Edinburgh Gut Cell Atlas Models

Bill Hill [2] Derek Houghton [2] Mehran Sharghi [2] Michael Glinka [1] Mike Wicks [1] Shahida Din [3] Silvie Fexova [4]
Jana Eliasova [4] Irene Papatheodorou [4] David Adams [5] Albert Burger [2] Richard Baldock [1] Mark Arends [1]

[1]University of Edinburgh, Edinburgh, UK [2]Heriot-Watt University, Edinburgh, UK [3]Edinburgh IBD Unit, Western General Hospital, Edinburgh, UK

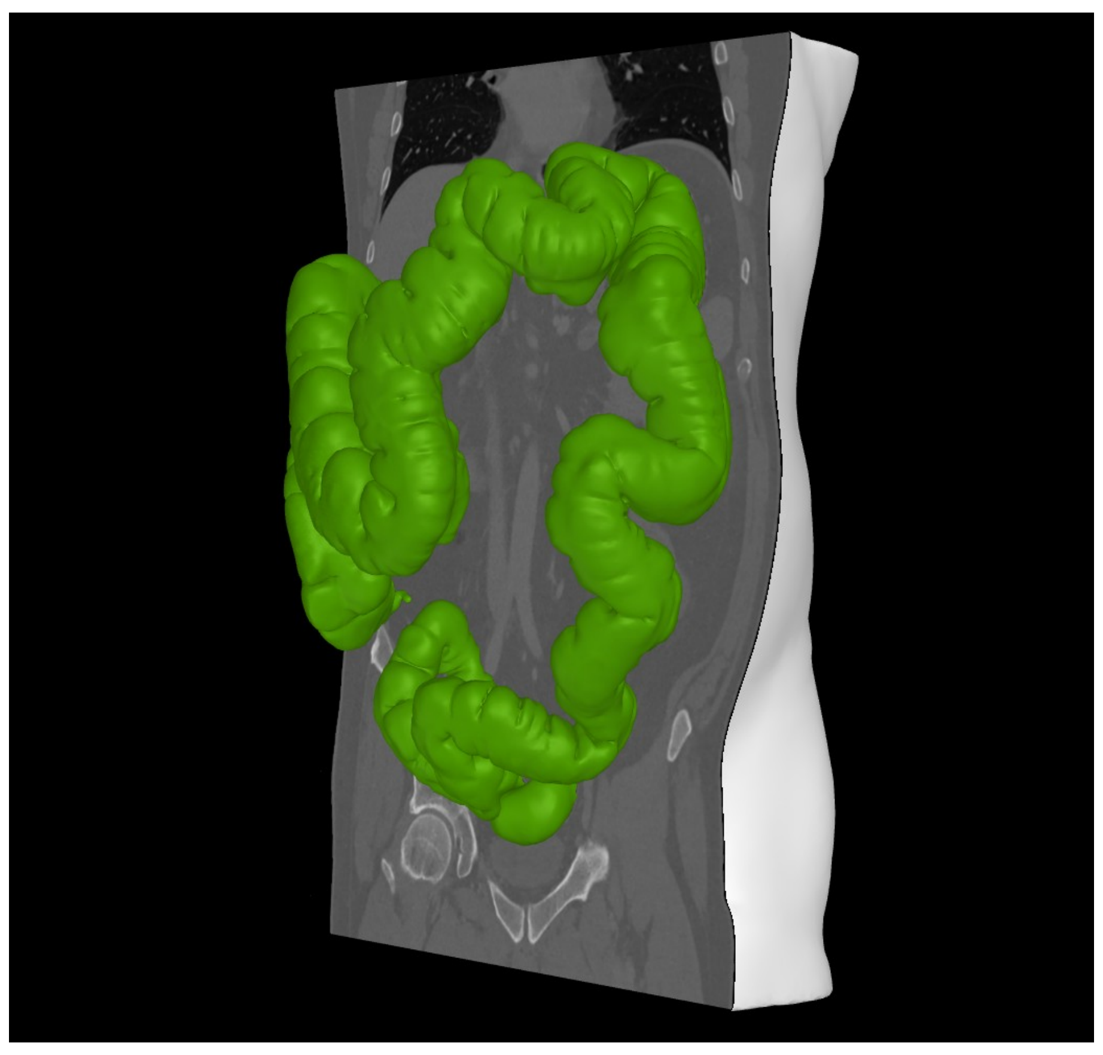
[4]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK [5]Wellcome Sanger Institute, Cambridge, UK

Introduction

We report a *common coordinate framework* (CCF) for spatial data in the context of the Helmsley Gut Cell Atlas (GCA) programme. In consultation with surgeons and clinical pathologists we developed a 1-D CCF capturing lineal distance in the gut from anus through the terminal ileum. In 2-D and 3-D this is represented as the centreline path through the gut and for sample location this is most useful metric for spatial comparison. Spatial location away from this path is captured using standard ontology terms for the gut layers, sub-layers, tissue and cell types as provided by the ASCT+B tables from the HCA/HuBMAP standards groups [1]. The 1D CCF has been mapped to the 2-D Anatomogram developed for the Single Cell Expression Atlas [2] and to four 3-D canonical models of the gut including both the male and female gut models from HuBMAP providing spatial integration and interoperability. Here we describe our models and web-based viewer.

Generation of models from patient data

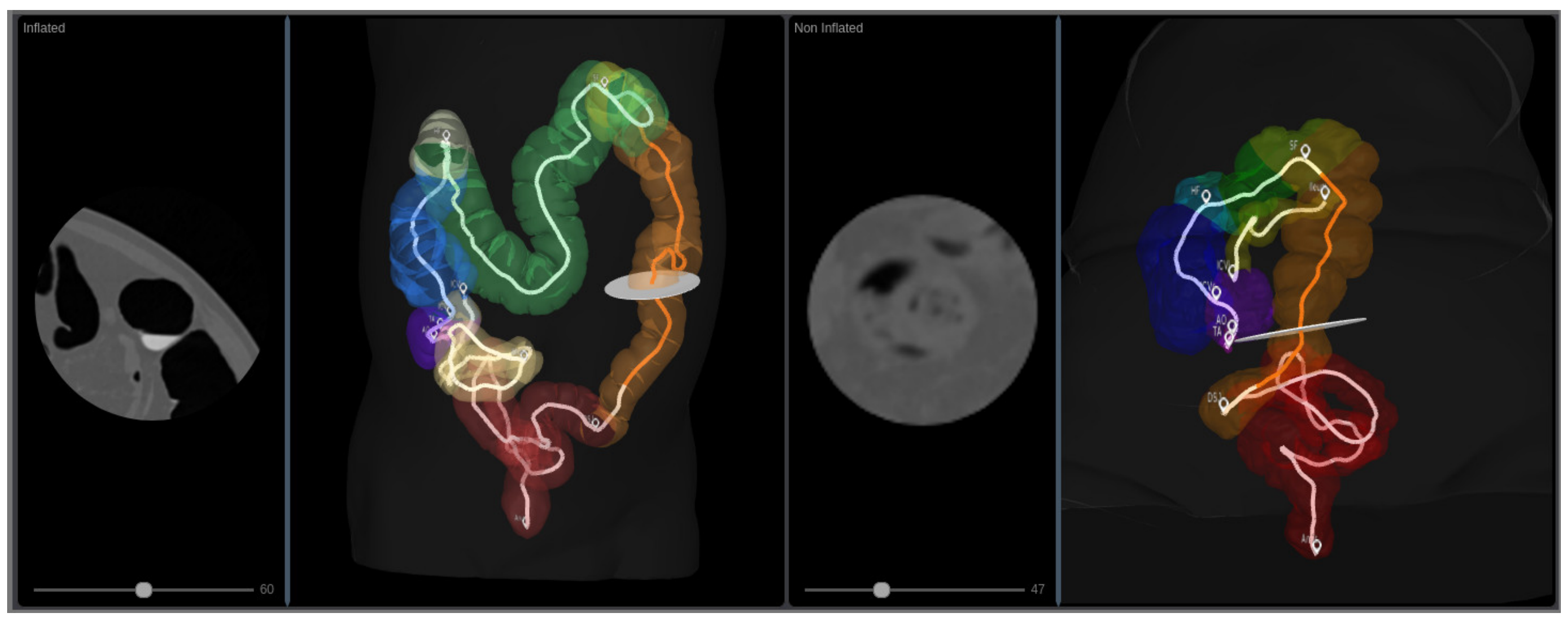
In addition to our 1D model, we have currently generated models from two anonymised patient CT image data sets; one in which the colon had been inflated and the other in which it had not. For the non-inflated dataset, due to lack of contrast, the images were first partially segmented using a U-Net [3] based convolutional neural network. Both models were segmented using a combination of region growing and morphological operators, with some manual intervention to separate touching tissues. The segmented image domains were used to extract midline paths using a path propagation algorithm in which a cost was minimised during propagation, with the cost based on a combination of distance from the end point and inverse distance from the lumen boundary. The ordered midline pixels/voxels were then used to fit least squares smooth B-spline curves. These curves form our midline path representation for 2 and 3D models. Most image processing and all path extraction was done using Woolz [4].



Colon segmented from patient CT.

Location and mapping in and between models

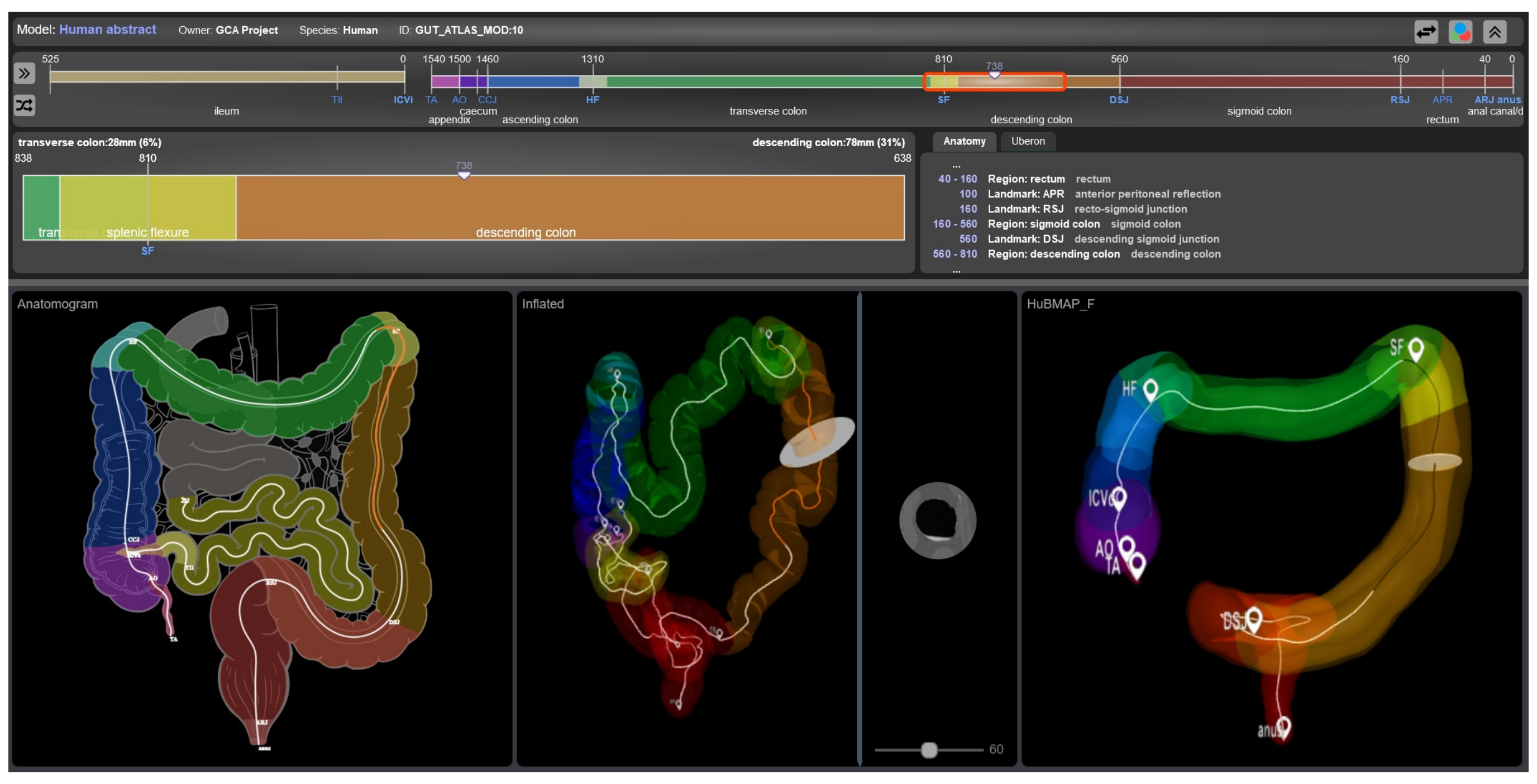
Locations within models are specified using the proportionate distances along midline paths between anatomical landmarks. Similarly, data are mapped between models using proportionate distances between landmarks along midline paths.
Locations off of midline paths (but within the domain of the colon or terminal ileum) may be mapped to the midline using a precomputed map of the nearest midline point, where the nearest evaluation is constrained to the appropriate sub-domain and the sub-domains are the portions of the colon or terminal ileum that are well defined anatomical regions.



Equivalent mapped positions in two models derived from patient CT images.

The Edinburgh model viewer

We have developed a web-based model viewer which will be used to display locations and mapped data in the context of our 1D and other mapped models. Here we show the 1D model above, a 2D anatomogram, a 3D model with inflated colon derived from an anonymised patient CT image and the HuBMAP VHF model. Where image data are associated with a model then image sections orthogonal to the midline paths may be viewed at all locations along the paths. Current locations and regions of interest are common to all linked models in the viewer so that selecting these in one model displays the corresponding locations and regions of interest in the other linked models. The viewer uses Fabric.js [5] and THREE.js [6] JavaScript libraries for interactive visualisation of 2 and 3D models. Details of anatomical structures, their locations and other configuration data unique to each model are obtained via an MVC based RESTful service layer with database connectivity developed as part of the project.



Model: Human abstract Owner: GCA Project Species: Human ID: GUT_ATLAS_MOD:10

transverse colon:28mm (6%)
descending colon:78mm (31%)

...
40 - 160 Region: rectum rectum
100 Landmark: APR anterior peritoneal reflection
160 Landmark: RSJ recto-sigmoid junction
160 - 560 Region: sigmoid colon sigmoid colon
560 Landmark: DSJ descending sigmoid junction
560 - 810 Region: descending colon descending colon
...

The Edinburgh model viewer showing 1-, 2- and 3-D models.

Conclusions and Future Work

We have developed a 1-D representation of the human large intestine and ileum. This is coupled with the HuBMAP ASCT+B standards to deliver a complete CCF for spatial annotation of patient samples and data including scRNA-Seq, histopathology from biopsy and surgical resection. Open source image analysis tools have been implemented to segment and map the 1-D framework onto 2-D (Anatomogram) and 3-D canonical models including the male and female model from HuBMAP. The tools will also allow mapping to individual patient 3-D clinical scans. We present a Javascript viewer providing browser-based model visualisation and exploration with real-time spatial cross-comparison between all models. This software in conjunction with the implementation of data "location cards" will be extended to visualise and query sample locations of data held in the Helmsley and HuBMAP data repositories. The model will be annotated with detailed clinical, histological, and transcriptomic data providing a platform for integrating and analysing healthy and disease tissue. In addition the underlying framework can also be mapped cross-species to allow comparative analysis.

References

[1] ASCT+B. https://hubmapconsortium.github.io/ccf/pages/ccf-anatomical-structures.html

[2] Expression atlas anatomogram. https://www.ebi.ac.uk/about/news/service-news/anatomograms-new-feature-single-cell-expression-atlas

[3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. pages 234–241, 2015.

[4] Woolz. https://github.com/ma-tech/Woolz

[5] Fabric.js. http://fabricjs.com/

[6] THREE.js. https://threejs.org/

Ackowledgements

We thank the Helmsley Charitable Trust for funding this project. We would like to acknowledge the work of Michael Cheeseman in aiding the development of this project. We acknowledge the work of Dr Stephen Glancy Consultant Radiologist who provided the CT images.

GUT CELL ATLAS
An Initiative Supported by the Helmsley Charitable Trust

THE LEONA M. AND HARRY B. HELMSLEY CHARITABLE TRUST

THE UNIVERSITY OF EDINBURGH

HERIOT WATT UNIVERSITY