and diagnosis of aberrant crypt foci,” *Colorectal Disease,* vol. 14, pp. E157-E164, 2012.

[19] W. M. Wong, N. Mandir, R. A. Goodlad *et al.*, “Histogenesis of human colorectal adenomas and hyperplastic polyps: the role of cell proliferation and crypt fission,” *Gut,* vol. 50, pp. 212-217, 2002.

[20] K. Araki, T. Ogata, M. Kobayashi *et al.*, “A Morphological-Study on the Histogenesis of Human Colorectal Hyperplastic Polyps,” *Gastroenterology,* vol. 109, pp. 1468-1474, 1995.

[21] M. Brittan, and N. A. Wright, “Stem cell in gastrointestinal structure and neoplastic development,” *Gut,* vol. 53, pp. 899-910, 2004.

[22] A. J. Langlands, A. A. Almet, P. L. Appleton *et al.*, “Paneth Cell-Rich Regions Separated by a Cluster of Lgr5+Cells Initiate Crypt Fission in the Intestinal Stem Cell Niche,” *Plos Biology,* vol. 14, pp. e1002491, 2016.

[23] A. A. Almet, B. D. Hughes, K. A. Landman *et al.*, “A Multicellular Model of Intestinal Crypt Buckling and Fission,” *Bulletin of Mathematical Biology,* vol. 80, pp. 335-359, 2018.

[24] B. M. Boman, and J. Z. Fields, “An APC:WNT Counter-Current-Like Mechanism Regulates Cell Division Along the Human Colonic Crypt Axis: A Mechanism That Explains How APC Mutations Induce Proliferative Abnormalities That Drive Colon Cancer Development,” *Front Oncol,* vol. 3, pp. 244, 2013.

[25] M. Bjerknes, “The Crypt Cycle and the Asymptotic Dynamics of the Proportion of Differently Sized Mutant Crypt Clones in the Mouse Intestine,” *Proceedings of the Royal Society B-Biological Sciences,* vol. 260, pp. 1-6, 1995.

[26] M. Bjerknes, “Expansion of mutant stem cell populations in the human colon,” *Journal of Theoretical Biology,* vol. 178, pp. 381-385, 1996.

[27] A. A. Almet, P. K. Maini, D. E. Moulton *et al.*, “Modeling perspectives on the intestinal crypt, a canonical system for growth, mechanics, and remodeling,” *Current Opinion in Biomedical Engineering,* vol. 15, pp. 32-39, 2020.

[28] Globocan 2012, "Epidemiology of colorectal cancer worldwide", Accessed on: 17 February 2021 [Online], Available: http://globocan.iarc.fr/Pages/fact_sheets_cancer.aspx.

[29] Cancer research UK, "Colorectal cancer causes and risks", Accessed on: 17 February 2021 [Online], Available: http://about-cancer.cancerresearchuk.org/about-cancer/bowel-cancer/risks-causes.

[30] V. Stigliano, L. Sanchez-Mete, A. Martayan *et al.*, “Early-onset colorectal cancer: A sporadic or inherited disease?,” *World Journal of Gastroenterology (WJG),* vol. 20, pp. 12420-12430, 2014.

[31] A. K. Gupta, T. P. Pretlow, and R. E. Schoen, “Aberrant crypt foci: What we know and what we need to know,” *Clinical Gastroenterology and Hepatology,* vol. 5, pp. 526-533, 2007.

[32] C. S. Potten, M. Kellett, S. A. Roberts *et al.*, “Measurement of Invivo Proliferation in Human Colorectal Mucosa Using Bromodeoxyuridine,” *Gut,* vol. 33, pp. 71-78, 1992.

[33] S. A. Roberts, J. H. Hendry, and C. S. Potten, “Deduction of the Clonogen Content of Intestinal Crypts - a Direct Comparison of 2-Dose and Multiple-Dose Methodologies,” *Radiation Research,* vol. 141, pp. 303-308, 1995.

[34] M. P. de Leon, and C. di Gregorio, “Pathology of colorectal cancer,” *Digestive and Liver Disease,* vol. 33, pp. 372-388, 2001.

[35] E. C. Klatt, *Robbins and Cotran atlas of pathology*: Elsevier Health Sciences, Philadelphia, 2014.

[36] F. Hugenholtz, and W. M. de Vos, “Mouse models for human intestinal microbiota research: a critical evaluation,” *Cell Mol Life Sci,* vol. 75, pp. 149-160, 2018.

[37] C. L. Scudamore, *A Practical Guide to the Histology of the Mouse*: John Wiley & Sons, Chichester, 2014.

[38] T. Nolte, P. Brander-Weber, C. Dangler *et al.*, “Nonproliferative and Proliferative Lesions of the Gastrointestinal Tract, Pancreas and Salivary Glands of the Rat and Mouse,” *J Toxicol Pathol,* vol. 29, pp. 1S-125S, 2016.

[39] I. M. M. van Leeuwen, H. M. Byrne, O. E. Jensen *et al.*, “Crypt dynamics and colorectal cancer: advances in mathematical modelling,” *Cell Proliferation,* vol. 39, pp. 157-181, 2006.

[40] C. S. Potten, and M. Loeffler, “Stem-Cells - Attributes, Cycles, Spirals, Pitfalls and Uncertainties - Lessons for and from the Crypt,” *Development,* vol. 110, pp. 1001-1020, 1990.

[41] F. A. Meineke, C. S. Potten, and M. Loeffler, “Cell migration and organization in the intestinal crypt using a lattice-free model,” *Cell Proliferation,* vol. 34, pp. 253-266, 2001.

[42] R. H. Stephens, J. Tanianis-Hughes, N. B. Higgs *et al.*, “Region-dependent modulation of intestinal permeability by drug efflux transporters: In vitro studies in mdr1a(-/-) mouse intestine,” *Journal of Pharmacology and Experimental Therapeutics,* vol. 303, pp. 1095-1101, 2002.

[43] H. Cheng, and C. P. Leblond, “Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine V. Unitarian theory of the origin of the four epithelial cell types,” *American Journal of Anatomy,* vol. 141, pp. 537-561, 1974.

[44] C. S. Potten, R. Schofield, and L. G. Lajtha, “A comparison of cell replacement in bone marrow, testis and three regions of surface epithelium,” *Biochim Biophys Acta,* vol. 560, pp. 281-299, 1979.

[45] M. Bjerknes, and H. Cheng, “The stem-cell zone of the small intestinal epithelium. III. Evidence from columnar, enteroendocrine, and mucous cells in the adult mouse,” *Am J Anat,* vol. 160, pp. 77-91, 1981.

[46] S. M. Karam, and C. P. Leblond, “Dynamics of epithelial cells in the corpus of the mouse stomach. I. Identification of proliferative cell types and pinpointing of the stem cell,” *Anat Rec,* vol. 236, pp. 259-279, 1993.

[47] C. S. Potten, C. Booth, and D. M. Pritchard, “The intestinal epithelial stem cell: the mucosal governor,” *International Journal of Experimental Pathology,* vol. 78, pp. 219-243, 1997.

[48] H. Cheng, and C. P. Leblond, “Origin, differentiation and renewal of the four main epithelial cell types in the mouse small intestine I. Columnar cell,” *American Journal of Anatomy,* vol. 141, pp. 461-479, 1974.

[49] N. Barker, J. H. van Es, J. Kuipers *et al.*, “Identification of stem cells in small intestine and colon by marker gene Lgr5,” *Nature,* vol. 449, pp. 1003-1007, 2007.

[50] R. May, S. M. Sureban, N. Hoang *et al.*, “Doublecortin and CaM Kinase-like-1 and Leucine-Rich-Repeat-Containing G-Protein-Coupled Receptor Mark Quiescent and Cycling Intestinal Stem Cells, Respectively,” *Stem Cells,* vol. 27, pp. 2571-2579, 2009.

[51] J. M. de Pereda, M. P. Lillo, and A. Sonnenberg, “Structural basis of the interaction between integrin alpha 6 beta 4 and plectin at the hemidesmosomes,” *Embo Journal,* vol. 28, pp. 1180-1190, 2009.

[52] N. A. Wright, “Epithelial stem cell repertoire in the gut: clues to the origin of cell lineages, proliferative units and cancer,” *International Journal of Experimental Pathology,* vol. 81, pp. 117-143, 2000.

[53] H. Cheng, and M. Bjerknes, “Whole Population Cell-Kinetics and Postnatal-Development of the Mouse Intestinal Epithelium,” *Anatomical Record,* vol. 211, pp. 420-426, 1985.

[54] A. P. Maskens, and R. M. Dujardinloits, “Kinetics of Tissue Proliferation in Colorectal Mucosa during Postnatal-Growth,” *Cell and Tissue Kinetics,* vol. 14, pp. 467-477, 1981.

[55] A. Cairnie, and B. Millen, “Fission of crypts in the small intestine of the irradiated mouse,” *Cell Proliferation,* vol. 8, pp. 189-196, 1975.

[56] N. A. Wright, and A. Alnafussi, “The Kinetics of Villus Cell-Populations in the Mouse Small-Intestine .2. Studies on Growth-Control after Death of Proliferative Cells Induced by Cytosine-Arabinoside, with Special Reference to Negative Feedback Mechanisms,” *Cell and Tissue Kinetics,* vol. 15, pp. 611-621, 1982.

[57] H. Lodish, A. Berk, C. A. Kaiser *et al.*, *Molecular Cell Biology*, 7 ed.: W. H. Freeman and Company, New York, 2013.

[58] C. M. Caldwell, and K. B. Kaplan, “The Role of APC in Mitosis and in Chromosome Instability,” *Apc Proteins,* vol. 656, pp. 51-64, 2009.

[59] S. Bahmanyar, W. J. Nelson, and A. I. M. Barth, “Role of APC and Its Binding Partners in Regulating Microtubules in Mitosis,” *Apc Proteins,* vol. 656, pp. 65-74, 2009.

[60] S. Nakamura, I. Kino, and S. Baba, “Cell-Kinetics Analysis of Background Colonic Mucosa of Patients with Intestinal Neoplasms by Exvivo Autoradiography,” *Gut,* vol. 29, pp. 997-1002, 1988.

[61] S. Nakamura, I. Kino, and S. Baba, “Nuclear-DNA Content of Isolated Crypts of Background Colonic Mucosa from Patients with Familial Adenomatous Polyposis and Sporadic Colorectal-Cancer,” *Gut,* vol. 34, pp. 1240-1244, 1993.

[62] H. S. Park, R. A. Goodlad, D. J. Ahnen *et al.*, “Effects of epidermal growth factor and dimethylhydrazine on crypt size, cell proliferation, and crypt fission in the rat colon. Cell proliferation and crypt fission are controlled independently,” *The American journal of pathology,* vol. 151, pp. 843-852, 1997.

[63] M. Bjerknes, H. Cheng, K. Hay *et al.*, “APC mutation and the crypt cycle in murine and human intestine,” *American Journal of Pathology,* vol. 150, pp. 833-839, 1997.

[64] H. Gehart, and H. Clevers, “Tales from the crypt: new insights into intestinal stem cells,” *Nature Reviews Gastroenterology & Hepatology,* vol. 16, pp. 19-34, 2019.

[65] M. Spit, B. K. Koo, and M. M. Maurice, “Tales from the crypt: intestinal niche signals in tissue renewal, plasticity and cancer,” *Open Biology,* vol. 8, pp. 180120, 2018.

[66] A. Hirata, J. Utikal, S. Yamashita *et al.*, “Dose-dependent roles for canonical Wnt signalling in de novo crypt formation and cell cycle properties of the colonic epithelium,” *Development,* vol. 140, pp. 66-75, 2013.

[67] C. Crosnier, D. Stamataki, and J. Lewis, “Organizing cell renewal in the intestine: stem cells, signals and combinatorial control,” *Nature Reviews Genetics,* vol. 7, pp. 349-359, 2006.

[68] M. Shtutman, J. Zhurinsky, I. Simcha *et al.*, “The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway,” *Proceedings of the National Academy of Sciences of the United States of America,* vol. 96, pp. 5522-5527, 1999.

[69] T. Zhang, T. Otevrel, Z. Q. Gao *et al.*, “Evidence that APC regulates survivin expression: A possible mechanism contributing to the stem cell origin of colon cancer,” *Cancer Research,* vol. 61, pp. 8664-8667, 2001.

[70] T. C. He, A. B. Sparks, C. Rago *et al.*, “Identification of c-MYC as a target of the APC pathway,” *Science,* vol. 281, pp. 1509-1512, 1998.

[71] E. Batlle, J. T. Henderson, H. Beghtel *et al.*, “beta-catenin and TCF mediate cell positioning in the intestinal epithelium by controlling the expression of EphB/EphrinB,” *Cell,* vol. 111, pp. 251-263, 2002.

[72] J. Kennell, and K. M. Cadigan, “APC and beta-Catenin Degradation,” *Apc Proteins,* vol. 656, pp. 1-12, 2009.

[73] P. Polakis, “Wnt signaling and cancer,” *Genes & development,* vol. 14, pp. 1837-1851, 2000.

[74] T. F. R. Homfray, S. E. Cottrell, M. Ilyas *et al.*, “Defects in mismatch repair occur after APC mutations in the pathogenesis of sporadic colorectal tumours,” *Human Mutation,* vol. 11, pp. 114-120, 1998.

[75] M. Bienz, and H. Clevers, “Linking colorectal cancer to Wnt signaling,” *Cell,* vol. 103, pp. 311-320, 2000.

[76] H. Suzuki, D. N. Watkins, K. W. Jair *et al.*, “Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer,” *Nature Genetics,* vol. 36, pp. 417-422, 2004.

[77] M. F. Kielman, M. Rindapaa, C. Gaspar *et al.*, “Apc modulates embryonic stem-cell differentiation by controlling the dosage of beta-catenin signaling,” *Nature Genetics,* vol. 32, pp. 594-605, 2002.

[78] R. H. Giles, J. H. van Es, and H. Clevers, “Caught up in a Wnt storm: Wnt signaling in cancer,” *Biochimica Et Biophysica Acta-Reviews on Cancer,* vol. 1653, pp. 1-24, 2003.

[79] M. Ilyas, “Wnt signalling and the mechanistic basis of tumour development,” *Journal of Pathology,* vol. 205, pp. 130-144, 2005.

[80] C. Gaspar, and R. Fodde, “APC dosage effects in tumorigenesis and stem cell differentiation,” *International Journal of Developmental Biology,* vol. 48, pp. 377-386, 2004.

[81] T. Zhang, J. Z. Fields, L. Opdenaker *et al.*, “Survivin-Induced Aurora-B Kinase Activation A Mechanism by Which APC Mutations Contribute to Increased Mitoses during Colon Cancer Development,” *American Journal of Pathology,* vol. 177, pp. 2816-2826, 2010.

[82] E. Hopmann, W. Arlt, and M. Minceva, “Solvent system selection in counter-current chromatography using conductor-like screening model for real solvents,” *Journal of Chromatography A,* vol. 1218, pp. 242-250, 2011.

[83] M. Brittan, and N. A. Wright, “Gastrointestinal stem cells,” *Journal of Pathology,* vol. 197, pp. 492-509, 2002.

[84] S. B. Garcia, H. S. Park, M. Novelli *et al.*, “Field cancerization clonality, and epithelial stem cells: The spread of mutated clones in epithelial sheets,” *Journal of Pathology,* vol. 187, pp. 61-81, 1999.

[85] H. S. Park, R. A. Goodlad, and N. A. Wright, “Crypt Fission in the Small-Intestine and Colon - a Mechanism for the Emergence of G6pd Locus-Mutated Crypts after Treatment with Mutagens,” *American Journal of Pathology,* vol. 147, pp. 1416-1427, 1995.

[86] E. A. Mclellan, A. Medline, and R. P. Bird, “Sequential-Analyses of the Growth and Morphological-Characteristics of Aberrant Crypt Foci - Putative Preneoplastic Lesions,” *Cancer Research,* vol. 51, pp. 5270-5274, 1991.

[87] C. M. Fenoglio, and N. Lane, “The anatomical precursor of colorectal carcinoma,” *Cancer,* vol. 34, pp. suppl:819-823, 1974.

[88] T. Muto, H. J. Bussey, and B. C. Morson, “The evolution of cancer of the colon and rectum,” *Cancer,* vol. 36, pp. 2251-2270, 1975.

[89] J. Madara, P. Harte, J. Deasy *et al.*, “Evidence for an adenoma-carcinoma sequence in dimethylhydrazine-induced neoplasms of rat intestinal epithelium,” *The American journal of pathology,* vol. 110, pp. 230, 1983.

[90] C. Pin, A. Parker, A. P. Gunning *et al.*, “An individual based computational model of intestinal crypt fission and its application to predicting unrestrictive growth of the intestinal epithelium,” *Integrative Biology,* vol. 7, pp. 213-228, 2015.

[91] M. de Robertis, E. Massi, M. L. Poeta *et al.*, “The AOM/DSS murine model for the study of colon carcinogenesis: From pathways to diagnosis and therapy studies,” *J Carcinog,* vol. 10, pp. 9, 2011.

[92] G. Joslyn, M. Carlson, A. Thliveris *et al.*, “Identification of deletion mutations and three new genes at the familial polyposis locus,” *Cell,* vol. 66, pp. 601-613, 1991.

[93] I. Nishisho, Y. Nakamura, Y. Miyoshi *et al.*, “Mutations of chromosome 5q21 genes in FAP and colorectal cancer patients,” *Science,* vol. 253, pp. 665-669, 1991.

[94] L. K. Su, K. W. Kinzler, B. Vogelstein *et al.*, “Multiple Intestinal Neoplasia Caused by a Mutation in the Murine Homolog of the Apc Gene,” *Science,* vol. 256, pp. 668-670, 1992.

[95] A. R. Moser, C. Luongo, K. A. Gould *et al.*, “Apc(Min) - a Mouse Model for Intestinal and Mammary Tumorigenesis,” *European Journal of Cancer,* vol. 31a, pp. 1061-1064, 1995.

[96] A. R. Shoemaker, C. Luongo, A. R. Moser *et al.*, "Somatic mutational mechanisms involved in intestinal tumor formation in Min mice," *Cancer Research,* vol. 57, pp. 1999-2006, 1997.

[97] D. E. Corpet, and F. Pierre, "Point: From animal models to prevention of colon cancer. Systematic review of chemoprevention in Min mice and choice of the model system," *Cancer Epidemiology Biomarkers & Prevention,* vol. 12, pp. 391-400, 2003.

[98] M. H. Wong, J. Huelsken, W. Birchmeier *et al.*, "Selection of multipotent stem cells during morphogenesis of small intestinal crypts of Lieberkuhn is perturbed by stimulation of Lef-1/beta-catenin signaling," *Journal of Biological Chemistry,* vol. 277, pp. 15843-15850, 2002.

[99] J. M. Ward, and P. M. Treuting, "Rodent intestinal epithelial carcinogenesis: pathology and preclinical models," *Toxicol Pathol,* vol. 42, pp. 148-61, 2014.

[100] Y. Fujimitsu, H. Nakanishi, K. Inada *et al.*, "Development of aberrant crypt foci involves a fission mechanism as revealed by isolation of aberrant crypts," *Japanese Journal of Cancer Research,* vol. 87, pp. 1199-1203, 1996.

[101] T. Tsukamoto, H. Sakai, A. Hirata *et al.*, "Three-dimensional analysis of isolated hexosaminidase-altered aberrant crypts from colons of 1,2-dimethylhydrazine-treated rats," *Experimental and Toxicologic Pathology,* vol. 57, pp. 283-289, 2006.

[102] T. Tanaka, H. Kohno, R. Suzuki *et al.*, "A novel inflammation-related mouse colon carcinogenesis model induced by azoxymethane and dextran sodium sulfate," *Cancer Science,* vol. 94, pp. 965-973, 2003.

[103] C. Neufert, C. Becker, and M. F. Neurath, "An inducible mouse model of colon carcinogenesis for the analysis of sporadic and inflammation-driven tumor progression," *Nat Protoc,* vol. 2, pp. 1998-2004, 2007.

[104] C. Becker, M. Fantini, C. Schramm *et al.*, "TGF-beta suppresses tumor progression in Th2 colitis and associated colon cancer by inhibition of IL-6 trans-signaling," *Gastroenterology,* vol. 126, pp. A79-A79, 2004.

[105] C. Becker, M. C. Fantini, S. Wirtz *et al.*, "In vivo imaging of colitis and colon cancer development in mice using high resolution chromoendoscopy," *Gut,* vol. 54, pp. 950-954, 2005.

[106] M. Kobaek-Larsen, I. Thorup, A. Diederichsen *et al.*, "Review of colorectal cancer and its metastases in rodent models: Comparative aspects with those in humans," *Comparative Medicine,* vol. 50, pp. 16-26, 2000.

[107] C. de Filippo, G. Caderni, M. Bazzicalupo *et al.*, "Mutations of the Apc gene in experimental colorectal carcinogenesis induced by azoxymethane in F344 rats," *British Journal of Cancer,* vol. 77, pp. 2148-2151, 1998.

[108] D. S. Levine, and R. C. Haggitt, "Normal Histology of the Colon," *American Journal of Surgical Pathology,* vol. 13, pp. 966-984, 1989.

[109] P. M. Gignac, N. J. Kley, J. A. Clarke *et al.*, "Diffusible iodine-based contrast-enhanced computed tomography (diceCT): an emerging tool for rapid, high-resolution, 3-D imaging of metazoan soft tissues," *J Anat,* vol. 228, pp. 889-909, 2016.

[110] A. I. Thaker, A. Shaker, M. S. Rao *et al.*, "Modeling Colitis-Associated Cancer with Azoxymethane (AOM) and Dextran Sulfate Sodium (DSS)," *Journal of Visualized Experiments (JoVE)*, pp. e4100, 2012.

[111] J. C. Booth, "The Theory and Practice of Histological Techniques," *Pathology,* vol. 15, pp. 118-119, 1983.

[112] M. Titford, "The long history of hematoxylin," *Biotechnic & Histochemistry,* vol. 80, pp. 73-78, 2005.

[113] Alberts B, Johnson A, Lewis J *et al.*, *Molecular biology of the cell*, 4th ed.: Garland Science, New York, 2002.

[114] T. S. Mair, C. E. Sherlock, D. Fews *et al.*, "Idiopathic Fibrosis of the Tunica Muscularis of the Large Intestine in Five Horses with Colic," *J Comp Pathol,* vol. 154, pp. 231-234, 2016.

[115] U. Erben, C. Loddenkennper, S. Spieckermann *et al.*, "Histomorphology of intestinal inflammation in inflammatory bowel diseases (IBD) mouse models and its relevance for IBD in men," *International Journal of Clinical and Experimental Pathology,* vol. 9, pp. 408-442, 2016.

[116] L. M. Shih, T. L. Wang, G. Traverso *et al.*, "Top-down morphogenesis of colorectal tumors," *Proceedings of the National Academy of Sciences of the United States of America,* vol. 98, pp. 2640-2645, 2001.

[117] R. A. Goodlad, S. Levi, C. Y. Lee *et al.*, "Morphometry and Cell-Proliferation in Endoscopic Biopsies - Evaluation of a Technique," *Gastroenterology,* vol. 101, pp. 1235-1241, 1991.

[118] R. H. Whitehead, A. Brown, and P. S. Bhathal, "A Method for the Isolation and Culture of Human Colonic Crypts in Collagen Gels," *In Vitro Cellular & Developmental Biology,* vol. 23, pp. 436-442, 1987.

[119] R. H. Whitehead, K. Demmler, S. P. Rockman *et al.*, "Clonogenic growth of epithelial cells from normal colonic mucosa from both mice and humans," *Gastroenterology,* vol. 117, pp. 858-865, 1999.

[120] T. Sato, R. G. Vries, H. J. Snippert *et al.*, "Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche," *Nature,* vol. 459, pp. 262-265, 2009.

[121] H. Furukawa, S. Asai, T. Moriyasu *et al.*, "Three‐Dimensional Reconstruction of the Mucosa from Sequential Sections of Biopsy Specimens for Understanding Disease Progress in Ulcerative Colitis," *Digestive Endoscopy,* vol. 12, pp. 317-321, 2000.

[122] H. Furukawa, F. Inoue, N. Miyake *et al.*, "Three‐dimensional reconstruction of the mucosa from sequential sections of biopsy specimens of patients with ulcerative colitis: Relationship between crypt structure and vascular architecture," *Digestive Endoscopy,* vol. 16, pp. 107-112, 2004.

[123] E. McInnes, "Artefacts in histopathology," *Comparative Clinical Pathology,* vol. 13, pp. 100-108, 2005.

[124] D. Serra, U. Mayr, A. Boni *et al.*, "Self-organization and symmetry breaking in intestinal organoid development," *Nature,* vol. 569, pp. 66-72, 2019.

[125] C. W. Tan, Y. Hirokawa, and A. W. Burgess, "Analysis of Wnt signalling dynamics during colon crypt development in 3D culture," *Scientific Reports,* vol. 5, pp. 1-18, 2015.

[126] T. Takayama, S. Katsuki, Y. Takahashi *et al.*, “Aberrant crypt foci of the colon as precursors of adenoma and cancer,” *New England Journal of Medicine,* vol. 339, pp. 1277-1284, 1998.

[127] A. L. Polglase, W. J. McLaren, S. A. Skinner *et al.*, “A fluorescence confocal endomicroscope for in vivo microscopy of the upper- and the lower-GI tract,” *Gastrointestinal Endoscopy,* vol. 62, pp. 686-695, 2005.

[128] R. Kiesslich, J. Burg, M. Vieth *et al.*, “Confocal laser endoscopy for diagnosing intraepithelial neoplasias and colorectal cancer in vivo,” *Gastroenterology,* vol. 127, pp. 706-713, 2004.

[129] X. Qi, Y. S. Pan, Z. L. Hu *et al.*, “Automated quantification of colonic crypt morphology using integrated microscopy and optical coherence tomography,” *Journal of Biomedical Optics,* vol. 13, pp. 054055, 2008.

[130] S. A. Yuan, C. A. Roney, J. Wierwille *et al.*, “Co-registered optical coherence tomography and fluorescence molecular imaging for simultaneous morphological and molecular imaging,” *Physics in Medicine and Biology,* vol. 55, pp. 191-206, 2010.

[131] A. M. Winkler, P. F. S. Rice, R. A. Drezek *et al.*, “Quantitative tool for rapid disease mapping using optical coherence tomography images of azoxymethane-treated mouse colon,” *Journal of Biomedical Optics,* vol. 15, pp. 041512, 2010.

[132] S. W. Huang, S. Y. Yang, W. C. Huang *et al.*, "Study on image feature extraction and classification for human colorectal cancer using optical coherence tomography," presented at the European Conference on Biomedical Optics, vol. 8091, 2011.

[133] Y. F. Zeng, B. Rao, W. C. Chapman *et al.*, “The Angular Spectrum of the Scattering Coefficient Map Reveals Subsurface Colorectal Cancer,” *Scientific Reports,* vol. 9, pp. 1-11, 2019.

[134] E. S. Choi, W. J. Choi, S. Y. Ryu *et al.*, “Feasibility of Full-field Optical Coherence Microscopy in Ultra-structural Imaging of Human Colon Tissues,” *Journal of the Korean Physical Society,* vol. 57, pp. 79-83, 2010.

[135] L. Ritsma, S. I. J. Ellenbroek, A. Zomer *et al.*, “Intestinal crypt homeostasis revealed at single-stem-cell level by in vivo live imaging,” *Nature,* vol. 507, pp. 362-365, 2014.

[136] L. Ritsma, E. J. A. Steller, S. I. J. Ellenbroek *et al.*, “Surgical implantation of an abdominal imaging window for intravital microscopy,” *Nature Protocols,* vol. 8, pp. 583-594, 2013.

[137] T. P. Pretlow, B. J. Barrow, W. S. Ashton *et al.*, “Aberrant Crypts - Putative Preneoplastic Foci in Human Colonic Mucosa,” *Cancer Research,* vol. 51, pp. 1564-1567, 1991.

[138] A. K. Gupta, P. Pinsky, C. Rall *et al.*, “Reliability and accuracy of the endoscopic appearance in the identification of aberrant crypt foci,” *Gastrointestinal Endoscopy,* vol. 70, pp. 322-330, 2009.

[139] R. E. Rudolph, J. A. Dominitz, J. W. Lampe *et al.*, “Risk factors for colorectal cancer in relation to number and size of aberrant crypt foci in humans,” *Cancer Epidemiology Biomarkers & Prevention,* vol. 14, pp. 605-608, 2005.

[140] L. Roncucci, D. Stamp, A. Medline *et al.*, “Identification and Quantification of Aberrant Crypt Foci and Microadenomas in the Human Colon,” *Human Pathology,* vol. 22, pp. 287-294, 1991.

[141] L. Roncucci, A. Medline, and W. R. Bruce, “Classification of Aberrant Crypt Foci and Microadenomas in Human Colon,” *Cancer Epidemiology Biomarkers & Prevention,* vol. 1, pp. 57-60, 1991.

[142] T. P. Pretlow, M. A. Oriordan, T. G. Pretlow *et al.*, “Aberrant Crypts in Human Colonic Mucosa - Putative Preneoplastic Lesions,” *Journal of Cellular Biochemistry*, pp. 55-62, 1992.

[143] D. G. Adler, C. J. Gostout, D. Sorbi *et al.*, “Endoscopic identification and quantification of aberrant crypt foci in the human colon,” *Gastrointestinal Endoscopy,* vol. 56, pp. 657-662, 2002.

[144] R. P. Bird, and C. K. Good, “The significance of aberrant crypt foci in understanding the pathogenesis of colon cancer,” *Toxicology Letters,* vol. 112, pp. 395-402, 2000.

[145] T. Yokota, K. Sugano, H. Kondo *et al.*, “Detection of aberrant crypt foci by magnifying colonoscopy,” *Gastrointestinal Endoscopy,* vol. 46, pp. 61-65, 1997.

[146] L. Roncucci, S. Modica, M. Pedroni *et al.*, “Aberrant crypt foci in patients with colorectal cancer,” *British Journal of Cancer,* vol. 77, pp. 2343-2348, 1998.

[147] R. G. Stevens, H. Swede, C. D. Heinen *et al.*, “Aberrant crypt foci in patients with a positive family history of sporadic colorectal cancer,” *Cancer Letters,* vol. 248, pp. 262-268, 2007.

[148] H. Y. Lee, J. O. Kim, J. Shim *et al.*, “Multivariate discriminant analysis for branching classification of colonic tubular adenoma glands,” *Cytometry Part B-Clinical Cytometry*, pp. 429-440, 2020.

[149] K. A. Williams, N. J. Gostling, J. W. Steer *et al.*, “Quantifying intracortical bone microstructure: A critical appraisal of 2D and 3D approaches for assessing vascular canals and osteocyte lacunae,” *J Anat,* vol. 238, pp. 653–668, 2020.

[150] I. M. M. van Leeuwen, G. R. Mirams, A. Walter *et al.*, “An integrative computational model for intestinal tissue renewal,” *Cell Proliferation,* vol. 42, pp. 617-636, 2009.

[151] A. K. Gupta, and R. E. Schoen, “Aberrant crypt foci: are they intermediate endpoints of colon carcinogenesis in humans?,” *Current Opinion in Gastroenterology,* vol. 25, pp. 59-65, 2009.

[152] A. J. Carulli, L. C. Samuelson, and S. Schnell, “Unraveling intestinal stem cell behavior with models of crypt dynamics,” *Integrative Biology,* vol. 6, pp. 243-257, 2014.

[153] C. M. Edwards, and S. J. Chapman, “Biomechanical modelling of colorectal crypt budding and fission,” *Bulletin of Mathematical Biology,* vol. 69, pp. 1927-1942, 2007.

[154] C. Pin, A. J. Watson, and S. R. Carding, “Modelling the spatio-temporal cell dynamics reveals novel insights on cell differentiation and proliferation in the small intestinal crypt,” *PLoS One,* vol. 7, pp. e37115, 2012.

[155] T. Ingham-Dempster, D. C. Walker, and B. M. Corfe, “An agent-based model of anoikis in the colon crypt displays novel emergent behaviour consistent with biological observations,” *Royal Society Open Science,* vol. 4, pp. 160858, 2017.

[156] D. Muraro, A. Parker, L. Vaux *et al.*, “Chronic TNFalpha-driven injury delays cell migration to villi in the intestinal epithelium,” *J R Soc Interface,* vol. 15, pp. 20180037, 2018.

[157] P. Buske, J. Galle, N. Barker *et al.*, "A Comprehensive Model of the Spatio-Temporal Stem Cell and Tissue Organisation in the Intestinal Crypt," *Plos Computational Biology,* vol. 7, pp. e1001045, 2011.

[158] S. J. Dunn, I. S. Nathke, and J. M. Osborne, "Computational models reveal a passive mechanism for cell migration in the crypt," *PLoS One,* vol. 8, pp. e80516, 2013.

[159] H. J. Du, Q. Nie, and W. R. Holmes, "The Interplay between Wnt Mediated Expansion and Negative Regulation of Growth Promotes Robust Intestinal Crypt Structure and Homeostasis," *Plos Computational Biology,* vol. 11, pp. e1004285, 2015.

[160] A. G. Fletcher, C. J. W. Breward, and S. J. Chapman, "Mathematical modeling of monoclonal conversion in the colonic crypt," *Journal of Theoretical Biology,* vol. 300, pp. 118-133, 2012.

[161] P. Buske, J. Przybilla, M. Loeffler *et al.*, "The intestinal stem cell niche: a computational tissue approach," *Biochemical Society Transactions,* vol. 42, pp. 671-677, 2014.

[162] T. Thalheim, P. Buske, J. Przybilla *et al.*, "Stem cell competition in the gut: insights from multi-scale computational modelling," *Journal of the Royal Society Interface,* vol. 13, pp. 20160218, 2016.

[163] R. H. Smallwood, W. M. L. Holcombe, and D. C. Walker, "Development and validation of computational models of cellular interaction," *Journal of Molecular Histology,* vol. 35, pp. 659-665, 2004.

[164] W. A. Kalender, *Computed Tomography: Fundamentals, System Technology, Image Quality, Applications*, 3rd ed.: Publicis, Erlangen, 2011.

[165] P. Willmott, *An Introduction to Synchrotron Radiation: Techniques and Applications*, First ed.: John Wiley & Sons Ltd., Chichester, 2011.

[166] P. Tafforeau, R. Boistel, E. Boller *et al.*, "Applications of X-ray synchrotron microtomography for non-destructive 3D studies of paleontological specimens," *Applied Physics a-Materials Science & Processing,* vol. 83, pp. 195-202, 2006.

[167] J. C. Elliott, F. S. L. Wong, P. Anderson *et al.*, "Determination of mineral concentration in dental enamel from X-ray attenuation measurements," *Connective Tissue Research,* vol. 38, pp. 61-72, 1998.

[168] S. Seltzer, "Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients, NIST Standard Reference Database 126", Accessed on: 17 February 2021 [Online], Available: http://www.nist.gov/pml/data/xraycoef/index.cfm.

[169] S. C. Mayo, A. W. Stevenson, and S. W. Wilkins, "In-Line Phase-Contrast X-ray Imaging and Tomography for Materials Science," *Materials (Basel),* vol. 5, pp. 937-965, 2012.

[170] M. J. Kitchen, G. A. Buckley, T. E. Gureyev *et al.*, "CT dose reduction factors in the thousands using X-ray phase contrast," *Scientific reports,* vol. 7, pp. 1-9, 2017.

[171] B. Zeller-Plumhoff, J. L. Mead, D. Tan *et al.*, "Soft tissue 3D imaging in the lab through optimised propagation-based phase contrast computed tomography," *Optics Express,* vol. 25, pp. 33451, 2017.

[172] F. Pfeiffer, T. Weitkamp, O. Bunk *et al.*, "Phase retrieval and differential phase-contrast imaging with low-brilliance X-ray sources," *Nature Physics,* vol. 2, pp. 258-261, 2006.

[173] S. C. Mayo, T. J. Davis, T. E. Gureyev *et al.*, "X-ray phase-contrast microscopy and microtomography," *Optics Express,* vol. 11, pp. 2289-2302, 2003.

[174] O. Hemberg, M. Otendal, and H. M. Hertz, "Liquid-metal-jet anode electron-impact x-ray source," *Applied Physics Letters,* vol. 83, pp. 1483-1485, 2003.

[175] E. N. Landis, and D. T. Keane, "X-ray microtomography," *Materials Characterization,* vol. 61, pp. 1305-1316, 2010.

[176] L. Q. Ren, M. U. Ghani, D. Wu *et al.*, "The impact of spectral filtration on image quality in micro-CT system," *Journal of Applied Clinical Medical Physics,* vol. 17, pp. 301-315, 2016.

[177] B. L. Henke, E. M. Gullikson, and J. C. Davis, "X-ray interactions: photoabsorption, scattering, transmission and reflection E= 50-30,000 eV, Z= 1-92," *Atomic data and nuclear data tables,* vol. 54, pp. 181-342, 1993.

[178] Center for X-Ray Optics, "Index of refraction calculator," Accessed on: 17 February 2021 [Online], Available: https://henke.lbl.gov/optical_constants/getdb2.html.

[179] M. A. Beltran, D. M. Paganin, K. Uesugi *et al.*, "2D and 3D X-ray phase retrieval of multi-material objects using a single defocus distance," *Optics Express,* vol. 18, pp. 6423-6436, 2010.

[180] A. Pogany, D. Gao, and S. W. Wilkins, "Contrast and resolution in imaging with a microfocus x-ray source," *Review of Scientific Instruments,* vol. 68, pp. 2774-2782, 1997.

[181] A. Sarapata, M. Willner, M. Walter *et al.*, "Quantitative imaging using high-energy X-ray phase-contrast CT with a 70 kVp polychromatic X-ray spectrum," *Optics Express,* vol. 23, pp. 523-535, 2015.

[182] R. Murty, "Effective atomic numbers of heterogeneous materials," *Nature,* vol. 207, pp. 398-399, 1965.

[183] J. H. Hubbell, W. J. Veigele, E. A. Briggs *et al.*, "Atomic form factors, incoherent scattering functions, and photon scattering cross sections," *Journal of Physical and Chemical Reference Data,* vol. 4, pp. 471-538, 1975.

[184] H. A. Haus, and J. R. Melcher, *Electromagnetic fields and energy*: Prentice Hall, New Jersey, 1989.

[185] B. Zeller-Plumhof, "An Image-Based Computation of Microvascular Dysfunction", PhD thesis, Engineering Sciences, Faculty of Engineering and the Environment, University of Southampton, 2017.

[186] J. Radon, "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten," *Akad. Wiss.,* vol. 69, pp. 262-277, 1917.

[187] J. Radon, "On the determination of functions from their integral values along certain manifolds," *IEEE Transactions on Medical Imaging,* vol. 5, pp. 170-176, 1986.

[188] C. K. Avinash, and M. Slaney, *Principles of computerized tomographic imaging*: IEEE press, New York, 1988.

[189] P. M. Bidola, I. Zanette, K. Achterhold *et al.*, "Optimization of propagation-based phase-contrast imaging at a laboratory setup," *Optics Express,* vol. 23, pp. 30000-30013, 2015.

[190] R. Coisson, "Spatial Coherence of Synchrotron-Radiation," *Applied Optics,* vol. 34, pp. 904-908, 1995.

[191] P. Cloetens, R. Barrett, J. Baruchel *et al.*, “Phase objects in synchrotron radiation hard x-ray imaging,” *Journal of Physics D-Applied Physics,* vol. 29, pp. 133-146, 1996.

[192] F. van der Veen, and F. Pfeiffer, “Coherent x-ray scattering,” *Journal of Physics-Condensed Matter,* vol. 16, pp. 5003-5030, 2004.

[193] S. A. Tsekenis, N. Tait, and H. McCann, “Spatially resolved and observer-free experimental quantification of spatial resolution in tomographic images,” *Review of Scientific Instruments,* vol. 86, pp. 035104, 2015.

[194] C. E. Shannon, “Communication in the presence of noise (Reprinted from the Proceedings of the IRE, vol 37, pg 10-21, 1949),” *Proceedings of the Ieee,* vol. 86, pp. 447-457, 1998.

[195] Pioneer in Cameras and Optoelectronics, "pco.edge 5.5 data sheet", Accessed on: 17 February 2021 [Online], Available: https://www.pco.de/fileadmin/user_upload/pco-product_sheets/pco.edge_55_data_sheet.pdf.

[196] A. V. Bronnikov, “Reconstruction formulas in phase-contrast tomography,” *Optics Communications,* vol. 171, pp. 239-244, 1999.

[197] A. Groso, R. Abela, and M. Stampanoni, “Implementation of a fast method for high resolution phase contrast tomography,” *Optics Express,* vol. 14, pp. 8103-8110, 2006.

[198] A. Groso, M. Stampanoni, R. Abela *et al.*, “Phase contrast tomography: An alternative approach,” *Applied Physics Letters,* vol. 88, pp. 214104, 2006.

[199] X. Wu, H. Liu, and A. Yan, “X-ray phase-attenuation duality and phase retrieval,” *Optics Letters,* vol. 30, pp. 379-381, 2005.

[200] D. Paganin, S. C. Mayo, T. E. Gureyev *et al.*, “Simultaneous phase and amplitude extraction from a single defocused image of a homogeneous object,” *Journal of Microscopy,* vol. 206, pp. 33-40, 2002.

[201] T. Weitkamp, D. Haas, D. Wegrzynek *et al.*, “ANKAphase: software for single-distance phase retrieval from inline X-ray phase-contrast radiographs,” *Journal of Synchrotron Radiation,* vol. 18, pp. 617-629, 2011.

[202] B. Yang, “Fresnel and refractive lenses for X-rays,” *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment,* vol. 328, pp. 578-587, 1993.

[203] Paul Scherrer Institut (PSI), "TOMCAT beamline - Swiss Light Source (SLS)", Accessed on: 17 February 2021 [Online], Available: https://www.psi.ch/en/sls/tomcat.

[204] Diamond Light Source (DLS), "I13-2 beamline ", Accessed on: 17 February 2021 [Online], Available: http://www.diamond.ac.uk/Beamlines/Materials/I13/Diamond-Manchester_Imaging_Branchline.html

[205] National Centre for the Replacement Refinement and Reduction of Animals in Research (NC3R), "ARRIVE guidelines", Accessed on: 17 February 2021 [Online], Available: https://www.nc3rs.org.uk/arrive-guidelines.

[206] National Centre for the Replacement Refinement and Reduction of Animals in Research (NC3R), "Replacement Refinement and Reduction (3R) principle", Accessed on: 17 February 2021 [Online], Available: https://www.nc3rs.org.uk/the-3rs.

[207] O. S. Sohn, E. S. Fiala, S. P. Requeijo *et al.*, "Differential effects of CYP2E1 status on the metabolic activation of the colon carcinogens azoxymethane and methylazoxymethanol," *Cancer Research,* vol. 61, pp. 8435-8440, 2001.

[208] I. Okayasu, S. Hatakeyama, M. Yamada *et al.*, "A Novel Method in the Induction of Reliable Experimental Acute and Chronic Ulcerative-Colitis in Mice," *Gastroenterology,* vol. 98, pp. 694-702, 1990.

[209] B. Parang, C. W. Barrett, and C. S. Williams, "AOM/DSS model of colitis-associated cancer," *Methods Mol Biol,* vol. 1422, pp. 297-307, 2016.

[210] C. G. Whittem, A. D. Williams, and C. S. Williams, "Murine Colitis modeling using Dextran Sulfate Sodium (DSS)," *J Vis Exp*, 2010.

[211] P. Ariel, "A beginner's guide to tissue clearing," *Int J Biochem Cell Biol,* vol. 84, pp. 35-39, 2017.

[212] X. Huo, D. Liu, L. Gao *et al.*, "Flavonoids Extracted from Licorice Prevents Colitis-Associated Carcinogenesis in AOM/DSS Mouse Model," *Int J Mol Sci,* vol. 17, pp. 1343, 2016.

[213] L. B. Meira, J. M. Bugni, S. L. Green *et al.*, "DNA damage induced by chronic inflammation contributes to colon carcinogenesis in mice," *Journal of Clinical Investigation,* vol. 118, pp. 2516-2525, 2008.

[214] S. Dolly, "TASMIP spectra calculator", Accessed on: 17 February 2021 [Online], Available: http://solutioinsilico.com/medical-physics/applications/tasmip-app.php.

[215] J. M. Boone, and J. A. Seibert, "An accurate method for computer-generating tungsten anode x-ray spectra from 30 to 140 kV," *Med Phys,* vol. 24, pp. 1661-1670, 1997.

[216] M. Gallis, "X-Ray Spectrum Model", Accessed on: 17 February 2021 [Online], Available: https://www.compadre.org/Repository/document/ServeFile.cfm?ID=13241&DocID=3775.

[217] D. M. Tucker, G. T. Barnes, and D. P. Chakraborty, "Semiempirical Model for Generating Tungsten Target X-Ray-Spectra," *Medical Physics,* vol. 18, pp. 211-218, 1991.

[218] A. M. Hernandez, and J. M. Boone, "Tungsten anode spectral model using interpolating cubic splines: Unfiltered x-ray spectra from 20 kV to 640 kV," *Medical Physics,* vol. 41, pp. 042101, 2014.

[219] Diamond Light Source (DLS), "Beamline I13-2 facilities and equipment", Accessed on: 17 February 2021 [Online], Available: http://www.diamond.ac.uk/Beamlines/Materials/I13/Diamond-Manchester_Imaging_Branchline/Facilities_and_equipment_Imaging.html.

[220] N. Wadeson, and M. Basham, "Savu: a Python-based, MPI framework for simultaneous processing of multiple, N-dimensional, large tomography datasets," *arXiv preprint arXiv:1610.08015*, 2016.

[221] Diamond Light Source (DLS), "Savu plugins API", Accessed on: 17 February 2021 [Online], Available: https://savu.readthedocs.io/en/latest/plugin_autosummary/#savu-plugins.

[222] M. Basham, "Savu dezinger filter", Accessed on: 17 February 2021 [Online], Available: https://github.com/DiamondLightSource/Savu/blob/8a7838fd0bea2b7ceec22f0d10e9033110608f35/savu/plugins/filters/dezinger.py.

[223] S. Titarenko, P. J. Withers, and A. Yagola, “An analytical formula for ring artefact suppression in X-ray tomography,” *Applied Mathematics Letters,* vol. 23, pp. 1489-1495, 2010.

[224] G. Ramachandran, and A. Lakshminarayanan, “Three-dimensional reconstruction from radiographs and electron micrographs: application of convolutions instead of Fourier transforms,” *Proceedings of the National Academy of Sciences,* vol. 68, pp. 2236-2240, 1971.

[225] W. van Aarle, W. J. Palenstijn, J. Cant *et al.*, “Fast and flexible X-ray tomography using the ASTRA toolbox,” *Optics Express,* vol. 24, pp. 25129-25147, 2016.

[226] W. van Aarle, W. J. Palenstijn, J. De Beenhouwer *et al.*, “The ASTRA Toolbox: A platform for advanced algorithm development in electron tomography,” *Ultramicroscopy,* vol. 157, pp. 35-47, 2015.

[227] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, “Performance improvements for iterative electron tomography reconstruction using graphics processing units (GPUs),” *Journal of Structural Biology,* vol. 176, pp. 250-253, 2011.

[228] USA department of defence, "Military communication system technical standards," Department of defense, Washington, D.C. 20301, 1969.

[229] J. Schindelin, I. Arganda-Carreras, E. Frise *et al.*, “Fiji: an open-source platform for biological-image analysis,” *Nature Methods,* vol. 9, pp. 676, 2012.

[230] J. Schindelin, C. T. Rueden, M. C. Hiner *et al.*, “The ImageJ ecosystem: An open platform for biomedical image analysis,” *Molecular reproduction and development,* vol. 82, pp. 518-529, 2015.

[231] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, “NIH Image to ImageJ: 25 years of image analysis,” *Nature Methods,* vol. 9, pp. 671, 2012.

[232] Fisher scientific, "Paraffin wax", Accessed on: 17 February 2021 [Online], Available: https://www.fishersci.co.uk/shop/products/paraffin-wax-pure-melting-point-about-60-c-pastillated/10395900.

[233] F. Marone, and M. Stampanoni, “Regridding reconstruction algorithm for real-time tomographic imaging,” *Journal of Synchrotron Radiation,* vol. 19, pp. 1029-1037, 2012.

[234] Olympus America Inc., "Olympus VS110 manual", Accessed on: 17 February 2021 [Online], Available: http://www.olympusamerica.com/files/seg_bio/vs110brochure.pdf.

[235] M. S. Brown, and C. J. Pisula, "Conformal deskewing of non-planar documents," presented at the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 998-1004.

[236] M. S. Brown, and W. B. Seales, "Document restoration using 3D shape: A general deskewing algorithm for arbitrarily warped documents," in Proceedings Eighth IEEE International Conference on Computer Vision (ICCV 2001), vol. 2, pp. 367-374.

[237] M. X. Sun, R. G. Yang, L. Yun *et al.*, "Geometric and photometric restoration of distorted documents," presented at the Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1, vol. 2, 2005, pp. 1117-1123.

[238] K. Pal, M. Terras, and T. Weyrich, “Interactive Exploration and Flattening of Deformed Historical Documents,” *Computer Graphics Forum,* vol. 32, pp. 327-334, 2013.

[239] M. S. Brown, M. X. Sun, R. G. Yang *et al.*, "Restoring 2D content from distorted documents," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 29, pp. 1904-1916, 2007.

[240] L. Zhang, Y. Zhang, and C. L. Tan, "An improved physically-based method for geometric restoration of distorted document images," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 30, pp. 728-734, 2008.

[241] D. Baum, N. Lindow, H. C. Hege *et al.*, "Revealing hidden text in rolled and folded papyri," *Applied Physics A,* vol. 123, pp. 171, 2017.

[242] I. Bukreeva, A. Mittone, A. Bravin *et al.*, "Virtual unrolling and deciphering of Herculaneum papyri by X-ray phase-contrast tomography," *Scientific Reports,* vol. 6, pp. 1-7, 2016.

[243] O. Samko, Y. K. Lai, D. Marshall *et al.*, "Virtual unrolling and information recovery from scanned scrolled historical documents," *Pattern Recognition,* vol. 47, pp. 248-259, 2014.

[244] D. Allegra, E. Ciliberto, P. Ciliberto *et al.*, "Virtual Unrolling Using X-Ray Computed Tomography," *2015 23rd European Signal Processing Conference (Eusipco)*, pp. 2864-2868, 2015.

[245] S. Stabile, F. Palermo, I. Bukreeva *et al.*, "A computational platform for the virtual unfolding of Herculaneum Papyri," *Scientific Reports,* vol. 11, pp. 1-11, 2021.

[246] M. D. R. Kok, J. B. Robinson, J. S. Weaving *et al.*, "Virtual unrolling of spirally-wound lithium-ion cells for correlative degradation studies and predictive fault detection," *Sustainable Energy & Fuels,* vol. 3, pp. 2972-2976, 2019.

[247] R. F. Ziesche, T. Arlt, D. P. Finegan *et al.*, "4D imaging of lithium-batteries using correlative neutron and X-ray tomography with a virtual unrolling technique," *Nat Commun,* vol. 11, pp. 777, 2020.

[248] P. M. Goggin, K. C. Zygalakis, R. O. C. Oreffo *et al.*, "High-resolution 3D imaging of osteocytes and computational modelling in mechanobiology," *European Cells and Materials (ECM),* vol. 31, pp. 264-295, 2016.

[249] P. Goggin, E. M. L. Ho, H. Gnaegi *et al.*, "Development of protocols for the first serial block-face scanning electron microscopy (SBF SEM) studies of bone tissue," *Bone,* vol. 131, pp. 115107, 2020.

[250] D. J. Griffiths, *Introduction to electrodynamics*, 4th ed.: Cambridge University Press, Cambridge, UK, 2017.

[251] S. Salvador, and P. Chan, "Toward accurate dynamic time warping in linear time and space," *Intelligent Data Analysis,* vol. 11, pp. 561-580, 2007.

[252] M. Spivak, *A comprehensive introduction to differential geometry*, 3rd ed.: Publish or perish, INC., Houston, Texas, USA, 1999.

[253] J. B. Tenenbaum, V. d. Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science,* vol. 290, pp. 2319-2323, 2000.

[254] S. Schaefer, T. McPhail, and J. Warren, "Image deformation using moving least squares," *Acm Transactions on Graphics,* vol. 25, pp. 533-540, 2006.

[255] ImageJ, "ImageJ jython API", Accessed on: 17 February 2021 [Online], Available: https://imagej.nih.gov/ij/developer/api/overview-summary.html.

[256] P. A. Yushkevich, J. Piven, H. C. Hazlett *et al.*, “User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability,” *Neuroimage,* vol. 31, pp. 1116-1128, 2006.

[257] T. W. Ridler, and S. Calvard, “Picture Thresholding Using an Iterative Selection Method,” *IEEE Transactions on Systems Man and Cybernetics,* vol. 8, pp. 630-632, 1978.

[258] D. Legland, I. Arganda-Carreras, and P. Andrey, “MorphoLibJ: integrated library and plugins for mathematical morphology with ImageJ,” *Bioinformatics,* vol. 32, pp. 3532-3534, 2016.

[259] M. Doube, M. M. Klosowski, I. Arganda-Carreras *et al.*, “BoneJ Free and extensible bone image analysis in ImageJ,” *Bone,* vol. 47, pp. 1076-1079, 2010.

[260] B. Bollobás, *Modern graph theory*: Springer Science & Business Media, New York, 2013.

[261] G. Lehmann, and D. Legland, "Efficient N-dimensional surface estimation using Crofton formula and run-length encoding". 2012. *hal-02811118*.

[262] F. Pedregosa, G. Varoquaux, A. Gramfort *et al.*, “Scikit-learn: Machine learning in Python,” *The Journal of machine Learning research,* vol. 12, pp. 2825-2830, 2011.

[263] B. Efron, “Censored Data and the Bootstrap,” *Journal of the American Statistical Association,* vol. 76, pp. 312-319, 1981.

[264] T. J. DiCiccio, and B. Efron, “Bootstrap confidence intervals,” *Statistical science,* vol. 11, pp. 189-212, 1996.

[265] B. Efron, and R. J. Tibshirani, *An introduction to the bootstrap*: CRC press, Abingdon, 1994.

[266] E. J. Pitman, “Significance tests which may be applied to samples from any populations,” *Supplement to the Journal of the Royal Statistical Society,* vol. 4, pp. 119-130, 1937.

[267] S. Raschka, “MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack,” *Journal of open source software,* vol. 3, pp. 638, 2018.

[268] J. Unpingco, *Python for probability, statistics, and machine learning*: Springer Nature, Switzerland, 2016.

[269] C. E. Bonferroni, “Teoria statistica delle classi e calcolo delle probabilita,” *Publ. R. Istit. Super. Sci. Econ. Commerc. Firenze,* vol. 8, pp. 1-62, 1936.

[270] E. W. Weisstein, "Bonferroni correction", Accessed on: 17 February 2021 [Online], Available: https://mathworld.wolfram.com/BonferroniCorrection.html.

[271] U. Erben, C. Loddenkemper, K. Doerfel *et al.*, “A guide to histomorphological evaluation of intestinal inflammation in mouse models,” *International Journal of Clinical and Experimental Pathology,* vol. 7, pp. 4557, 2014.

[272] A. E. Scott, D. M. Vasilescu, K. A. D. Seal *et al.*, “Three dimensional imaging of paraffin embedded human lung tissue samples by micro-computed tomography,” *PLoS One,* vol. 10, pp. 1-10, 2015.

[273] R. M. Spencer, J. D. Schroeter, and T. B. Martonen, “Computer simulations of lung airway structures using data-driven surface modeling techniques,” *Computers in Biology and Medicine,* vol. 31, pp. 499-511, 2001.

[274] M. Welter, K. Bartha, and H. Rieger, "Vascular remodelling of an arterio-venous blood vessel network during solid tumour growth," *Journal of Theoretical Biology,* vol. 259, pp. 405-422, 2009.

[275] S. Y. Ho, C. Y. Chao, H. L. Huang *et al.*, "NeurphologyJ: An automatic neuronal morphology quantification method and its application in pharmacological discovery," *Bmc Bioinformatics,* vol. 12, pp. 1-18, 2011.

[276] P. Wright, X. Fu, I. Sinclair *et al.*, "Ultra high resolution computed tomography of damage in notched carbon fiber-epoxy composites," *Journal of Composite Materials,* vol. 42, pp. 1993-2002, 2008.

[277] P. J. Schilling, B. P. R. Karedla, A. K. Tatiparthi *et al.*, "X-ray computed microtomography of internal damage in fiber reinforced polymer matrix composites," *Composites Science and Technology,* vol. 65, pp. 2071-2078, 2005.

[278] L. Wollatz, S. J. Johnston, P. M. Lackie *et al.*, "3D Histopathology - a Lung Tissue Segmentation Workflow for Microfocus X-ray-Computed Tomography Scans," *Journal of Digital Imaging,* vol. 30, pp. 772-781, 2017.

[279] T. Tuohimaa, M. Otendal, and H. M. Hertz, "Phase-contrast x-ray imaging with a liquid-metal-jet-anode microfocus source," *Applied Physics Letters,* vol. 91, pp. 074104, 2007.

[280] D. H. Larsson, P. A. C. Takman, U. Lundstrom *et al.*, "A 24 keV liquid-metal-jet x-ray source for biomedical applications," *Review of Scientific Instruments,* vol. 82, pp. 123701, 2011.

[281] Y. I. Nesterets, S. W. Wilkins, T. E. Gureyev *et al.*, "On the optimization of experimental parameters for x-ray in-line phase-contrast imaging," *Review of Scientific Instruments,* vol. 76, pp. 093706, 2005.

[282] D. Kalasova, T. Zikmund, L. Pina *et al.*, "Characterization of a laboratory-based X-ray computed nanotomography system for propagation-based method of phase contrast imaging," *IEEE Transactions on Instrumentation and Measurement,* vol. 69, pp. 1-1, 2019.

[283] A. Balles, S. Zabler, T. Ebensperger *et al.*, "Propagator based formalism for optimizing in-line phase contrast imaging in laboratory X-ray setups," *Rev Sci Instrum,* vol. 87, pp. 093707, 2016.

[284] B. Zeller-Plumhoff, T. Roose, G. F. Clough *et al.*, "Image-based modelling of skeletal muscle oxygenation," *Journal of The Royal Society Interface,* vol. 14, pp. 20160992, 2017.

[285] B. Zeller-Plumhoff, K. R. Daly, G. F. Clough *et al.*, "Investigation of microvascular morphological measures for skeletal muscle tissue oxygenation by image-based modelling in three dimensions," *J R Soc Interface,* vol. 14, pp. 20170635, 2017.

[286] M. G. Jones, A. Fabre, P. Schneider *et al.*, "Three-dimensional characterization of fibroblast foci in idiopathic pulmonary fibrosis," *JCI Insight,* vol. 1, pp. e86375, 2016.

[287] S. K. Robinson, J. J. Ramsden, J. Warner *et al.*, "Correlative 3D Imaging and Microfluidic Modelling of Human Pulmonary Lymphatics using Immunohistochemistry and High-resolution mu CT," *Scientific Reports,* vol. 9, pp. 1-14, 2019.

[288] O. L. Katsamenis, M. Olding, J. A. Warner *et al.*, "X-Ray Microcomputed Tomography for Nondestructive Three-Dimensional X-Ray Histology," *The American journal of pathology,* vol. 189, pp. 1608-1620, 2019.

[289] S. W. Wilkins, Y. I. Nesterets, T. E. Gureyev *et al.*, "On the evolution and relative merits of hard X-ray phase-contrast imaging methods," *Philos Trans A Math Phys Eng Sci,* vol. 372, pp. 20130021, 2014.

[290] M. Fratini, I. Bukreeva, G. Campi *et al.*, "Simultaneous submicrometric 3D imaging of the micro-vascular network and the neuronal system in a mouse spinal cord (vol 5, 8514, 2015)," *Scientific Reports,* vol. 5, 2015.

[291] B. Zeller-Plumhoff, T. Roose, O. L. Katsamenis *et al.*, "Phase contrast synchrotron radiation computed tomography of muscle spindles in the mouse soleus muscle," *Journal of Anatomy,* vol. 230, pp. 859-865, 2017.

[292] E. Pauwels, D. Van Loo, P. Cornillie *et al.*, "An exploratory study of contrast agents for soft tissue visualization by means of high resolution X-ray computed tomography imaging," *J Microsc,* vol. 250, pp. 21-31, 2013.

[293] B. D. Metscher, "MicroCT for developmental biology: a versatile tool for high-contrast 3D imaging at histological resolutions," *Dev Dyn,* vol. 238, pp. 632-40, 2009.

[294] D. L. Pham, C. Y. Xu, and J. L. Prince, "Current methods in medical image segmentation," *Annual Review of Biomedical Engineering,* vol. 2, pp. 315-337, 2000.

[295] R. M. Haralick, and L. G. Shapiro, "Image Segmentation Techniques," *Computer Vision Graphics and Image Processing,* vol. 29, pp. 100-132, 1985.

[296] A. Bali, and S. N. Singh, "A Review on the Strategies and Techniques of image Segmentation," *5th International Conference on Advanced Computing & Communication Technologies (ACCT 2015)*, pp. 113-120.

[297] S. Powell, V. A. Magnotta, H. Johnson *et al.*, "Registration and machine learning-based automated segmentation of subcortical and cerebellar brain structures," *Neuroimage,* vol. 39, pp. 238-247, 2008.

[298] M. Baiker, J. Milles, J. Dijkstra *et al.*, "Atlas-based whole-body segmentation of mice from low-contrast Micro-CT data," *Medical Image Analysis,* vol. 14, pp. 723-737, 2010.

[299] S. Bakas, M. Reyes, A. Jakab *et al.*, "Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the BRATS challenge," *arXiv preprint arXiv:1811.02629*, 2018.

[300] I. Arganda-Carreras, V. Kaynig, C. Rueden *et al.*, "Trainable Weka Segmentation: a machine learning tool for microscopy pixel classification," *Bioinformatics,* vol. 33, pp. 2424-2426, 2017.

[301] C. Sommer, C. Straehle, U. Kothe *et al.*, "Ilastik: Interactive Learning and Segmentation Toolkit," *2011 8th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 230-233, 2011.

[302] T. A. Ingham-Dempster, R. Rosser, B. M. Corfe *et al.*, "From cell to multi-crypt: Agent-based models of the human colon suggests novel processes of Field cancerisation," *Journal of Computational Science,* vol. 41, pp. 101066, 2020.

[303] J. A. Kiernan, "Formaldehyde, formalin, paraformaldehyde and glutaraldehyde: what they are and what they do," *Microscopy today,* vol. 8, pp. 8-13, 2000.

[304] E. R. Weibel, and R. A. Vidone, "Fixation of the lung by formalin steam in a controlled state of air inflation," *Am Rev Respir Dis,* vol. 84, pp. 856-61, 1961.

[305] B. D. Metscher, “MicroCT for comparative morphology: simple staining methods allow high-contrast 3D imaging of diverse non-mineralized animal tissues,” *BMC Physiol,* vol. 9, pp. 11, 2009.

[306] M. J. Lawson, O. L. Katsamenis, D. S. Chatelet *et al.*, "Immunofluorescence-guided segmentation of three-dimensional features in micro-computed tomography datasets of human lung tissue", Under revision, 2021.

[307] S. Mairhofer, S. Zappala, S. R. Tracy *et al.*, “RooTrak: Automated Recovery of Three-Dimensional Plant Root Architecture in Soil from X-Ray Microcomputed Tomography Images Using Visual Tracking,” *Plant Physiology,* vol. 158, pp. 561-569, 2012.

[308] S. R. Tracy, C. R. Black, J. A. Roberts *et al.*, “Quantifying the impact of soil compaction on root system architecture in tomato (Solanum lycopersicum) by X-ray micro-computed tomography,” *Annals of Botany,* vol. 110, pp. 511-519, 2012.

[309] C. Raymond, and L. Siddall, “Hidden Text: Imaging and reading an ancient tablet encased in an envelope,” *Buried History,* vol. 54, pp. 3-10, 2018.

[310] S. Klein, M. Staring, K. Murphy *et al.*, “Elastix: a toolbox for intensity-based medical image registration,” *IEEE transactions on medical imaging,* vol. 29, pp. 196-205, 2010.

[311] C. I. Gioumouxouzis, A. T. Chatzitaki, C. Karavasili *et al.*, “Controlled Release of 5-Fluorouracil from Alginate Beads Encapsulated in 3D Printed pH-Responsive Solid Dosage Forms,” *Aaps Pharmscitech,* vol. 19, pp. 3362-3375, 2018.

[312] G. K. Eleftheriadis, C. Ritzoulis, N. Bouropoulos *et al.*, “Unidirectional drug release from 3D printed mucoadhesive buccal films using FDM technology: In vitro and ex vivo evaluation,” *European Journal of Pharmaceutics and Biopharmaceutics,* vol. 144, pp. 180-192, 2019.

[313] C. A. Raymond, C. Haberl, J. Lundberg *et al.*, "Virtual recovery of text from an ancient inscribed lead scroll using neutron tomography," in ANSTO User Meeting 2019.